The Wistar Institute of Anatomy and Biology

Annual Progress Report: 2009 Formula Grant

Reporting Period

July 1, 2010 – June 30, 2011

Formula Grant Overview

The Wistar Institute of Anatomy and Biology received $1,778,081 in formula funds for the grant award period January 1, 2010 through June 30, 2011. Accomplishments for the reporting period are described below.

Research Project 1: Project Title and Purpose

Modeling the Epigenetic Changes in Alternative Promoters of Cancer Genes - We propose a combination of computational, statistical and high-throughput experimental approaches to better characterize the use of alternative promoters of human genes in cancer cells, which will lead to improved biomarkers and drug discovery, one of the major challenges of cancer research. We will (1) develop an information resource of alternative promoters of cancer genes; (2) experimentally assess the active or repressed state of alternative promoters in different cancer cells relative to normal cells, using massive parallel sequencing approaches; and (3) develop statistical models to define unique and common epigenetic signatures across the promoters in different cancer cells.

Duration of Project

1/1/2010 - 6/30/2011

Project Overview

Cancer results from the accumulation of genetic and epigenetic events arising over a long period of time. Despite the fact that many important advances have been made in understanding the differences between normal and cancer cells, achieving a high level of specificity and efficacy in cancer diagnosis and treatment still remains a great challenge. A key insight provided by the use of high-throughput molecular technologies, such as microarrays, is the complexity and heterogeneity of cancer. A phenomenon that adds an important dimension to this complexity is the presence of alternative first exons and promoters within a gene. Sparing neighboring healthy tissues is an important consideration when designing therapies that suppress cancer cells. In a specific cell-type or tissue-type, only a subset of genes (or gene isoforms) is active, generally determined by a piece of DNA, known as the promoter, which lies upstream of the corresponding gene. Recent annotations of the human genome suggest that up to 70% of protein-coding genes make one or more alternative forms. Notably, the incidence of multiple-promoters among cancer genes is much higher than the proportion among the remainder of genes, and there

is growing evidence linking aberrant use of multiple promoters and cancer formation ─ several oncogenes and tumor-suppressor genes are already known to have multiple promoters, and moreover, it is known that the aberrant use of one promoter over another in some of these genes is directly linked to cancerous cell growth.

We hypothesize that alternative promoters in a gene locus are aberrantly used (either activated or silenced) in cancer genomes. Our rationale is that regional epigenetic modifications in alternative promoters are tissue-specific and that disruptions to these processes occur during cancer initiation and progression. These aberrant events may lead to the eventual silencing or activation of critical promoters of cancer genes. We propose a combination of computational, statistical and high-throughput experimental approaches to better characterize the use of alternative promoters of human genes in cancer cells, which will lead to improved biomarkers and drug discovery, one of the major challenges of cancer research. Specifically we will (1) develop an information resource of alternative promoters of cancer genes, (2) experimentally assess the active or repressed state of alternative promoters in different cancer cells relative to normal cells using massive parallel sequencing approaches, and (3) develop statistical models to define unique and common epigenetic signatures across the promoters in different cancer cells.

Principal Investigator

Ramana V. Davuluri, PhD
Associate Professor
The Wistar Institute
3601 Spruce Street
Philadelphia, PA 19104

Other Participating Researchers

Louise Showe, PhD, Hyunsoo Kim, PhD, Yingtao Bi, PhD, Sharmistha Pal, PhD, Ravi Gupta, PhD - employed by The Wistar Institute

Expected Research Outcomes and Benefits

Two important questions of immediate attention that will be addressed by this research proposal are (i) which alternative promoters of human genes are preferentially silenced or activated in cancer cells, and (ii) what are the differences in the epigenetic modifications across the alternative promoters that cause differential use of the corresponding promoters in different tissues? The successful completion of this research will provide (i) an information resource of the alternative promoters of cancer genes, and (ii) a comprehensive survey of tissue-specific and cancer-specific alternative promoters. The combined use of bioinformatics and massive parallel sequencing technologies will help better understand the epigenetic mechanisms that regulate alternative promoters in a tissue-specific manner and provide clues to their aberrant regulation in cancer cells. We believe this research will identify a large number of novel targets for the diagnosis and treatment of different cancers. For example, imagine a drug that suppresses expression of a specific form of a human gene called CYP19 promoter II, which is aberrantly activated in breast cancer, while sparing the other five promoters that control expression of this

gene in healthy tissues. By targeting the specific problematic promoter, a more effective treatment with fewer side effects can be administered. It is not hard to imagine how specific targeting of individual promoters could lead to more accurate and sensitive diagnoses as well.

Summary of Research Completed

Further Development of CancerPromDb database: Integrative analysis of exon-array data-sets revealed that isoform-level expression profiles provide better cancer signatures than gene-level expression profiles.

Although genome-wide, gene-level expression changes have been studied extensively, the dynamic changes at the transcript/isoform level between normal and cancer cells remain largely unexplored. We hypothesized that the isoform-level expression profiles can better discriminate normal and cancer cells than gene-level expression profiles. To test this, Affymetrix exon-array datasets, comprising 160 cell lines of normal or tumor tissue origins, were analyzed. Hierarchical clustering, based on isoform-level expressions, effectively grouped the normal and oncogenic cell lines with a virtually perfect homogeneity grouping rate (97.5%), regardless of the tissue origins of the cell lines. However, at the gene level, this rate was much lower and was 75% at best (Figure 1). Statistical analyses at the isoform level between groups of cancer and normal samples revealed the existence of numerous genes having differentially expressed isoforms, which otherwise were not significant at the gene level. The opposite expression pattern of two *TPM4* isoforms observed in the cell lines was confirmed in breast cancer patients' tumor tissues over normal neighboring breast tissues (Figure 2). Finally, pathway-based enrichment analysis revealed "protein ubiquitination," "purine metabolism," and "breast cancer regulation by stathmin1" as significantly enriched canonical pathways with the gene set that showed differential expression at the isoform level but not at the gene level. In summary, a common, isoform-level signature exists that can be used to effectively discriminate cancer and non-cancer cell lines and provide potentially important biological implications for a better understanding of common defining characteristics of tumors of varied types. A manuscript describing these results has been submitted.

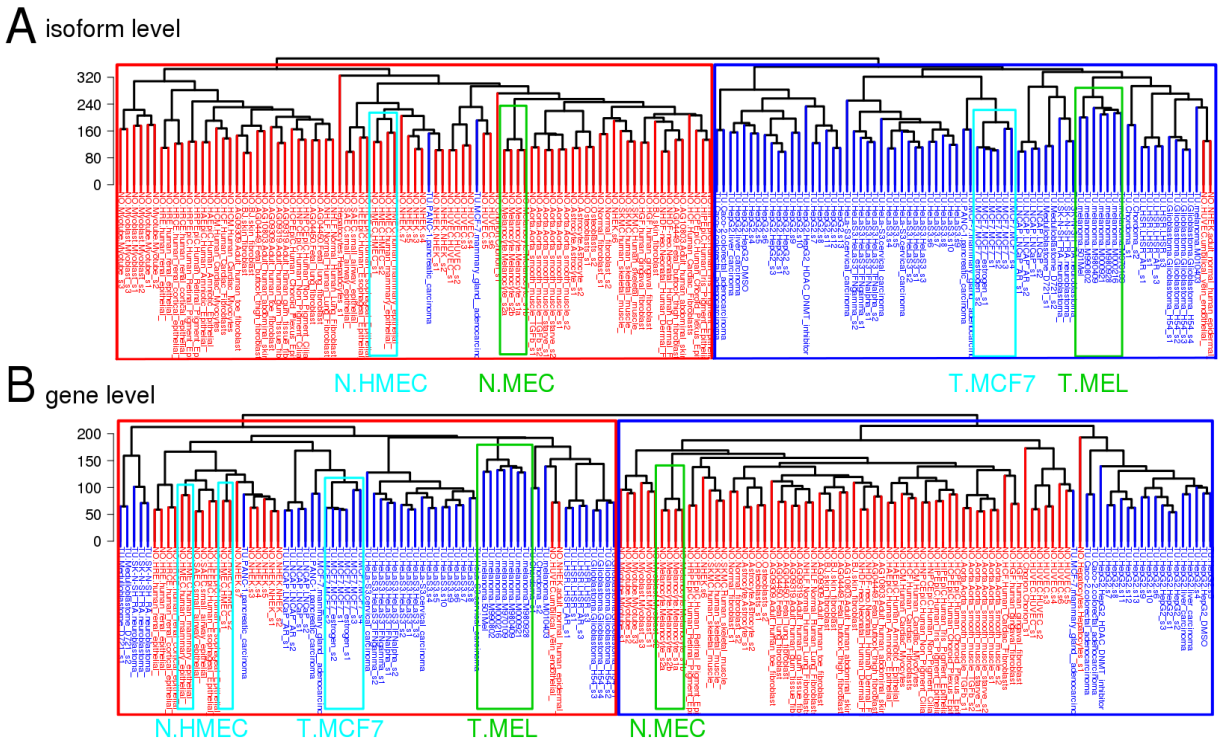

Figure 1: Hierarchical clustering of 160 (73 oncogenic – blue color; 87 non-oncogenic – red color) cell-lines of various tissue origins, using expression estimates at (**A**) isoform-level (87,345 transcripts) and (**B**) gene-level (27,063 genes). The top 76% number of genes/transcripts that showed the highest coefficient of variance (CV) of expression profile across all the samples was used for the clustering. The normal melanocyte (N.MEL) and melanoma (T.MEL) cell-lines were clustered together and separated into normal and oncogenic groups in dendrogram A; whereas in dendrogram B, samples belonging to “N MEC” were not grouped together and clustered with the overall oncogenic group.

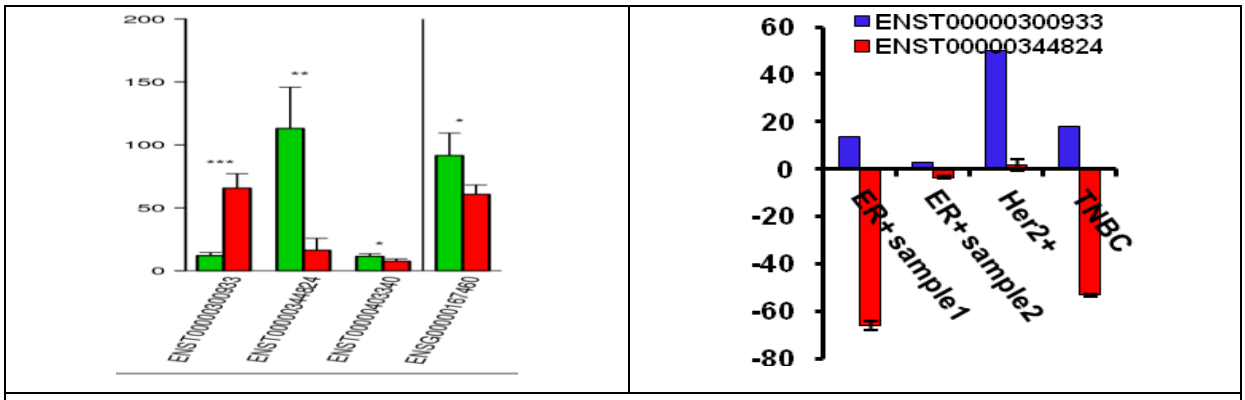

Figure 2: Expression of *TPM4* and its transcript variants in HMEC (N) and MCF7 (T) cell-lines. The two transcript variants of TPM4 showed opposing expression patterns between HMEC normal and MCF7 breast cancer cell lines, in exon-array datasets (Left panel). Validation results via RT-PCR experiment confirmed the opposing expression patterns of the two major transcript variants (ENST00000300933 and ENST00000344824) in various human breast cancer tissue subtypes compared to the surrounding normal matched normal breast tissues (Right panel). Here the subtypes are ER+: estrogen receptor positive, Her2+: Her2 gene positive, TNBC: Triple negative breast cancer

Research Project 2: Project Title and Purpose

Characterizing Mechanisms of Transcriptional Activation Using Live Cell Imaging - Numerous diseases, including cancer, are caused by aberrations in transcriptional regulation. Gene activation requires the tight coordination of regulatory factors at the site of transcription. We propose to study the recruitment dynamics of these factors at a transcription site in single living cells in order to produce a high-resolution map of their recruitment order. A comprehensive understanding of transcriptional regulation is important for the development of new strategies to cure human diseases.

Duration of Project

1/1/2010 - 6/30/2011

Project Overview

We will investigate the recruitment dynamics of transcriptional regulatory factors to a transcription site in single living cells during activation because the exact timing and order of recruitment is not known and cannot be fully determined using biochemical techniques alone. We have developed a method to visualize the activation of a chromatinized transcription site in single living cells with high temporal and spatial resolution. We plan to use it to better understand the temporal mechanisms regulating transcription initiation. It is our goal to use this methodology to develop a high-resolution regulatory pathway map of transcriptional activation, much like what has been done in the signal transduction field. These studies will also provide new insight into the kinetic control of human transcription activation pathways.

In order to achieve these goals, we propose the following specific aims:

Aim 1. Interrogate by shRNA depletion the requirement of known regulatory factors on (i) chromatin decondensation, (ii) pre-initiation complex assembly and (iii) RNA synthesis, using kinetic, live cell imaging of single cells.

Aim 2. Use p53 as a model transcriptional activator to interrogate the timing of its activation kinetics using live cell imaging analysis.

Principal Investigator

Susan M. Janicki, PhD
Assistant Professor
The Wistar Institute
3601 Spruce Street
Philadelphia, PA 19104

Other Participating Researchers

Ilona Rafalska-Metcalf, PhD, Alyshia Newhart, BS - employed by The Wistar Institute

Expected Research Outcomes and Benefits

We will use shRNA depletion to evaluate the requirement of known regulatory factors on chromatin decondensation, pre-initiation complex assembly and RNA synthesis. We will also determine the kinetic and spatial dynamics of post-translational modifications (PTM) and PTM binding proteins during transcription activation and mitosis. In order to expand the use of the live cell imaging methodology, we also propose to use p53 as a model transcriptional activator and to interrogate the timing of its activation kinetics. These studies are significant because they cannot be done using other existing techniques that average effects in cell populations. Additionally, they will provide a portrait of unperturbed gene regulation in single cells. A comprehensive understanding of transcriptional regulation is important for the development of new strategies to cure human diseases.

Summary of Research Completed

An important component of using single cell live cell imaging to study mechanisms of gene regulation is the continual characterization of the factors which function at the specific reporter site being studied in these experiments. Our previous results demonstrated that p53 is not a very strong activator, which has made it difficult to study regulatory factor recruitment when it is used. However, we have discovered that this site is regulated by Daxx and ATRX (alpha-thalassemia/mental retardation, X-linked), which were recently identified as regulators of replication independent histone H3.3 chromatin assembly. The results described below detail our characterization of their regulatory functions.

Results

The histone H3.3 chaperone, Daxx, but not HIRA, is recruited to a transgene array.

HIRA was the first histone H3.3 chaperone identified. To determine whether it is recruited to the activated transgene array in human U2OS 2-6-3 cells, we expressed yellow fluorescent protein (YFP)-tagged HIRA and examined its localization in relation to the transcriptional activator, Cherry-tTA-ER, a modified form of the tetracycline-regulated transcriptional activator (tTA). HIRA-YFP was not recruited (Fig. 1, panels a-d). Therefore, we used a HIRA antibody to examine the endogenous factor in case the YFP tag interfered with recruitment. Recruitment of endogenous HIRA was also not detected (Fig. 1, panels e-h).

In contrast to HIRA, YFP-tagged Daxx was enriched at both the inactive and activated transgene array where it co-localized with the Cherry-tagged lac repressor and the transcriptional activator, Cherry-tTA-ER, as shown in the intensity profiles of the merged images (Fig. 1, panels i-p). YFP-Daxx was detected at 74.5 ± 6.3% of inactive sites and 100% of activated sites, indicating that the signal which recruits it is present at the majority of inactive arrays but is produced at all sites during activation. These results also indicate that this transgene array possesses characteristics which cause it to be recognized as heterochromatin by the histone H3.3 chaperones.

ATRX is required for transgene array silencing.

Daxx forms a complex with the chromatin-remodeling factor, ATRX, and both factors are required for H3.3 incorporation into telomeres and centromeres. Interestingly, U2OS cells do not express ATRX (Fig. 2A). However, when ATRX-YFP was expressed in U2OS 2-6-3 cells, it co-localized with Cherry-lac repressor at the inactive array, suggesting that it functions at the site (data not shown). Previously, we showed that activator binding induces significant chromatin decondensation; the pixel area of the array bound by Cherry-tTA-ER is ~5.2X larger than the inactive array marked by Cherry-lac repressor (Fig. 2C). However, in ATRX-YFP expressing cells, the activator-bound array remained highly condensed (~1.7X the area of the inactive array) and the cyan fluorescent protein with peroxisomal targeting signal SKL (CFP-SKL) protein was barely detected in the peroxisomes in the cytoplasm (Fig. 2B, panels a-d and 2C). This indicates that ATRX functions to transcriptionally repress the transgene array.

To determine whether the catalytic activity of ATRX is required for this repression, we expressed ATRX(K1600R)-YFP, which contains a mutation commonly used to abolish the ATPase activity of SWI/SNF proteins. The array decondensation and CFP-SKL were higher in cells expressing this mutant, thus indicating that the helicase activity of ATRX is required for transcriptional repression (Fig. 2B, panels e-d and 2C). Interestingly, the area occupied by the activator-bound array in ATRX(K1600R) expressing cells was only ~3X larger than the inactive site (Fig. 2C). This suggests that there are other mechanisms, in addition to its helicase activity, through which ATRX inhibits chromatin decondensation. RNA pol II was enriched at the activator-bound array in ATRX(K1600R) but not WT-ATRX expressing cells, indicating that ATRX represses transcription, at least in part, by inhibiting RNA pol II recruitment (Fig. 2B, panels i-p).

Materials and Methods

Plasmid Construction

YFP-Daxx-C1 was moved from GFP-DAXX-C1 (gift of Dr. Gerd Maul, The Wistar Institute) (BamHI). HIRA-YFP-N1 was made by cloning the PCR product (gift of Dr. Genevieve Almouzni, CNRS Institute Curie), which included a 23-aa linker (XhoI/SacII). ATRX-YFP was made by cloning a PCR product of full length ATRX (gift from Dr. David Picketts, Ottawa Hospital Research Institute) into pLU (XbaI/SalI) then replacing the 3' end with a fusion between the 3' end (BamHI/AgeI) and YFP-N1 (BamHI/SalI). pLU-ATRX(K1600R)-YFP (6) was made by introducing a point mutation using Quickchange kit (Stratagene).

Image Analysis

Intensity profiles for merged pictures were generated using SimplePCI software (Compix). Pixel areas of the transgene array were made by manually selecting regions of interest (ROIs) using ImageJ software.

Immunofluorescence

For immunofluorescence, 2-6-3 cells were transfected overnight with the activator, Cherry-tTA-ER. Cells were first pre-extracted with 0.5 % Triton X-100 in CSK buffer (10 mM PIPES, pH 7.0, 100 mM NaCl, 300 mM sucrose, 3 mM $MgCl_2$ with freshly added protease inhibitors) for 5-10 minutes and then fixed in 3% formaldehyde in 1× PBS for 15 minutes, both steps at room temperature (RT). Cells were blocked with 3% BSA in 1× PBS for 1 hour at RT and incubated with 1˚ Abs diluted in blocking buffer. The following Abs were used: RNA pol II 4H8 (Covance, 1:500) and HIRA (gift of Peter Adams, PhD, University of Glasgow). 2˚ antibodies were Alexa Fluor AF488 conjugated (Invitrogen) and diluted 1:3000 in 1× PBS.

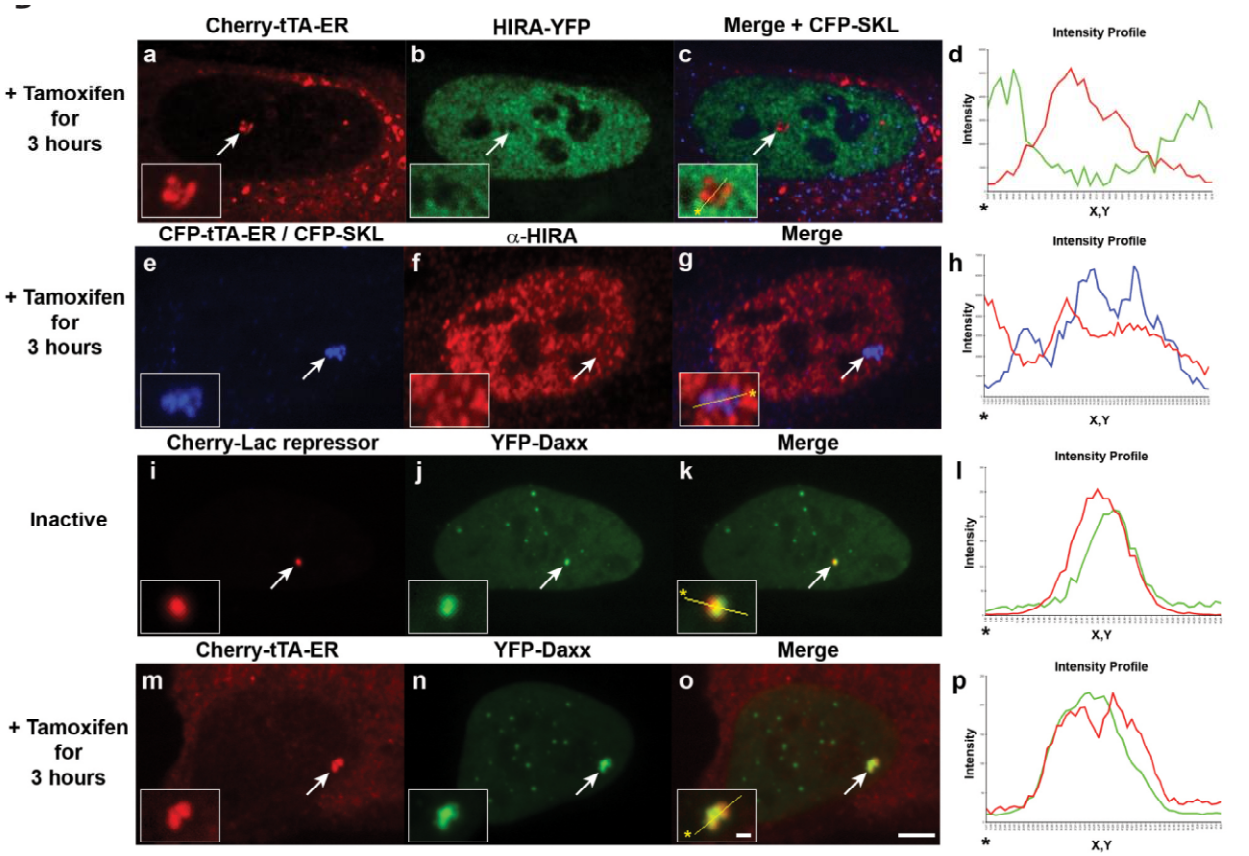

Figure 1: The histone H3.3 chaperone, Daxx, but not HIRA, is recruited to a transgene array.

Localization of HIRA-YFP (**a-d**) and endogenous HIRA labeled with α-HIRA antibody (**e-h**) in relation to the Cherry and CFP tagged activator, tTA-ER, at the activated transgene array. YFP-Daxx is enriched at the inactive array, marked by Cherry-lac repressor (**i-l**), and activated array, marked by Cherry-tTA-ER (**m-p**). Yellow lines in enlarged merge insets show the path through which the red, green and blue intensities were measured in the intensity profiles to the right of the images. Asterisks mark the start of the line. Scale bar=5 µm. Scale bars in the enlarged inset=1 µm.

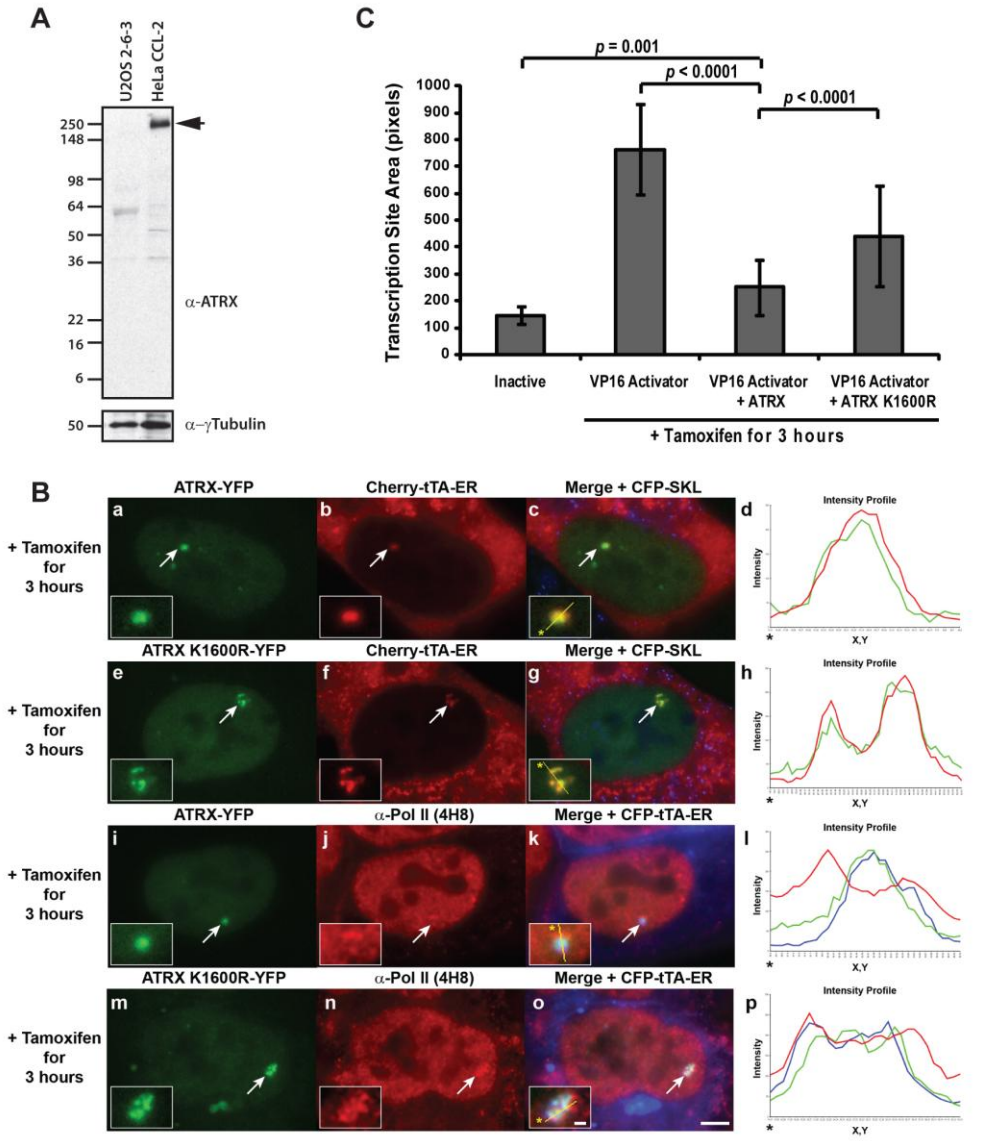

Figure 2: ATRX is required for transgene array silencing.

(A) Western blot analysis of endogenous ATRX levels in U2OS 2-6-3 and HeLa cells showing that ATRX is not expressed in the U2OS-derived cell line. γ-Tubulin is used as a loading control. (**B**) Wild type ATRX-YFP (**a-d**) and ATRX(K1600R)-YFP (**e-h**) are enriched at the activated array marked by Cherry-tTA-ER. RNA pol II, visualized with 4H8 antibodies, is not recruited to the activated array in the presence of wild-type ATRX-YFP (**i-l**) but it is enriched in cells expressing ATRX(K1600R)-YFP (**m-p**). Yellow lines in enlarged merge insets show the path through which the red, green and blue intensities were measured in the intensity profiles to the right of the images. Asterisks mark the start of the line. Scale bar=5 µm. Scale bars in the enlarged inset=1 µm. (**C**) Measurement of the pixel area of the inactive transgene array, marked by Cherry-lac repressor and the activated array marked by the activator, Cherry-tTA-ER, in U2OS 2-6-3 cells expressing either YFP, wild type ATRX-YFP and ATRX(K1600R)-YFP; SDs (in the form of error bars) are shown in the graph.

Research Project 3: Project Title and Purpose

Elucidation of the Integrator Composition and Function - This project is focused at elucidating the role of the Integrator, a complex that was discovered by research investigators at The Wistar Institute, in the control of transcription and cellular proliferation. We showed that Integrator is critical for transcription and processing of small nuclear ribonucleic acid (RNA) genes. Elucidation of the mechanism through which small nuclear RNAs are transcribed and processed to mature RNAs is of seminal importance in understanding regulation of gene expression and growth control. Aberrant regulation of small nuclear RNA processing could result in defective messenger RNA processing through deregulation of gene splicing and may lead to cellular transformation and oncogenesis.

Duration of Project

1/1/2010 - 6/30/2011

Project Overview

Small nuclear ribonucleic acids (snRNAs) are components of the spliceosome, a complex assembly of ribonucleoparticles involved in pre-micro RNA (mRNA) processing. The majority of snRNAs (i.e., U1 to U5) are transcribed by RNA polymerase II (RNAPII) to yield short non-polyadenylated 3'-extended precursors. These precursors are then exported to the cytoplasm for further 3' trimming and incorporation into the small nuclear ribonucleoparticles (snRNPs). The formation of proper snRNA precursors depends on a *cis*-acting sequence, named the 3' box, located 9-19 nucleotides downstream of the 3' end of the mature snRNA and on the presence of an snRNA compatible promoter at the 5' end of the gene. Furthermore, the C-terminal domain (CTD) of the largest subunit of RNAPII has been shown to play a crucial role in snRNA 3'-end processing. The identity of the factor(s) that associate with the CTD and result in the processing of the snRNA genes remained elusive for over two decades. Recently, we isolated a multi-protein complex termed the Integrator that associates with the CTD of RNAPII and mediates the 3'-end processing of snRNAs. This project is focused on characterization of the Integrator polypeptide composition and functional elucidation of its mechanism of action.

Specific Aims:

1. *Analysis of subunits of the Integrator and reconstitution of the complex using recombinant components.* Several of the proteins in the Integrator complex do not correspond to genes whose function or expression patterns have been described. Biochemical and immunological analyses will be performed to determine the role of these proteins in regulation of small nuclear RNA transcription and processing.

2. *Identification of the subunit(s) of the Integrator that mediate association with the CTD.* We will examine the role of the CTD phosphorylation in modulating the association of the Integrator complex and RNAPII.

Principal Investigator

Ramin Shiekhattar, PhD
Professor
The Wistar Institute
3601 Spruce Street
Philadelphia, PA 19104

Other Participating Researchers

David Baillat, PhD, Clement Carre, PhD, Kaiping Yan, PhD, Matteo Cesaroni, PhD, Alessandro Gardini, PhD, Derly Acevedo, Fan Lai, PhD – employed by The Wistar Institute

Expected Research Outcomes and Benefits

This project will lead to a greater understanding of the mechanism by which eukaryotic gene expression is regulated. A few years ago, we discovered a novel complex termed the Integrator that mediates the transcription and processing of small nuclear ribonucleic acid (RNA) genes. Small nuclear RNA genes are critical components of the machinery that regulates gene splicing of nearly all human messenger RNAs. Therefore, understanding the scope and mode of action of Integrator will provide us with a greater insight into the precise mechanism by which genes are regulated. This knowledge will be important as aberrant regulation of both transcription and splicing could result in development of cancer.

Summary of Research Completed

To directly assess the role of Integrator complex and its individual subunits in snRNA processing, we utilized a reporter construct which allows the detection of the processed and unprocessed transcript using RNAase protection. As shown in Figure 1, depletion of INTs 1, 2, 3, 4, 9 and 11 results in a defect of snRNA processing. We have generated knock-down vectors for the rest of the Integrator subunits and are in the process of systematic analysis of all subunits. It will also be important to know the subunit composition of the complex following the knock-down of individual subunits. We have generated a number of epitope-tagged cell lines using different Integrator subunits where Integrator could be affinity-purified. Such an analysis will not only begin to reveal the subunit architecture of the complex, but also will provide valuable information regarding individual subunits whose presence is required to maintain the integrity of a large multi-protein complex. Such information will also aid in precise explanation of the functional defects following the knock-down of the individual subunits. For example, we will be able to discriminate between functional defects that result from the loss of the catalytic subunit INTS11, which may reveal a functional requirement for an individual subunit of the complex. Taken together, these studies combined with *in vitro* reconstitution should provide us with great insight into the structure and function of the Integrator complex.

Integrator displays a role in transcription beyond that of RNA processing

To further dissect the Integrator role in snRNA transcription as well as RNA processing, we designed a reporter plasmid comprising the entire U2 snRNA gene (promoter, U2 snRNA sequence and 3´ box) fused at its 3´end with the renilla luciferase (pU2G-RL in Figure 4a) and another comprising the U2 snRNA promoter directly fused to the renilla luciferase coding sequence (pU2P-RL, in Figure 2a). The pU2G-RL should express luciferase only when the processing is impaired, while pU2P-RL should indicate the transcription rate of the U2 promoter in the different conditions used. The shRNA mediated knock-down of INTS1 and INTS11, known to impair U2 snRNA 3´end processing, resulted in an important increase of luciferase activity (Figure 2b, left panel). Unexpectedly, knock-down of INTS1 and INTS11 also resulted in the increase of transcription from the U2P construct (Figure 2a, right panel: white columns). Ratio of luciferase activity from U2G over U2P is still in favour of a processing defect over just a transcriptional activation (Figure 2a, right panel). Nevertheless, the increase in luciferase from pU2P-RL reporter seen following Integrator knock-down is suggestive of a role for Integrator in transcription of the U2 promoter distinct from its role in 3´-box-mediated snRNA processing. Importantly, we find a similar effect following the knock-down of most Integrator subunits (Figure 3). It is formally possible that Integrator has a greater role in RNA processing. However, any additional role for Integrator in RNA processing would lead to decreased luciferase activity, rather than an enhanced luciferase expression.

One exciting hypothesis concerns the possible role for Integrator in suppressing aberrant transcripts generated due to read-through transcripts. Normally, Integrator suppresses such transcripts and its depletion leads to accumulation of such aberrant transcripts. In the following section we will present further data that is consistent with a role for Integrator in the regulation of a large number of RNAPII-mediated genes.

Purification of Integrator from the chromatin fraction and global analysis of its genome occupancy

We initially purified the Integrator complex using a nuclear extract of a stably-expressing Flag-INTS10 293T-derived cell line. In order to gain further insight into the composition of the Integrator complex that is associated with the chromatin fraction, we purified Integrator using a new protocol that was described for purification of chromatin bound factors. Interestingly, we found that not only was Integrator quite enriched in the chromatin fraction as expected from its stable association with RNAPII, but also we were able to obtain a cleaner purification of the complex using chromatin enriched fractions (also called nuclear pellet). Interestingly, the purification enabled us to avoid the excess of over-expressed flag-tagged protein. Flag-INTS10 is now eluted at nearly stoichiometric ratio with the other subunits (Figure 4a). Mass spectrometric sequencing of this chromatin bound Integrator complex confirmed the subunit composition we had previously established for the Integrator (Figure 4b). The enriched association of the Integrator with the chromatin persuaded us to analyze its global genome occupancy using chromatin immunopreciptation followed by high throughput Solexa sequencing. We anticipate a greater role for this complex on RNAPII-mediated transcription beyond that of snRNA transcription and RNA processing.

The close association of Integrator and chromatin described above and the recent identification of Integrator in large scale purifications of the spliceosome or the mRNA 3´end processing machinery suggested a broader function for the Integrator complex in RNAPII transcription. To investigate this possibility, we designed a ChIP-sequencing experiment. A stable 293T cell line stably expressing INTS11 fused to two Hemagglutinin (HA) epitopes was established and tested for enrichment of U1 snRNA gene 3´end as a control (Figure 5a).

The ChIP material was then used for high throughput sequencing using Solexa technology. Moreover, we used an HA antibody ChIP from an untagged 293T cell line as control and subjected those to Solexa sequencing as well. The analysis of high throughput data is performed in collaboration with the Bioinformatic facility at the Wistar institute headed by Dr. Ramana Davuluri. We obtained 4,379,099 unambiguous reads from HA-Int11 cell line and 4,815,758 unambiguous reads from parental untagged control cell line. Following clustering analysis and subtraction of signals (peaks) that did not reveal a statistically significant difference between HA-Int11 cell line and the parental control cell line, we were left with 2,321 specific HA-Int11 peaks. Eighty-five percent of these peaks (1,973) displayed a close association with the transcriptional start sites of RNAPII-dependent genes (Figure 5b and c). Gene ontology analysis indicated that a large number of these genes (greater than 50) are involved in establishment or maintenance of chromatin architecture (Figure 5c).

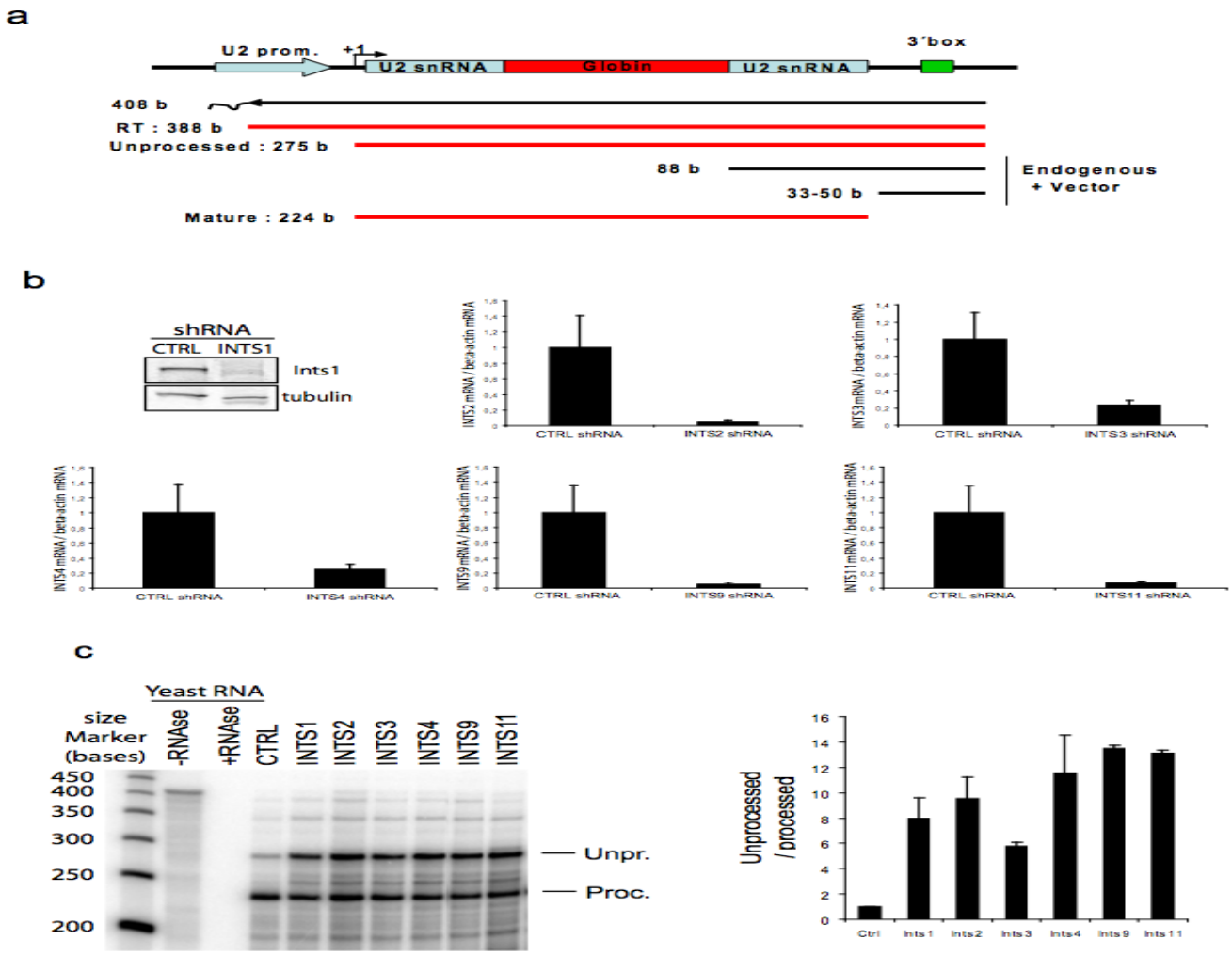

Figure 1. Integrator is essential for the processing of snRNA genes.

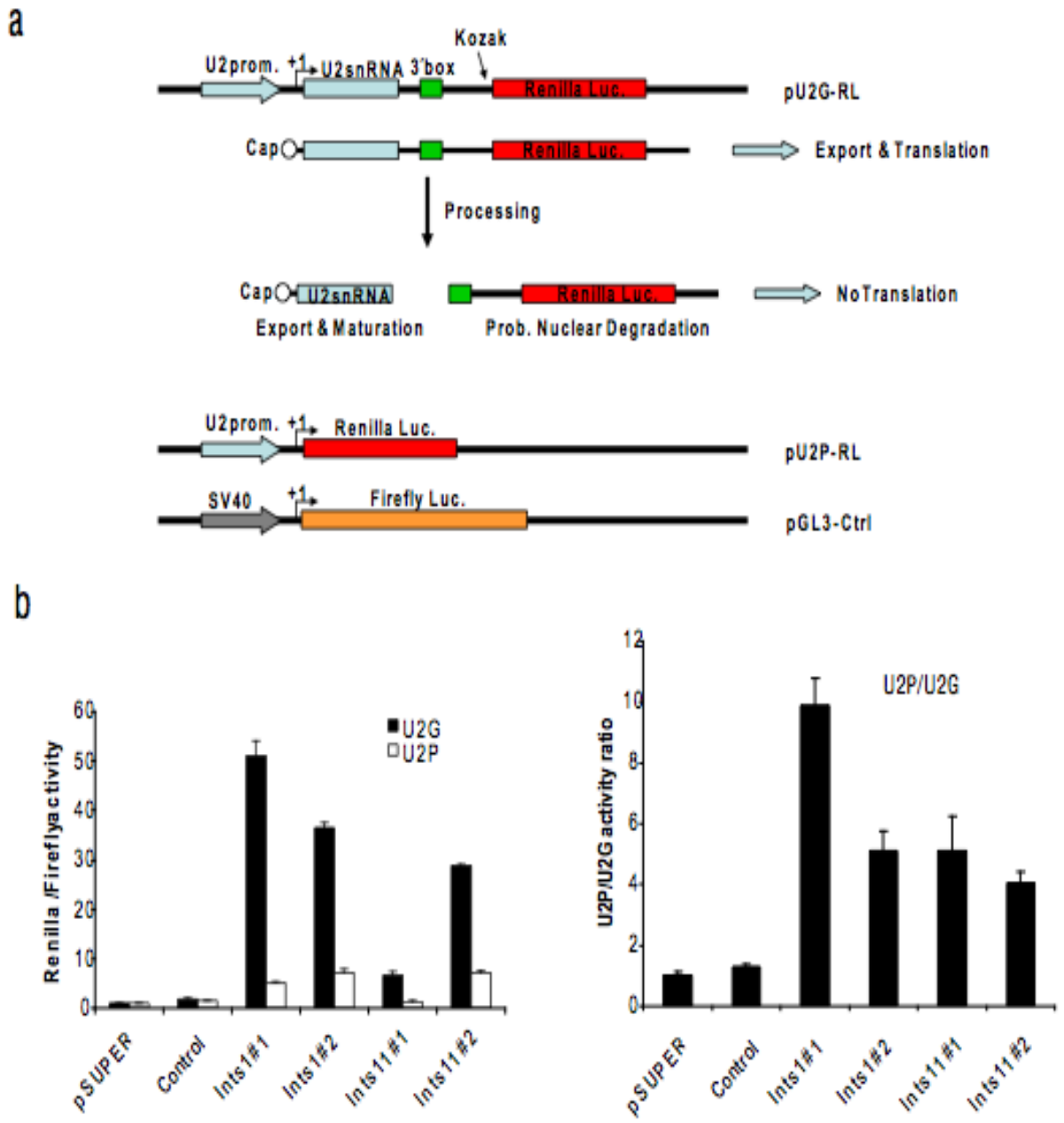

Figure 2. Integrator displays a function beyond that of 3'-end processing.

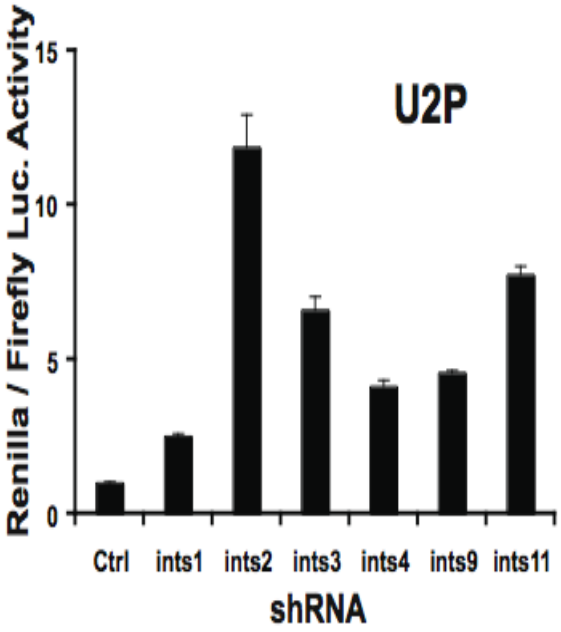

Figure 3. Depletion of multiple Integrator subunits leads to enhanced transcription.

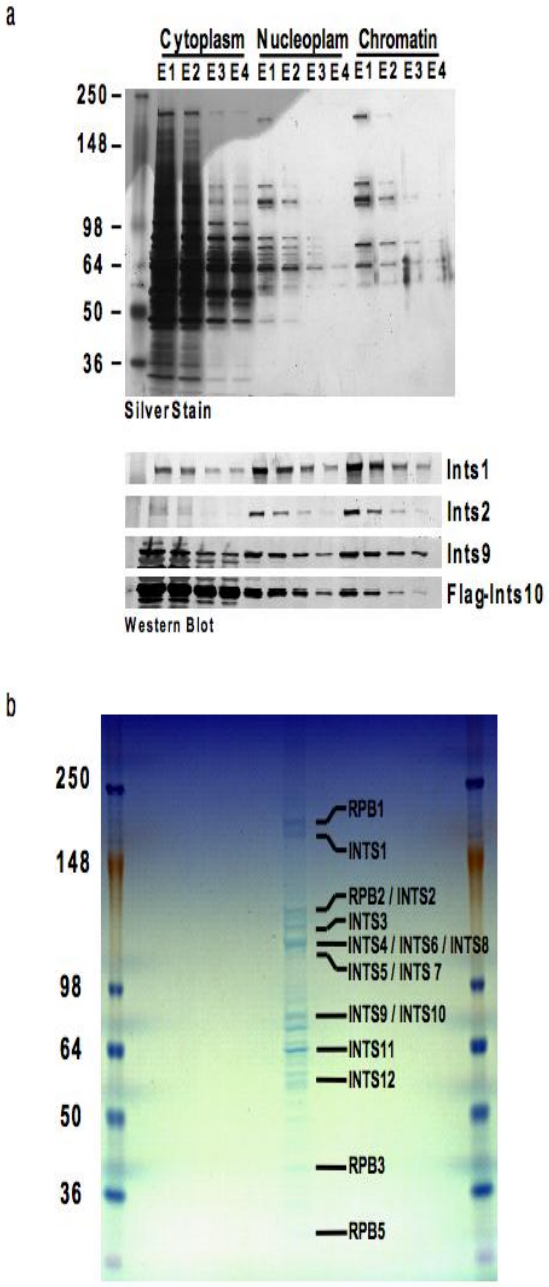

Figure 4. Purification of Integrator from chromatin fraction.
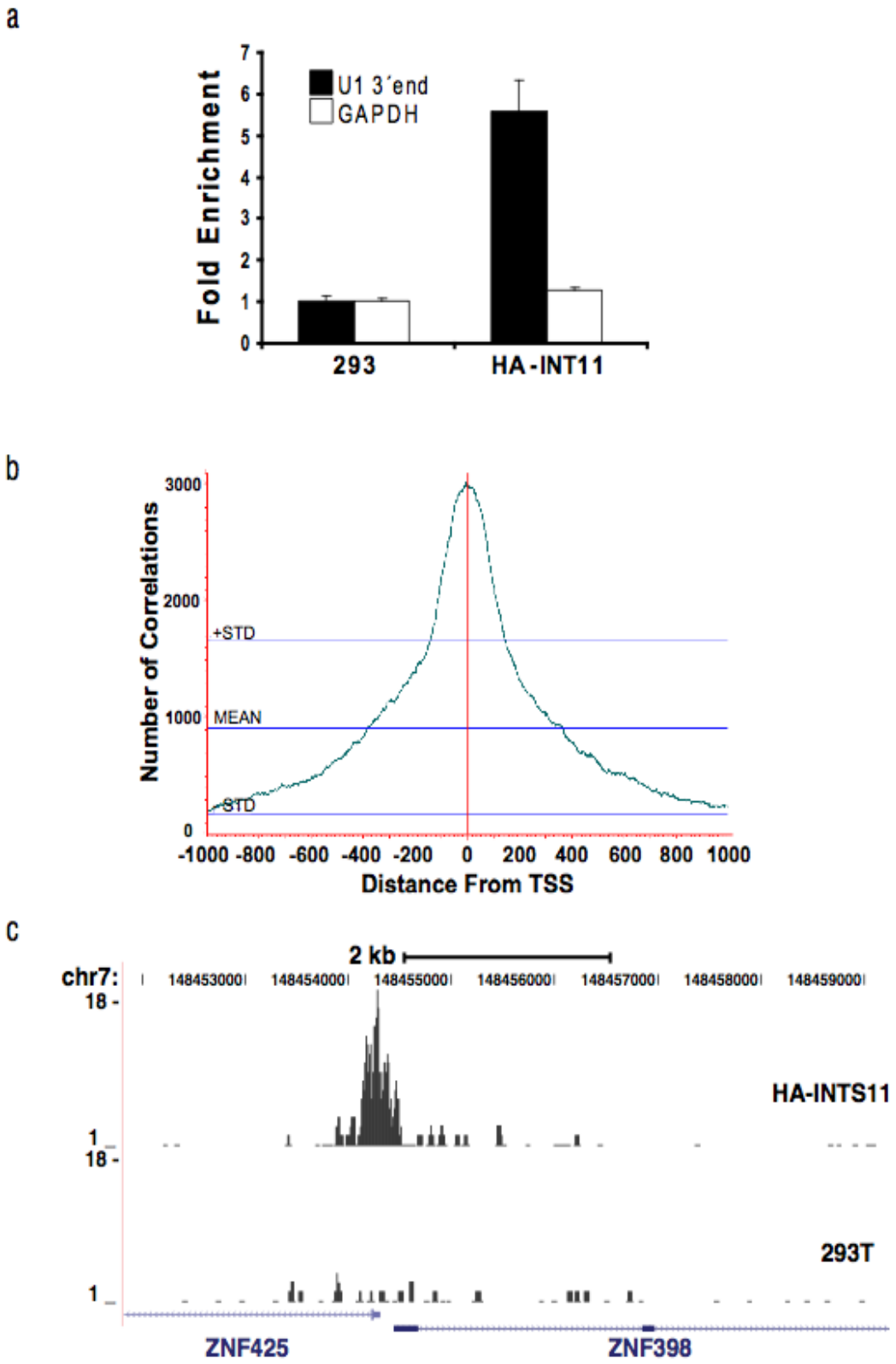
Figure 5. Global analysis of Integrator using ChIP-Sequencing