Extensive Alternative Splicing of the Repressor Element Silencing Transcription Factor Linked to Cancer

Citation

Published Version
doi:10.1371/journal.pone.0062217

Permanent link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:11179841

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

Share Your Story
The Harvard community has made this article openly available. Please share how this access benefits you. Submit a story.

Accessibility
Extensive Alternative Splicing of the Repressor Element Silencing Transcription Factor Linked to Cancer

Guo-Lin Chen*, Gregory M. Miller
Division of Neuroscience, New England Primate Research Center, Harvard Medical School, Southborough, Massachusetts, United States of America

Abstract
The repressor element silencing transcription factor (REST) is a coordinate transcriptional and epigenetic regulator which functions as a tumor suppressor or an oncogene depending on cellular context, and a truncated splice variant REST4 has been linked to various types of cancer. We performed a comprehensive analysis of alternative splicing (AS) of REST by rapid amplification of cDNA ends and PCR amplification of CDNAS from various tissues and cell lines with specific primers. We identified 8 novel alternative exons including an alternate last exon which doubles the REST gene boundary, along with numerous 5′/3′ splice sites and ends in the constitutive exons. With the combination of various splicing patterns (e.g. exon skipping and alternative usage of the first and last exons) that are predictive of altered REST activity, at least 45 alternatively spliced variants of coding and non-coding mRNA were expressed in a species- and cell-type/tissue-specific manner with individual differences. By examining the repertoire of REST pre-mRNA splicing in 27 patients with kidney, liver and lung cancer, we found that all patients without exception showed differential expression of various REST splice variants between paired tumor and adjacent normal tissues, with striking cell-type/tissue and individual differences. Moreover, we revealed that exon 3 skipping, which causes no frame shift but loss of a domain essential for nuclear translocation, was affected by pioglitazone, a highly selective activator of the peroxisome proliferator-activated receptor gamma (PPARγ) which contributes to cell differentiation and tumorigenesis besides its metabolic actions. Accordingly, this study demonstrates an extensive AS of REST pre-mRNA which redefines REST gene boundary and structure, along with a general but differential link between REST pre-mRNA splicing and various types of cancer. These findings advance our understanding of the complex, context-dependent regulation of REST gene expression and function, and provide potential biomarkers and therapeutic targets for cancer.


Editor: Emanuele Buratti, International Centre for Genetic Engineering and Biotechnology, Italy
Received February 14, 2013; Accepted March 18, 2013; Published April 16, 2013
Copyright: © 2013 Chen, Miller. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by National Institutes of Health grants DA025697 (GMM), DA030177 (GMM), and OD11103 (NEPRC). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: guo-lin.chen@hms.harvard.edu

Introduction
Alternative splicing (AS), a process to differentially link exons in a single precursor mRNA (pre-mRNA) to produce two or more different mature mRNAs, is a major contributor to transcriptome and proteome diversity, with >90% of human genes undergoing AS in a tissue- and developmental stage-specific manner [1,2]. It is now recognized that the coupling between transcription and splicing is crucial for AS regulation [3,4], and that exon-intron junctions or splice sites (SS) are specified by epigenetic modifications dependent on cellular context [4,5]. Accordingly, epigenetic modifications affect not only transcription, but also the co-transcriptional splicing [6,7]. Epigenetic regulation of pre-mRNA splicing, in line with the spatiotemporal selection of SS, suggest that AS cross-talks with environmental cues to contribute to adaptive responses and disease pathophysiology. Indeed, AS is modulated by the circadian clock, psychological stress, and numerous hormones and chemicals [8–10], while an estimated 15–60% of human genetic diseases, ranging from neurological to tumorigenic and metabolic disorders, involve splicing mutations [2,11–14].

The repressor element silencing transcription factor (REST, also known as NRSF for neuron-restrictive silencing factor), originally identified as a repressor of neuronal genes in non-neuronal cells [15,16], is now recognized as a coordinate transcriptional and epigenetic regulator that orchestrates the cellular epigenome in both neuronal and non-neuronal cells [17]. REST binds to widely distributed genomic regulatory sequences including the repressor element-1 (RE1) by a DNA-binding domain (DBD) which comprises 8 zinc finger motifs (ZFM), while its effect on gene expression is mediated by two independent repression domains (RD1 and RD2) which directly or indirectly recruit numerous transcriptional and epigenetic cofactors. Briefly, the N-terminal RD1 recruits mSin3, a scaffold for histone deacetylases (HDACs), while the C-terminal RD2 partners with the REST co-repressor (CoREST) which additionally recruits HDACs, methyl-CpG binding protein 2 (MeCP2), histone H3K4 lysine demethylase (LSD1) and H3K9 methyltransferases (G9a), as well as a component of the SWI/SNF chromatin remodeling complex-Brg1. By recruiting numerous cofactors to target gene loci, REST promotes dynamic, context-dependent chromatin organization and repression/activation of thousands of genes involved in many cellular processes including tumorigenesis, for which it functions as a tumor suppressor or an oncogene depending on cellular context [18,19]. The diverse, context-dependent function of REST is specified by multiple mechanisms.
including proteasomal degradation, nuclear translocation and pre-mRNA splicing [20–23], as well as the modulation by non-coding RNAs (ncRNAs) and binding affinity of REST to diverse RE1 and non-RE1 sites [24,25].

REST undergoes AS with a limited number of splice variants having been reported, of which a C-terminal truncated variant REST4, which contains RD1 and ZFMs 1–5, has been well documented [26–28]. Another two splice variants, REST1 which contains RD1 and ZFMs 1–4 [16], and REST-5F have also been documented. Notably, REST4 with ZFM-5 is transported to the nucleus while REST1 without ZFM-5 is not [32], and it was later demonstrated that ZFM-5 is essential for the nuclear targeting of REST [33].

In this study, we performed a comprehensive analysis of the AS of REST pre-mRNA and examined its relevance to cancer. We demonstrate that: 1) REST undergoes extensive AS across a gene boundary now doubled by a novel last exon (E5), with numerous coding and non-coding mRNAs being formed with a species- and cell-type/tissue-specific expression; 2) numerous REST splice variants, which are caused by various splicing patterns (e.g. exon skipping and alternative usage of the first and last exons) of altered REST activity, are generally but differentially linked to altered REST activity; and 3) exon 3 (E3) skipping, which causes no frame shift but loss of ZFM-5 essential for nuclear translocation, is remarkably affected by pioglitazone, a highly selective agonist for PPARγ which modulates cell differentiation and tumorigenesis besides its metabolic actions. These findings advance our understanding of the complexity of REST gene regulation and function, and provide potential biomarkers and therapeutic targets for cancer.

Materials and Methods
Ethics Statement
The use of human tissues was approved by the Harvard Institutional Review Board, and the related projects for which macaques and mice were euthanized were approved by the Institutional Animal Care and Use Committee for Harvard Medical School. The Harvard Medical School animal management program is accredited by the American Association for the Accreditation of Laboratory Animal Care (AAALAC) and meets National Institutes of Health standards as set forth in the Guide for the Care and Use of Laboratory Animals (DHHS Publication No. (NIH) 85–23 Revised 1983). The institution also accepts as mandatory the PHS Policy on Humane Care and Use of Laboratory Animals by Awardee Institutions and NIH Principles for the Utilization and Care of Vertebrate Animals Used in Testing, Research, and Training.

Materials (macaques and mice) involved in this study were cared for in compliance with National Institutes of Health, US Department of Agriculture, and Harvard Medical School guidelines for animal research. Macaques were single-housed and all efforts were made to reduce discomfort and provide enrichment opportunities (e.g. varied food supplements, foraging and task-oriented feeding methods, and interaction with caregivers and research staff). Specifically, macaques were fed twice daily (AM and PM) with a balanced commercially available Old World Primate Diet (e.g. Harlan Teklad 8714 Monkey Diet). Fruit and/or vegetable supplements were provided to all animals daily, and drinking water was provided ad libitum by automatic water fountains or plastic water bottles. Tissue samples were collected from five macaques which had been utilized in other experiments at the time of their necropsy. The macaques were euthanized following being anesthetized with ketamine HCl by an intravenous pentobarbital overdose, and exsanguinated. Mice were euthanized with carbon dioxide gas inhalation followed by cervical dislocation, and all efforts were made to minimize suffering.

Human and animal tissues
The cDNA samples derived from adult human tissues (kidney, liver, lung, pituitary, hippocampus, amygda and pons, 1 for each) were purchased from the BioChain Institute, Inc (Newark, CA). 27 pairs of tumor and adjacent normal tissues from patients diagnosed clinically with kidney, liver and lung cancers (9 pairs for each) were obtained from the UMass Cancer Center Tissue Bank (5 pairs for each cancer as tissue in RNAlater) and the BioChain Institute Inc (4 pairs for each cancer as total RNA). The demographic information of the patients briefly shown in Table 1. The human peripheral blood mononuclear cells (PBMCs), which were purified as described previously [34], were kindly gifted by Dr. Fred Wang at the Brigham & Women’s Hospital. We also collected tissues from rhesus monkeys and mice euthanized for other projects.

Cell lines and drug treatment
A total of 18 cell lines derived from human (HEK293, HEK293T, HepG2, NCCIT, SH-SYSY, A549, MCF7, K562, SK-N-MC, HeLa, Raji, TE671, Jurkat, Sup-T1 and induced pluripotent stem (iPS)), nonhuman primate (COS-7) and rodents (RN46A and PC12) were employed in this study. Except for RN46A and iPS, all the other 16 cell lines were obtained from American Type Culture Collection (Manassas, VA). RN46A cells were kindly provided by Dr. Scott Whittemore [35], while the iPS cells were originated from Dr. Stephen J. Hagarty [36]. To examine the effect of pioglitazone on REST pre-mRNA splicing, the NCCIT, HEK293T and HepG2 cells were treated with either 10 μM of pioglitazone (Sigma-Aldrich) or a matched concentration of the solvent (0.04% DMSO), and cells were harvested at 48 hours following treatment. Treatments were performed in duplicate on 3 independent occasions.

RNA isolation and cDNA synthesis
Total RNA was extracted using Trizol® reagent (Invitrogen). An aliquot of total RNA was reverse transcribed into cDNA using the Quantitect® Reverse Transcription Kit (Qiagen), while another aliquot was reverse transcribed into cDNA by using an anchored oligo-dT (anchor sequence given in Table 2). Synthesized cDNA was diluted to 50 ng/μl for use.

PCR amplification and DNA sequencing
A touchdown PCR protocol was employed for both standard and nested PCRs, and amplifications were performed in a MJ Research PTC-200 Peltier Thermal Cycler (GMI) in a total volume of 20 μl comprising 1 μl of template (50 ng/μl cDNA for standard or 1st step PCR), and 1:20 diluted product of 1st step PCR for 2nd step of nested PCR), 10 pmoles of each primer (Table 2), and 10 μl of GoTaq® Green Master Mix (Promega). The first step of nested PCR was performed by using cDNA made by anchored oligo-dT as template and oligo-dT anchor paired with E1, F2, E1, F2 and F1, F2 as the primer sets. For adult normal human tissues without cDNA made by anchored oligo-dT available, the oligo-dT anchor was replaced by E1, E2 and E3 (outside E1 and E4, respectively). Primers for rhesus macaque and rodents...
products with poor quality sequencing data were cloned into a service by the Functional Biosciences Inc (Madison, WI). PCR and sequenced. DNA sequencing was performed as commercial extension of 5 min at 72°C then fixed at 55°C.

Generation) from Roche (Indianapolis, IN) and GeneRacer Kit, a RNA Oligo was firstly ligated to RNA, followed by 28 (for 1st step of nested PCR) or 40 (for standard or 2nd step of nested PCR) cycles of 30 s denaturation at 95°C, 30 s annealing (temperature starting at 61°C and decreased by 0.5°C/cycle for the initial 12 cycles, then fixed at 55°C), and 60~180 s extension at 72°C, with a final extension of 5 min at 72°C. PCR products were loaded on a 2% agarose gel and amplicons of distinct size were excised, purified and sequenced. DNA sequencing was performed as commercial service by the Functional Biosciences Inc (Madison, WI). PCR products with poor quality sequencing data were cloned into pGEM-T vector (Promega) for further sequencing.

5’/3’ Rapid amplification of cDNA ends (RACE)

Total RNA generated from HEK293, HEK293T, HepG2 and SH-SY5Y were used to perform the 5’ and 3’ RACE, which were carried out by two commercial kits, the 5’/3’ RACE Kit (2nd Generation) from Roche (Indianapolis, IN) and GeneRacer® from Invitrogen, which differ in strategies for 5’ RACE. For 5’ RACE with GeneRacer® kit, a RNA Oligo was firstly ligated to RNA, followed by cDNA synthesis using Oligo-dT and nested PCR using REST-specific reverse primers (e.g. E1aF1 and E1bF1) paired with a GeneRacer® 5’ primer (homologous to the RNA Oligo). For 5’ RACE with Roche’s kit, CDNA was first synthesized from total RNA using a REST-specific reverse primer (e.g. E1R2, outside), followed by tailing with dATP and terminase transferase and subsequent nested PCR using anchored Oligo-dT or anchor primer paired with a REST-specific reverse primer (e.g. E1R1, E1bR2).

Table 1. Demographic data for the patients with cancer.

<table>
<thead>
<tr>
<th>ID#</th>
<th>Gender</th>
<th>Age(y)</th>
<th>Basic Diagnosis</th>
<th>Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ki#1</td>
<td>Male</td>
<td>2</td>
<td>Nephroblastoma</td>
<td>n.a.</td>
</tr>
<tr>
<td>Ki#2</td>
<td>Male</td>
<td>45</td>
<td>Renal Cell Carcinoma</td>
<td>2</td>
</tr>
<tr>
<td>Ki#3</td>
<td>Male</td>
<td>43</td>
<td>Renal Cell Carcinoma</td>
<td>2</td>
</tr>
<tr>
<td>Ki#4</td>
<td>Female</td>
<td>50</td>
<td>Renal Cell Carcinoma</td>
<td>2</td>
</tr>
<tr>
<td>Ki#5</td>
<td>Male</td>
<td>52</td>
<td>Renal Cell Carcinoma</td>
<td>2</td>
</tr>
<tr>
<td>Ki#6</td>
<td>Female</td>
<td>75</td>
<td>Renal Cell Carcinoma</td>
<td>2</td>
</tr>
<tr>
<td>Ki#7</td>
<td>Male</td>
<td>53</td>
<td>Renal Cell Carcinoma</td>
<td>n.a.</td>
</tr>
<tr>
<td>Ki#8</td>
<td>Male</td>
<td>63</td>
<td>Renal Cell Carcinoma</td>
<td>n.a.</td>
</tr>
<tr>
<td>Ki#9</td>
<td>Female</td>
<td>55</td>
<td>Adenocarcinoma</td>
<td>n.a.</td>
</tr>
<tr>
<td>Li#1</td>
<td>Male</td>
<td>59</td>
<td>Hepatocellular Carcinoma</td>
<td>3</td>
</tr>
<tr>
<td>Li#2</td>
<td>Male</td>
<td>59</td>
<td>Hepatocellular Carcinoma</td>
<td>n.a.</td>
</tr>
<tr>
<td>Li#3</td>
<td>Female</td>
<td>71</td>
<td>Hepatocellular Carcinoma</td>
<td>3</td>
</tr>
<tr>
<td>Li#4</td>
<td>Male</td>
<td>58</td>
<td>Cholangiocarcinoma</td>
<td>2</td>
</tr>
<tr>
<td>Li#5</td>
<td>Male</td>
<td>59</td>
<td>Cholangiocarcinoma</td>
<td>2</td>
</tr>
<tr>
<td>Li#6</td>
<td>Female</td>
<td>36</td>
<td>Cholangiocarcinoma</td>
<td>2</td>
</tr>
<tr>
<td>Li#7</td>
<td>Male</td>
<td>60</td>
<td>Hepatocellular Carcinoma</td>
<td>n.a.</td>
</tr>
<tr>
<td>Li#8</td>
<td>Male</td>
<td>33</td>
<td>Hepatocellular Carcinoma</td>
<td>n.a.</td>
</tr>
<tr>
<td>Li#9</td>
<td>Male</td>
<td>47</td>
<td>Hepatocellular Carcinoma</td>
<td>n.a.</td>
</tr>
<tr>
<td>Lu#1</td>
<td>Male</td>
<td>51</td>
<td>Adenocarcinoma</td>
<td>3</td>
</tr>
<tr>
<td>Lu#2</td>
<td>Male</td>
<td>63</td>
<td>Squamous Cell Carcinoma</td>
<td>n.a.</td>
</tr>
<tr>
<td>Lu#3</td>
<td>Male</td>
<td>64</td>
<td>Adenocarcinoma</td>
<td>3</td>
</tr>
<tr>
<td>Lu#4</td>
<td>Male</td>
<td>51</td>
<td>Adenocarcinoma</td>
<td>3</td>
</tr>
<tr>
<td>Lu#5</td>
<td>Male</td>
<td>61</td>
<td>Adenocarcinoma</td>
<td>3</td>
</tr>
<tr>
<td>Lu#6</td>
<td>Male</td>
<td>63</td>
<td>Adenocarcinoma (100%)</td>
<td>3</td>
</tr>
<tr>
<td>Lu#7</td>
<td>Male</td>
<td>46</td>
<td>Squamous cell carcinoma</td>
<td>n.a.</td>
</tr>
<tr>
<td>Lu#8</td>
<td>Male</td>
<td>63</td>
<td>Squamous cell carcinoma</td>
<td>n.a.</td>
</tr>
<tr>
<td>Lu#9</td>
<td>Male</td>
<td>67</td>
<td>Squamous cell carcinoma</td>
<td>n.a.</td>
</tr>
</tbody>
</table>

Note: The grade information is not available for samples from BioChain Institute, Inc. doi:10.1371/journal.pone.0062217.t001

Table 2. Oligos used for the detection of specific REST splice variants.

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Standard/nested PCR</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E4R1</td>
<td>5'-ccagcaccacatattacacc-3'</td>
<td></td>
</tr>
<tr>
<td>E3F2</td>
<td>5’-ccagcaccacatattacacc-3’</td>
<td></td>
</tr>
<tr>
<td>E5F1</td>
<td>5’-ccagcaccacatattacacc-3’</td>
<td></td>
</tr>
<tr>
<td>E5R2</td>
<td>5’-ccagcaccacatattacacc-3’</td>
<td></td>
</tr>
<tr>
<td>E6F3</td>
<td>5’-ccagcaccacatattacacc-3’</td>
<td></td>
</tr>
<tr>
<td>E6R4</td>
<td>5’-ccagcaccacatattacacc-3’</td>
<td></td>
</tr>
<tr>
<td>E7F5</td>
<td>5’-ccagcaccacatattacacc-3’</td>
<td></td>
</tr>
<tr>
<td>E7R6</td>
<td>5’-ccagcaccacatattacacc-3’</td>
<td></td>
</tr>
<tr>
<td>E8F7</td>
<td>5’-ccagcaccacatattacacc-3’</td>
<td></td>
</tr>
<tr>
<td>E8R8</td>
<td>5’-ccagcaccacatattacacc-3’</td>
<td></td>
</tr>
<tr>
<td>E9F9</td>
<td>5’-ccagcaccacatattacacc-3’</td>
<td></td>
</tr>
<tr>
<td>E9R10</td>
<td>5’-ccagcaccacatattacacc-3’</td>
<td></td>
</tr>
<tr>
<td>E10F1</td>
<td>5’-ccagcaccacatattacacc-3’</td>
<td></td>
</tr>
<tr>
<td>E10R1</td>
<td>5’-ccagcaccacatattacacc-3’</td>
<td></td>
</tr>
<tr>
<td><strong>qRT-PCR</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E11R1</td>
<td>5’-ccagcaccacatattacacc-3’</td>
<td>SYBR Green I</td>
</tr>
<tr>
<td>E11F1</td>
<td>5’-ccagcaccacatattacacc-3’</td>
<td>SYBR Green I</td>
</tr>
<tr>
<td>E11R2</td>
<td>5’-ccagcaccacatattacacc-3’</td>
<td>SYBR Green I</td>
</tr>
<tr>
<td>E11F3</td>
<td>5’-ccagcaccacatattacacc-3’</td>
<td>SYBR Green I</td>
</tr>
<tr>
<td>E11R4</td>
<td>5’-ccagcaccacatattacacc-3’</td>
<td>SYBR Green I</td>
</tr>
<tr>
<td>E11F5</td>
<td>5’-ccagcaccacatattacacc-3’</td>
<td>SYBR Green I</td>
</tr>
<tr>
<td>E11R6</td>
<td>5’-ccagcaccacatattacacc-3’</td>
<td>SYBR Green I</td>
</tr>
<tr>
<td>E11F7</td>
<td>5’-ccagcaccacatattacacc-3’</td>
<td>SYBR Green I</td>
</tr>
<tr>
<td>E11R8</td>
<td>5’-ccagcaccacatattacacc-3’</td>
<td>SYBR Green I</td>
</tr>
<tr>
<td>E11F9</td>
<td>5’-ccagcaccacatattacacc-3’</td>
<td>SYBR Green I</td>
</tr>
<tr>
<td>E11R10</td>
<td>5’-ccagcaccacatattacacc-3’</td>
<td>SYBR Green I</td>
</tr>
</tbody>
</table>

*The E1a/E2,E 1b/E2,E 1c/E2 junctions can be quantified by SYBR Green I qRT-PCR using the E1aF1,E 1bF1 and E1cF1 paired with E2R1, respectively. doi:10.1371/journal.pone.0062217.t002
inside). To perform 3’ RACE, cDNA was synthesized from total RNA using anchored Oligo-dT, followed by nested PCR using REST-specific forward primers (e.g. E1aF1 and E1aF2) paired with the anchor primer.

Quantitative real-time PCR (qRT-PCR) assay
Using primers listed in Table 2, we developed SYBR Green I and/or Hybridization Probe qRT-PCR for specific exon-exon junctions and exon 2 (E2), presumably represents overall expression of coding mRNAs regardless of the first and last exons), as well as the housekeeping gene GAPDH. The qRT-PCR assays were performed as previously described on a Roche LightCycler 2.0 system [37], with the threshold cycle (Ct) values being employed to evaluate the expression change between paired tissues or cells by the 2^−ΔΔCt approach [38]. PCR reactions were run in duplicate. For E1a/E4 junction which is expressed at low levels in most cases, only SYBR Green I assay is available and the Ct values are likely confounded by primer dimer formed by excessive remaining primers, so we performed qRT-PCR by using the product of the 1st step PCR as template. Owing to the high sequence identity of 3′ ends of E2 and E4, qRT-PCR assay specific for E2/E4 junction which is expressed at low levels in most cases, not attainable.

Bioinformatics and data analysis
Prediction of the open reading frame (ORF) of specific REST variants was performed by using the StarORF program (http://star.mit.edu/orf/runapp.html), and epigenetic information at the REST locus was retrieved from the UCSC Genome Browser (http://genome.ucsc.edu/cgi-bin/hgGateway). Comparisons of qRT-PCR-assayed expression levels of specific exon-exon junctions between paired tissues or cells were carried out by the 2^−ΔΔCt approach using appropriate reference, and two-fold change was considered as significant.

Results
Identification of E2/E3 skipping expressed in a species-dependent manner
We performed standard and nested PCRs with cDNA samples derived from numerous human tissues (liver, kidney, lung, pituitary, hippocampus, amygdala and pons) and cell lines (HEK293, HEK293T, HepG2, NCCIT, SH-SY5Y, A549, MCF7 and iP5s). Using a reverse primer E4R1 targeting the proximal exon 4 (E4) paired with the forward primers E4aF1 and E4aF2 targeting exons 1a (E1a) and 2 (E2) (Figure 1A and Table 2), respectively, we identified several REST splice variants with E2 and/or E3 skipped (Figure 1B and 1C). The variant with E2 skipped is abundantly expressed in all tested cell lines and tissues except amygdala, while the variants with E3 alone or plus E2 skipped were expressed at low levels in a subset of tissues and cell lines. Notably, E2/E3 skipping is predominantly associated with E1a but rarely with E1b and E1c. Additionally, we found no variants containing both E4 and E5 suggesting that E4 and E5 are mutually exclusive and that E1a, like the case for E2 and E5, can be completely skipped. Accordingly, the novel last exon E5 doubles the human REST gene boundary from ~28 kb to ~59 kb (Figure 3A); E5 inclusion was not observed in nonhuman primates and rodents.

Discovery of a novel last exon which doubles human REST gene boundary
We performed RACE to determine the 5′ and 3′ ends of human REST mRNA. Unexpectedly, we identified a novel polyadenylated exon (E5) that locates ~30 kb downstream of E4 and partially overlaps in opposite direction with exon 5 of the nitric oxide-associated-1 (NOD1) gene (Figure 2A), which encodes a GTPase essential for mitochondrial protein synthesis [39]. By nested PCRs using an E5-specific reverse primer (E5R1) paired with E4aF1 and E4aF2, respectively, we found that E5 inclusion, occasionally in combination with E2/E3 skipping, is expressed in most human tissues and cell lines (Figure 2B, 2C and Figure S1). Like E2/E3 skipping, E5 inclusion is primarily associated with E1a but rarely with E1b and E1c. Notably, we found no variants containing both E4 and E5, suggesting that E4 and E5 are mutually exclusive and that E1a, like the case for E2 and E5, can be completely skipped. Accordingly, the novel last exon E5 doubles the human REST gene boundary from ~28 kb to ~59 kb (Figure 3A); E5 inclusion was not observed in nonhuman primates and rodents.

Extensive AS of REST pre-mRNA
By examining above-mentioned primate tissues and cell lines, as well as paired tumor and adjacent normal tissues from cancer patients mentioned hereinafter, we revealed an extensive AS of REST pre-mRNA. As shown in Figure 3A, besides E5 and the previously reported exon N (now renamed as N3c), we identified another 7 novel alternative exons (N1, N2a, N2b, N3a, N3b, N4a and N4b), along with numerous 5′/3′ SS and ends in the constitutive exons E2 and E4. Regardless of the 5′/3′ ends in E2 and E4, at least 45 variants (S1–S45) are formed by various splicing patterns including exon skipping, alternative 5′/3′ SS, mutual exclusion and alternative usage of the first and last exons (Figure 3B). Sequences of the novel variants have been deposited in GenBank with assigned accession numbers JX896957-JX896993 and KC117262-KC117266.

29 of the 45 splice variants involve full or partial skipping of E2 where translation initiates, such that they may function as ncRNAs due to lack of the translation start site (TSS), except that several variants (S16, S19 and S28) whose TSS was preserved are predictive of N-terminal truncated REST protein isoforms. Similarly, partial or complete skipping of E4, which is frequently accompanied by E2 skipping and/or E5 inclusion, produces numerous ncRNAs or coding mRNAs predictive of truncated REST protein isoforms. In contrast, skipping of the 84-bp E3 (S4 and S34) causes no frame shift but loss of ZFM-5. Notably, variants with intact E2 (or plus E3) followed by alternative exons (but not E4) are predicted to encode C-terminal truncated REST proteins, of which 8 variants with E3 encode REST4 (S3 and S12) or REST1-like (S2, S21, S31, S33, S39 and S41) proteins with RD1 and ZFMs 1–5, while another 4 variants (S34, S39–S41) with E3 skipped encode REST1 with RD1 and ZFMs 1–4 (Figure 3C).

Most of the 45 splice variants are expressed at low levels in a cell-type/tissue-specific manner; however, at least one variant was observed in all the tested tissues and cell lines (Figure 1, 2 and Figure S3).

Link between REST pre-mRNA splicing and cancer
We compared the expression of specific splicing patterns and variants between paired tumor (T) and adjacent normal (N) tissues...
from 27 patients with kidney (K), liver (L) and lung (Lu) cancers (9 pairs for each cancer, Table 1) by nested PCR and qRT-PCR assays, with a focus on E2/E3 skipping and alternative usage of the first and last exons. The qRT-PCR assays were designed targeting specific exon-exon junctions (Figure 4A), and expression changes of specific splicing between T and N are shown in folds (T over N) for each subject (Figure 4B). With the exception of E1a/E3 and E1a/E4 which represent S5 and S6 variants, respectively, the exon-exon junctions do not necessarily represent a single specific splice variant, which however can be detected by nested PCRs (Figure 4C). Accordingly, both qRT-PCR and nested PCR assays were taken into consideration for the comparison of REST pre-mRNA splicing profile between paired T and N tissues. While the wide expression of E2/E3 skipping and E3 inclusion in human was further validated, we found that all the 27 patients without exception showed differential expression of numerous REST splice variants caused by specific splicing patterns between paired T and N tissues, with a striking tissue-type and individual difference.

As shown in Figure 4B and 4C, variants with E2/E3 skipped were differentially expressed between paired T and N tissues for most patients. Of variants using E4 as the last exon, S6 (both E2 and E3 skipped) showed strikingly differential expression for all the patients except Lu#4 and 7. Specifically, 7, 5 and 1 patients with kidney, liver and lung cancer, respectively, showed increased S6 expression, whereas 2, 4 and 6 patients with kidney, liver and lung cancer, respectively, exhibited decreased S6 expression. Although qRT-PCR assay for E3 skipping only (i.e. E2/E4 junction) is not attainable, nested PCR with E2F1/E4R1 showed apparently...
differential expression of S4 between paired T and N for a portion of patients. Meanwhile, differential expression of E2-skipped variants S5 and S15 between paired T and N from some patients was shown by qRT-PCR and nested PCR, respectively. Moreover, some other variants (e.g. S14, S16 and S17) with E1b as the first exon and E2 partially skipped were differentially expressed between paired T and N for a few patients. For example, S14 and S16 were only observed in the T tissues of 4 patients (Ki#5, 6 for S14 and Ki#4, Lu#9 for S16, respectively). Likewise, of variants using E5 as the last exon, S34 (E3 skipped), S35 (E2 skipped) and S38 (both E2 and E3 skipped) showed apparently differential expression between paired T and N as indicated by nested PCRs with E1aF1/E5R1 and E2F1/E5R1. Accordingly, E2/E3 (especially E3) skipping is generally but differentially linked to different types of cancer with striking cell-type/tissue and individual differences.

As shown in Figure 4C, with the exception of 3 patients (Li#5, 6 and Lu#6) without detectable expression of E5 inclusion, all other patients showed differential expression of E5-included variants between paired T and N tissues. Particularly, the expression of S33/S39 without E2/E3 skipping was gained and lost in the T tissues of 4 patients (Ki#5, 6 for S14 and Ki#4, Lu#9 for S16, respectively). Likewise, of variants using E5 as the last exon, S34 (E3 skipped), S35 (E2 skipped) and S38 (both E2 and E3 skipped) showed apparently differential expression between paired T and N as indicated by nested PCRs with E1aF1/E5R1 and E2F1/E5R1. Accordingly, E2/E3 (especially E3) skipping is generally but differentially linked to different types of cancer with striking cell-type/tissue and individual differences.

As shown in Figure 4C, with the exception of 3 patients (Li#5, 6 and Lu#6) without detectable expression of E5 inclusion, all other patients showed differential expression of E5-included variants between paired T and N tissues. Particularly, the expression of S33/S39 without E2/E3 skipping was gained and lost in the T tissues of 4 patients (Ki#5, 6 for S14 and Ki#4, Lu#9 for S16, respectively). Notably, nested PCR with E1aF1/E5R1 and E2F1/E5R1 showed that all 9 kidney cancer patients expressed E5 in their N tissues, of them 3 (Ki#1, 7, 8) lost E5 expression in their T tissues. By contrast, all 9 patients with liver cancer showed no E5 expression in their N tissues, and of them 4 (Li#1, 2, 3, 7) gained E5 expression in their T tissues. The differential expression of E5-included variants between paired T and N tissues shown by nested PCR was supported by qRT-PCR assay of E3/E5 junction (Figure 4B). In contrast, pioglitazone (10 μM) slightly reduced E3 skipping in HepG2 cells while it exerted no effect on E5 skipping in HEK293T cells.

Discussion

We reveal that REST undergoes extensive AS across an unexpectedly large gene boundary defined by a novel alternate last exon (E5), with numerous ncRNAs and coding mRNAs being expressed in a species- and cell-type/tissue-specific manner with individual differences. Notably, we found that exon (E2, E3 and E4) skipping is preferentially associated with the usage of the conventional E1a as the first exon, i.e., AS of REST pre-mRNA
is promoter-dependent, suggesting that alternative promoters differ not only in the strength, but also in transcription elongation which is a major determent of pre-mRNA splicing. Since most splice variants involve a complete or partial skipping of exons (e.g. E2, E3 and E4) encoding specific functional domains of REST, their functional significance is generally predictable. Accordingly, these alternatively spliced REST variants, which are mostly expressed in a cell-type/tissue-specific manner with individual differences, presumably contribute to the diverse, context-dependent regulation of REST. Notably, we found that all tissues and cell lines without exception express at least one REST splice variant, and no apparent difference in the expression pattern of specific splice variants was observed between neuronal and non-neuronal tissues. In agreement, similar levels of the initial REST transcripts were reported between neuronal and non-neuronal cells [42], while the promoter of REST exhibits cell-type-independent active transcription as indicated by publicly available epigenetic data (Figure 6), suggesting that transcriptional regulation is unlikely a major contributor to cell-type/tissue-specific expression of REST. Hence, in line with epigenetic regulation of AS and the emerging role of AS in adaptive responses [2,6], our findings strongly suggest that pre-mRNA splicing, rather than transcription regulation, substantially contributes to the diverse, context-dependent REST gene expression and function.

REST functions as a tumor suppressor or an oncogene depending on cellular context, with both increased and decreased REST activity having been reported in different types of cancer [18,19]. In accordance, we found that numerous REST splice variants produced by E2/E3 skipping and alternative usage of the

---

**Figure 3. Details of the identified REST exons, splicing patterns and resultant splice variants.** (A) Identified REST exons and their 5'3' SS or ends. The UCSC genes track was used to briefly indicate the chromosomal location of REST gene locus, while the genomic sequence NG_029447 was employed as a reference to clarify the exact positions of the exons. Major domains of the REST protein (NP_005603) were illustrated parallel to the corresponding coding region so as to predict consequences of the AS of REST pre-mRNA at the protein level. The constitutive and alternative exons, as well as their 5'3' SS or ends, are color-coded as indicated. Note that: 1) exon N in Figure 1 is renamed as N3c here; 2) E2k is an extension of E2a without splicing; and 3) E5 and N4b are not actually presented in NG_029447. (B) Patterns of REST pre-mRNA splicing and resultant mRNA variants. Putative coding variants are color-coded (blue—no amino acid change predicted; red—truncated protein) or shadowed (loss of ZFM-5). (C) Prediction of C-terminal truncated REST proteins for specific mRNA variants. The exon-exon junctions and premature stop codons are underlined.

doi:10.1371/journal.pone.0062217.g003
first and last exons, which are predictive of altered REST activity, are generally but differentially linked to different types of cancer. Particularly, in accordance with the involvement of REST4 and another C-terminal truncated REST-FS (caused by a frameshift mutation) in tumorigenesis [26–28,43], we found that the usage of E5 as the last exon, which is predictive of a C-terminal truncated REST4-like protein, is differentially linked to various types of cancer. Unlike REST4 which exists in both primates and rodents, E5 inclusion is only expressed in human but it shows a wide tissue distribution, suggesting that it may contribute to the pathophysiology of a wide spectrum of human diseases. Besides encoding a REST4-like protein, E5-included variants overlap in opposite direction with NOA1 transcript(s), i.e., transcripts of REST and NOA1 act as natural antisense transcripts for each other, making it possible that transcripts of one gene regulate expression of the other gene through various mechanisms [44]. As a GTPase essential for mitochondrial protein synthesis, NOA1 is involved in oxidative stress and apoptosis [39,45,46]. Accordingly, E5 inclusion may mediate a coordinated effect of REST and NOA1 on cellular functions. In addition, E5 inclusion results in altered 3' UTR, which contributes to posttranscriptional regulation (e.g. mRNA stability) of gene expression.

Since E2 skipping eliminates the TSS, E2-skipped variants may function as ncRNAs which potentially modulate REST gene expression. Similar to E5 inclusion which is exclusively expressed in human, E2 skipping is ubiquitously expressed in human but rarely expressed in nonhuman primates and rodents, suggesting that human evolution involves a gain of much more complex REST pre-mRNA splicing, which may contribute to context-dependent human genome function that is by far more complex than other species. As for the E3 skipping, it causes no frame shift but loss of ZFM-5 essential for nuclear translocation [32,33], such that it provides an alternate mechanism for the regulation of REST activity. Notably, while REST-5F caused by E3 skipping was only previously observed in SH-SY5Y cells [27], we revealed a ubiquitous though usually non-abundant expression of E3 skipping in primates, suggesting that E3 skipping might be a common regulator which specifies the diverse, context-dependent function of REST. In accordance, we found that numerous E3-skipped variants were differentially expressed between paired tumor and adjacent normal tissues for most patients with cancer, and that E3 skipping is modulated by the PPARγ activator pioglitazone which contributes to cell differentiation and tumorigenesis [40,41,47]. In addition, our preliminary data showed that E3 skipping is linked to
virus-induced transformation of B-lymphocytes in both human and nonhuman primates (unpublished data). Hence, E3 skipping represents a potential biomarker for cancer. With respect to the alternative usage of the first exon, it alters the 5′-untranslated region which plays an important role in posttranscriptional regulation (e.g., mRNA stability, targeting and translation) of gene expression [37], such that it may affect the expression level of REST protein.

It has been documented that REST is a prognostic factor and therapeutic target for cancer [48,49]. Accordingly, correction of aberrant REST pre-mRNA splicing provides a new strategy for the treatment of cancer. Indeed, pioglitazone regulation of E3 skipping strongly supports the feasibility of manipulating REST activity through modulation of pre-mRNA splicing. Since E3 contains a conserved binding motif for PPARγ, it is likely that the binding of activated PPARγ to E3 interrupts the action of other splicing
factor(s) on E3 skipping/inclusion. To our knowledge, this is the first study reporting the regulation of pre-mRNA splicing by a ligand of PPARγ, which usually affects gene transcription [50]. PPARγ is implicated in a wide variety of biological processes including adipogenesis, glucose metabolism, inflammation and tumorigenesis, and it is the molecular target of the thiazolidinedione (TZD) class of antidiabetic drugs including pioglitazone and rosiglitazone [51]. It has been shown that TZDs suppress the growth of several cancer lines in vitro and in vivo, and lines of preclinical evidence supports the antineoplastic effects of PPARγ agonists; however, results from clinical trials show modest success [52]. Recently, the use of pioglitazone for type 2 diabetes mellitus is reportedly associated with an increased risk of bladder cancer [53], suggesting a context-dependent, bidirectional effect of PPARγ on tumorigenesis, which is in accordance with its cell-specific effect on E3 skipping (Fig.5), as well as the notion that both decreased and increased REST activity may contribute to tumorigenesis [18,19]. Mechanism(s) by which PPARγ modulates tumorigenesis are not yet fully understood; however, based on the role of REST in tumorigenesis and the regulation of REST activity by E3 skipping, we speculate that PPARγ regulation of REST pre-mRNA splicing is attributable. Accordingly, the context-dependent REST activity modulated by E3 skipping may underlie the context-dependent effect of PPARγ on tumorigenesis, such that E3 skipping represents a potential therapeutic target that might be utilized for personalized medicine for cancer. Moreover, REST targets numerous genes involved in metabolism [54], suggesting a context-dependent, bidirectional effect of PPARγ regulation of REST pre-mRNA splicing may contribute, at least in part, to the metabolic actions of PPARγ. In this regard, this study provides a novel mechanism underlying the biological actions of PPARγ and the close link between metabolism and tumorigenesis [55]. Furthermore, PPARγ exerts a neuroprotective effect on neurodegenerative disorders including Huntington disease, which is caused by disassociation of mutant Huntington protein with REST in the cytoplasm and therefore enhanced nuclear translocation and aberrant accumulation of REST in nucleus [21,36,57]. It is tempting to speculate that the neuroprotective effect of PPARγ might be explained by its regulation of E3 skipping (presumably increased in neurons), which may result in reduced translocation of REST into the nucleus and alleviation of the repressive effect of REST on neuronal genes essential for the maintenance of neurons. Thus, E3 skipping may have implications for a wide variety of human diseases.

In summary, this study revealed an extensive AS of REST pre-mRNA and a close link between aberrant REST pre-mRNA splicing and various types of cancer. The findings not only advance our understanding of the complexity of REST gene expression and function, but also provide potential biomarkers and therapeutic targets for the diagnosis and individualized treatment of cancer. However, our findings require further validation in a large population of patients with different types and prognosis of cancer, and warrant further investigation of mechanisms underlying REST pre-mRNA splicing regulation and biological functions of specific REST splice variants.

Supporting Information

Figure S1 Detection of E2/E3 skipping and E5 inclusion in additional human cell lines and PBMCs. Nested PCRs were performed by using the forward primers E1aF2 and E2F1 paired with the reverse primers E3R1 and E3R3, respectively. (TIF)

Figure S2 Individual difference in REST E2 skipping in rhesus monkeys. The primer set E1aF2/E4R1 was employed to perform the standard PCR using cDNA samples from pons and raphe tissues from 4 rhesus monkeys. E2 skipping was observed in 1 of the 4 macaque pons and raphe tissues, respectively. (TIF)

Figure S3 Expression profile of REST splice variants in human cell lines. The abundance of the expression was briefly estimated by the band of standard/nested PCRs and was color-encoded as indicated. The tissue distribution was given for variants (S9, S10, S14-S16, S24, S30, S36 and S41) that were not detected in cell lines. The glioblastoma tissue was obtained from UMass Cancer Center Tissue Bank. The GenBank accession numbers are given for 42 of the 43 variants. (TIF)

Acknowledgments

We thank the NEPRC Primate Genetics Core for bioinformatics support. We are grateful to Hong Yang, Lisa Ogawa, Danica Yang and Sherry Wu for their technical support. We thank Dr. Stephen R. Lyle at the University of Massachusetts Medical School, Drs. Susan V. Westmoreland, Bertha K. Mardras and Andrew D. Miller at the NEPRC, Dr. Fred Wang at the Brigham & Women’s Hospital, and Dr. Wen-Yuan Wang at the Massachusetts Institute of Technology for their kind helps with tissue and cell line collection.

Author Contributions

Conceived and designed the experiments: GLC GMM. Performed the experiments: GLC. Analyzed the data: GLC. Contributed reagents/materials/analysis tools: GLC. Wrote the paper: GLC GMM.

References


