Mono-Allelic Retrotransposon Insertion Addresses Epigenetic Transcriptional Repression in Human Genome

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters.

| Citation       | Byun, Hyang-Min, Kyu Heo, Kasey J Mitchell, and Allen S Yang. 2012. Mono-allelic retrotransposon insertion addresses epigenetic transcriptional repression in human genome. Journal of Biomedical Science 19(1): 13. |
|----------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Published Version | doi:10.1186/1423-0127-19-13                                                                                                                                                                                                                                   |
| Citable link    | http://nrs.harvard.edu/urn-3:HUL.InstRepos:8715417                                                                                                                                                                                                                 |
| 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 |

Mono-allelic retrotransposon insertion addresses epigenetic transcriptional repression in human genome

Hyang-Min Byun1,3*, Kyu Heo2,4, Kasey J Mitchell3 and Allen S Yang1,5

Abstract
Background: Retrotransposons have been extensively studied in plants and animals and have been shown to have an impact on human genome dynamics and evolution. Their ability to move within genomes gives retrotransposons to affect genome instability.

Methods: We examined the polymorphic inserted AluYa5, evolutionary young Alu, in the progesterone receptor gene to determine the effects of Alu insertion on molecular environment. We used mono-allelic inserted cell lines which carry both Alu-present and Alu-absent alleles. To determine the epigenetic change and gene expression, we performed restriction enzyme digestion, Pyrosequencing, and Chromatin Immunoprecipitation.

Results: We observed that the polymorphic insertion of evolutionally young Alu causes increasing levels of DNA methylation in the surrounding genomic area and generates inactive histone tail modifications. Consequently the Alu insertion deleteriously inactivates the neighboring gene expression.

Conclusion: The mono-allelic Alu insertion cell line clearly showed that polymorphic inserted repetitive elements cause the inactivation of neighboring gene expression, bringing aberrant epigenetic changes.

Keywords: Epigenetics, Retrotransposons, Long interspersed elements, Alu

Background
Retrotransposons have been extensively studied in plants and animals and have been shown to have an impact on human genome dynamics and evolution. About 42% of the human genome contains retrotransposons while DNA transposons account for around 2-3% [1-3]. According to the 2001 analysis, which has been confirmed overall by the 2004 update (International Human Genome Sequencing Consortium 2004), short interspersed elements (SINEs), such as Alu or SINE-R/VNTR/Alu (SVA), account for 13%, Long interspersed elements [LINE-1(L1)] for 20%, and long-terminal repeat (LTR) retrotransposons, such as endogenous retrovirus (ERV), for 8%, respectively, of the sequenced human genome. The retrotransposons increase their copy number by retrotransposition via RNA. Attempted or successful retrotranspositions carry a high risk of eliciting chromosome breaks, deletions, translocations, and recombinations [4]. It is estimated that there is one Alu retrotransposon insertion every 21 births [5] during gametogenesis, transferring the retrotransposon’s genetic information to the next generation [6,7]. These retrotransposition events are likely to change the activity of genes at the insertion site, including increased or decreased transcriptional activity. In some cases, this alteration of gene expression causes the development of several diseases or cancers [8]. DNA methylation on the retrotransposon is thought to be the mechanism that controls the retrotransposition rate. Recent vast numbers of publications uniformly address that complex disease, cancer, aging, and environmental challenges are associated with aberrant retrotransposon DNA methylation.

In fact, not all retrotransposons have the capability to retrotranspose to other genomic locations. Currently, most L1s are inactive and cannot retrotranspose to new genomic locations [9], while a small number of human
specific L1 (L1HS) elements remain retrotransposition competent. Retrotransposons seeded in the human genome several million years ago and have many subfamilies defined by distinct patterns of diagnostic base substitutions. Subfamilies may be classified as young, intermediate or old, reflecting the time since the start of retroposition by their members. The expansion of Alu subfamilies (Yc1, Ya5, Ya2, Yb9, Yb8, Y, Sg1, Sx, and J; young to old, respectively) is superimposed on primate evolution. The evolutionally young L1, Alu, and SVA are currently able to transspose in the human genome, hence the ongoing retrotranspositional insertions of the youngest subfamilies are not yet fixed in the human genome and represent polymorphic loci [10]. Some polymorphic insertions are known to be responsible for more than 30 human genetic diseases [11-13]. A genetic polymorphism names as PROGINS has been identified in the progesterone receptor (PGR) gene with insertion of Alu subfamily [14]. The correlations of Alu insertion polymorphism on PGR gene are associated with endometriosis [15,16], ovarian cancer with diethylstilbestrol exposure [17], breast cancer [18], and obesity [19]. Insertional polymorphic retrotransposons are often observed in a mono-allelic fashion, meaning retrotransposons are inserted into only one of the alleles in individuals. For instance, in chromosome 11, the PGR gene has a newly inserted AluYa5 subfamily between exon 5 and 6. In this study, we examine DNA methylation and histone modification of the locus which occurred mono-allelic young Alu, AluYa5 insertions and address the direct effect of retrotransposon in controlling gene expression.

Methods
Nucleic acid isolation and bisulfite treatment
Genomic DNA was isolated by standard proteinase K digestion and phenol-chloroform extraction [20]. Total RNA was collected and extracted from cultured cells with the RNeasy Protect minikit (QIAGEN Inc., Valencia, CA) according to the manufacturer’s recommended protocol. Reverse transcription was performed by using the first strand cDNA synthesis kit (NEB, Beverly, MA, USA). Bisulfite modification of genomic DNA has been described previously [21]. PCR primer sequences for Alu polymorphism with genomic DNA were forward: TTGAGTAAGCCCTCTAAAAAT and reverse: TTCTTGCTAAATCGTGTGTT, and with bisulfite DNA were forward: GAATTTGAAGGAAATAAATATTAGTG and reverse: CATTATTATCCAAAAATATTCTTAC TAA.

Quantitation of allele-specific gene expression by Pyrosequencing
PCR products from genomic DNA or cDNA were used for Pyrosequencing analysis as previously described [21]. Briefly, the PCR product of each gene was used for individual sequencing reactions. Streptavidin-Sepharose beads (Amersham Biosciences) and a Vacuum Prep Tool (Biotage AB) were used to purify the single-stranded biotinylated PCR products according to the manufacturer’s recommendation. The appropriate sequencing primer was annealed to the purified PCR product and used for a Pyrosequencing reaction using the PSQ 96HS system (Biotage). Raw data were analyzed with the allele quantitation algorithm using the PSQ 96 HS software. PCR primer sequences for Alu polymorphism by Pyrosequencing were forward: TTTTCGAAACTACATATTGA, reverse biotin labeled: TTTAGATATAGATCCGTGC, and sequencing primer: GATCTACAAGAA. For allele-specific expression, forward primers: TAGCTCAAGTTCTAATCTATTGC, reverse biotin labeled: TTTAGATTAGTAGTGTCG, and sequencing primer: GATCTCAACAAACA. To validate DNA methylation detection by Pyrosequencing, we designed control oligo for 100% DNA methylation (PSQ-C oligo: 5’-TATTAGATCGACGGGAACCAACGTGTATTGTTTGCT-3’) and 0% DNA methylation (PSQ-T oligo: 5’-TATTAGATCAACGGGGAAACGTGTATTGTTTGCT-3’). The sequencing primer for control oligo is 5’-CAAAGTTTTGCTCCCGT-3’. We mixed PSQ-C oligo (or PSQ-T oligo) with sequencing oligo in PyroMark Annealing Buffer (QIAGEN Inc., Valencia, CA) and performed Pyrosequencing with sequencing entry C/TCGT.

Chromatin Immunoprecipitation
ChIP assays were performed as described previously [22]. Briefly, 25 μg crosslinked protein-DNA complexes were immunoprecipitated using two different histone modification antibodies (H3K9ac, H3K9me3: Millipore) and eluted. Eluted DNA fragments were amplified by PCR. PCR primer sequences for the multiplex PCR reaction were L1 forward: GCCTTTGCGATTGTACCTCAG and reverse: GACGGGTGATTCTTCTGATT, AluY8 forward: GTGGCTCACGGCTGTATCCACG and reverse: GTGCCAGGGCAGGACTGGC, and Alu forward: TGGCTCAGCTTGTGAATCCCGG and reverse: GCCCTCAGGCTGTCAACCG. Analyzing density of gel bands was performed using ImageJ which is a public domain Java image processing program (http://rsb.info.nih.gov/ij/).

Results
Screening of AluYa5 insertional polymorphisms in cell lines
To find insertional polymorphic retrotransposons, we screened Raji, Jurkat, HT15, H1299, MCF, and K562 cell lines using the primer sets listed in the Methods section. The primers flanked the newly inserted retrotransposon AluYa5 in chr11:100,911,358-100,912,065 locus (Assembly:
hg19), thus presence of Alu insertion could be distinguished by length of PCR amplicon. The PCR amplicon with fully inserted Alu generates a 476 bp product, while the amplicon without Alu insertion produces a 150 bp product. Among the cell lines we tested, HT15 and H1299 showed two different sizes of bands after PCR amplification, indicating Alu has inserted in only one allele of the genome locus (Figure 1). MCF and K562 showed insertion of Alu into both alleles (476 bp products). Raji and Jurkat cell lines, however, did not carry an Alu insertion in either allele (150 bp products).

**Alu insertion dependent DNA cytosine methylation**

In order to examine the retrotransposon-derived DNA methylation spreading theory [23], we determined DNA methylation status on the Alu-present and Alu-absent alleles, using the mono-allelic inserted cell lines HT15 and H1299. The PCR amplicon with bisulfite treated DNA was digested with the restriction enzyme HpyCH4III, which cut the 5'-ACNGT-3' region located on the PCR amplicon sequence in only the methylated allele (Figure 2A). Both mono-allelic Alu inserted cell lines, HT15 and H1299, showed partial digestion of only the Alu-present allele, indicating DNA methylation exists in only the Alu inserted allele (Figure 2B). The Alu inserted allele in the H1299 cell line showed slightly more methylation than the Alu inserted allele in the HT15 cell line (Figure 2C). We did not observe digestion of the Alu-absent allele.

**Alu insertion derived inactive histone modification**

To determine whether Alu insertion causes histone tail modifications, we performed ChIP-PCR with two histone modification antibodies against H3K9ac or H3K9me3. Acetylation at Lys-9 on histone H3 (H3K9ac) is an active chromatin marker and often associated with positive gene expression; conversely, methylation at Lys-9 on histone H3 (H3K9me3) is an inactive chromatin marker and correlated with repressed gene expression [24]. After
chromatin immunoprecipitation with the two antibodies for active or repressive histone markers, followed by PCR amplification, we observed differential histone modification between \textit{Alu}-present and \textit{Alu}-absent alleles. The active marker H3K9ac is present in only the \textit{Alu}-absent allele; however, the inactive histone marker H3K9me3 exists in both allele of the genome locus (Figure 3). This difference in histone modification has only happened in young \textit{Alu} subfamilies, not all \textit{Alu} subfamilies. ChIP coupled with PCR amplification of \textit{Alu} \textit{J}, \textit{Alu} \textit{Yb8}, and \textit{L1HS} showed different distributions of histone modifications. \textit{Alu} \textit{J}, the oldest \textit{Alu} subfamily, co-located with both the active marker H3K9ac and the inactive marker H3K9me3. However, the young \textit{Alu} subfamily \textit{Alu} \textit{Yb8} had at least eight times more inactive histone marker H3K9me3. In addition, human-specific \textit{L1HS} did not show a different distribution of active or inactive histone markers (Figure 4).

\textbf{Gene expression repressed by \textit{Alu} insertion in the genome}

To examine differential gene expression in \textit{Alu}-present and \textit{Alu}-absent allele, we developed an allele-specific gene expression detection method using Pyrosequencing. To distinguish between the two alleles, we genotyped the single nucleotide polymorphism (SNP) at chr11:100921952-100922452 (2009 (GRCh37/hg19) assembly), reference SNP ID number is rs1042839, since this SNP is correlated with occurrence of \textit{Alu} insertion \cite{25,26}. To confirm this co-existence, we genotyped this SNP in the six cell lines we worked with and compared with their \textit{Alu} insertion statuses (Table 1). Hetero \textit{Alu} inserted cell lines HT15 and H1299 showed heterozygote C/T, \textit{Alu}-absent cell lines Hep3B2 and HL-60 had a C/C genotype, and \textit{Alu}-present cell lines MCF and K562 had a T/T genotype. We confirmed that the T allele co-exists with \textit{Alu} insertion, while the C allele co-exists with the absence of \textit{Alu} insertion in hetero \textit{Alu}-inserted cell lines. Next, we used this SNP to identify \textit{Alu}-inserted alleles for allele-specific gene expression detection in a Pyrosequencing reaction. After reverse transcription-PCR with mRNA from the hetero \textit{Alu}-inserted cell line H1299, we amplified the locus flanking the SNP to detect each allelic gene expression level (Figure 5). Surprisingly, we observed unequal gene expression levels between \textit{Alu}-present and \textit{Alu}-absent alleles, 10.5% and 89.5% respectively, having an equal distribution of both alleles in the genome (46.7% of \textit{Alu}-present allele vs 53.3% of \textit{Alu}-absent alleles with genotyping data). Thus the presence of \textit{Alu} in the gene body repressed gene expression at the allele containing the \textit{Alu} element.

\textbf{Discussion}

We examined the polymorphic inserted young \textit{Alu}, \textit{Alu} \textit{Ya5}, in the \textit{PGR} gene to determine the effects of \textit{Alu}
insertion on the near gene environment. We used mono-allelic inserted cell lines which carry both Alu-present and Alu-absent alleles. We observed that the polymorphic insertion of evolutionally young Alu causes increasing levels of DNA methylation in the surrounding genomic area and generates inactive histone tail modifications. Consequently, the Alu insertion deleteriously inactivates the neighboring gene expression (Figure 6).

It is a novel approach to address the cis-effects of retrotransposons or retrotransposition in neighboring genomic structures using a mono-allelic inserted young Alu subfamily. These effects were observed in a single cell line system, and virtually all conditions at the particular locus are the same; the only difference being the presence or absence of a retrotransposon insertion. Thus this system bypasses many concerns about experimental artifacts being solely responsible for deducing the function of retrotransposons in the genome.

Generally, our results agree with previous reports that retrotransposons may repress gene expression through an epigenetic mechanism. Our study strongly supports the observations that young active retrotransposons insert in areas that lack cytosine methylation. Retrotransposons spread DNA methylation into neighboring regions, generating repressive histone modifications. It causes a significant inactivation of gene expression. Hollister et al. reported the correlation of transposable elements and gene silencing; however the caveat was that the data do not show whether repetitive elements tend to preferentially insert near lowly expressed genes or...
whether the insertion of repetitive elements causes the low gene expression [27]. However, our mono-allelic inserted cell line system clearly showed that repetitive elements cause the inactivation of neighboring gene expression.

It has been estimated that approximately one out of every 21 births, 212 births, and 916 births has a new insertion of Alu, L1, and SVA retrotransposition, respectively [10]. Thus there is a great deal of retrotransposition in the current human genome. It has been know that evolutionally young repetitive elements have the capability to retrotranspose to other genomic locations. In our study, the

### Table 1 Alu insertion and genotyping in cell lines

|       | Hep3B2 | HL-60 | HT15 | H1299 | MCF   | K562 |
|-------|--------|-------|------|-------|-------|------|
| **Alu** | -/-    | -/-   | +/+  | +/+   | +/+   | +/+  |
| **SNP** | CC     | CC    | TC   | TC    | TT    | TT   |

Figure 4 Chromatin Immunoprecipitation assay to assess differential histone modifications of global L1HS, young Alu; AluYb8 and old Alu; AluJ. Representative gel showing chromatin preparations from immunoprecipitated with anti-IgG (lanes 4), anti-H3K9ac (lane 5), anti-H3K9me3 (lane 6), or the negative control non-immune serum (lanes 3). Lane 2 is non-immunoprecipitated DNA. Lane 1 is DNA size marker (A). The density of the gel band from L1HS, AluYb8, and AluJ was measured (B).
inactive histone modification solely existed in the young Alu subfamily and disappeared in the old Alu J subfamily. Coincidently, global Alu J has less DNA methylation than the young and active Alu Y family (data not shown). Hence this epigenetic difference may promote the mechanism that facilitates transposon mobility. However, what triggers this phenomenon is still not clear, though environmental cues are believed to be responsible for promoting movement of DNA transposons and retrotransposons. DNA methylation on retrotransposons is thought to be an intermediate of the retrotransposition mechanism. We have observed aberrant cytosine methylation patterns on retrotransposons with environmental challenges, but do not know what causes these events or what consequences follow them. Based on the observation of our data, aberrant cytosine methylation on retrotransposons caused by environmental challenges may trigger retrotransposon mobility, slowly reshaping human genome. In the future, it will be necessary to understand the function of other types of retrotransposons of different ages in order to finally resolve the meaning of this aberrant epigenetic phenomenon driven by environmental challenge.

Figure 5 Differential level of PGR gene expression on Alu-absent or -present allele. A) PCR combined with Pyrosequencing shows similar proportion of C (Alu-absent; 53.3%) and T allele (Alu-present; 46.7%) in genomic DNA. Reverse transcription-PCR coupled with Pyrosequencing shows different proportion of C (Alu-absent; 89.5%) and T allele (Alu-present; 10.5%). B) Representative program of genotyping and gene expression.

Figure 6 Summary of molecular signature on mono-allelic Alu insertion. Alu-present allele has cytosine methylation in neighbor CpG and repressive histone modification. Alu-absent allele shows no CpG methylation in near CpG site and has additionally active histone modification, hence show active gene expression.
Conclusions

The mono-allelic Alu insertion cell line clearly showed that polymorphic inserted repetitive elements cause the inactivation of neighboring gene expression, bringing aberrant epigenetic changes.

Author details

Jane Anne Nohl Division of Hematology, Norris Comprehensive Cancer Center, Keck School of Medicine, University of Southern California, Los Angeles, California, USA. 2Department of Biochemistry and Molecular Biology, Norris Comprehensive Cancer Center, University of Southern California, Los Angeles, California, USA. 3Department of Environmental Health, Harvard School of Public Health, Boston, Massachusetts, USA. 4Research Center, Dongnam Institute of Radiological and Medical Sciences, 40 Jwadong-gil, Jangan-eup, Gijang-gun, Busan 619-953, Republic of Korea. 5Amgen, Inc., 1 Amgen Center Drive, MS 38-2-A, Thousand Oaks, California 91320-1799, USA.

Authors' contributions

HMB designed the experiment. KH performed Chromatin immunocytochemistry and HMB performed rest of experiments. HMB, KH, KM, ASY prepared the manuscript and ASY oversaw the research. All authors have read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Received: 16 December 2011 Accepted: 2 February 2012 Published: 2 February 2012

References

1. SanMiguel P, Gaut BS, Tikhonov A, Nakajima Y, Bennetzen JL: The paleontology of intergenic retrotransposons of maize. Nat Genet 1998, 20:43-45.
2. Li W, Zhang P, Fellers JP, Friebe B, Gill BS: Sequence composition, organization, and evolution of the core Triticaceae genome. Plant J 2004, 40:500-511.
3. Lander ES, Linton LM, Birren B, et al: Initial sequencing and analysis of the human genome. Nature 2001, 409:860-921.
4. Syner DE, Connolly CG, Szak ST, Caputo EM, Cost GJ, Parmigiani G, Boeke JD: Human II retrotransposition is associated with genetic instability in vivo. Cell 2002, 110:327-338.
5. Xing J, Zhang Y, Han K, Salem Ah, Sen SK, Huff CD, Zhou Q, Kirkness EF, Levy S, Batzer MA, Jorde LB: Mobile elements create structural variation: analysis of a complete human genome. Genome Res 2009, 19:1516-1526.
6. Branciforte D, Martin SL: Developmental and cell type specificity of LINE-1 expression in mouse testis: implications for transposition. Mol Cell Biol 1994, 14:2584-2592.
7. Ergun S, Buschmann C, Heukeshoven J, Dammann K, Schnieders F, Lauke H, Chalupor F, Kilic N, Stratling WH, Schumann GG: Cell type-specific expression of LINE-1 open reading frames 1 and 2 in fetal and adult human tissues. J Biol Chem 2004, 279:27755-27763.
8. Wolff EM, Byun HM, Han HF, Sharma S, Nichols PW, Siegmund KD, Yang AS, Jones PA, Liang G: Hypomethylation of a LINE-1 promoter activates an alternate transcript of the MET oncogene in bladders with cancer. PLoS Genet 2010, 6:e1000917.
9. Brouha B, Chuskat J, Badge RM, Lutz-Pirgrige S, Farley AH, Moran JV, Kazazian HH Jr: Hot L1s account for the bulk of retrotransposition in the human population. Proc Natl Acad Sci USA 2003, 100:5280-5285.
10. Li X, Scarinche WA, Hill KA, Roberts S, Mengos A, Careni D, Pinto MT, Kasper CK, Sommer SS: Frequency of recent retrotransposition events in the human factor IX gene. Hum Mutat 2001, 17:511-519.
11. Deininger PL, Batzer MA: Alu repeats and human disease. Mol Genet Metab 1999, 67:183-193.
12. Mik Y: Retrotranspositional integration of mobile genetic elements in human diseases. J Hum Genet 1998, 43:77-84.
13. Ostertag EM, Kazazian HH Jr: Biology of mammalian L1 retrotransposons. Annu Rev Genet 2001, 35:501-538.