Evolutionary rate heterogeneity of *Alu* repeats upstream of the *APOA5* gene: do they regulate *APOA5* expression?

Edward A. Ruiz-Narváez · Hannia Campos

Received: 28 August 2007 / Accepted: 22 December 2007 / Published online: 12 January 2008
© The Japan Society of Human Genetics and Springer 2008

Abstract The *APOA5* gene, located in the *APOA1/C3/A4/A5* gene cluster, is a key regulator of triglyceride metabolism. ApoAV plasma concentration is much lower as compared to other apolipoproteins such as apoCIII and apoAI. This is due in part to the fact that the *APOC3*-enhancer, which up-regulates transcription of the *APOA1*, *APOC3*, and *APOA4* genes, does not increase expression of the *APOA5* gene. We postulated that intervening *Alu* repeats in the *APOA5-APOA4* intergenic region might be blocking action of the *APOC3*-enhancer over the *APOA5*-promoter. To search for evidence of functional significance of the intervening *Alu* sequences, we estimated nucleotide substitution rates of 21 pairs of *Alu* elements in the *APOA5-APOA4* intergenic region by comparing published sequences of human and chimpanzee. Also, we scanned the intergenic region for the presence of binding sites of the insulator protein CTCF. Seven out of the nine found CTCF binding sites were located in the first half of the intergenic region. Five out of those seven CTCF binding sites were placed inside *Alu* elements. Based on their substitution rates, we found two clearly defined groups of *Alu* sequences: a slow-evolving group (mean 0.98 ± 0.18) and a fast-evolving group (mean 2.74 ± 0.54). *Alu* repeats with lower substitution rate tended to be located up to 14-kb upstream of the *APOA5* gene, to belong to the oldest *J*-family and to have an opposite orientation to the *APOA5*

gene. Some *Alu* sequences may have functional relevance on the regulation of the *APOA5*-gene expression.

Keywords *APOA5* · *Alu* repeats · Apolipoproteins · Evolution · Substitution rate

Introduction

The *APOA1/C3/A4/A5* gene cluster, located in the human chromosome 11 (Fig. 1), is a key regulator of lipid metabolism (Talmud et al. 2002; Wang et al. 2004; Lai et al. 2005). These apolipoprotein genes show different patterns of tissue-specific expression even when they share *cis*-regulatory elements (Ogami et al. 1990; Ktistaki et al. 1994; Kardassis et al. 1996). In particular, the distal *APOC3*-enhancer can regulate tissue-specific expression of the three *APOA1*, *APOA4*, and *APOC3* genes (Ogami et al. 1990; Ktistaki et al. 1994; Zannis et al. 2001). A recent study found that this common enhancer did not regulate in vivo *APOA5* expression (Gao et al. 2005), which agrees with much lower plasma levels of apoAV compared to other apolipoproteins (O'Brien et al. 2005). For example, plasma apoAV levels (24–406 µg/l) (O'Brien et al. 2005) are 1,000–2,000-fold-lower than human apoCIII plasma concentrations (50–140 mg/l) (Marz et al. 1987; Sakurabayashi et al. 2001).

The reason for the apparent absence of action of the *APOC3*-enhancer over the *APOA5*-promoter is not clear. The distance between the *APOA5*-promoter and the *APOC3*-enhancer is about 35 kb, and it is known that eukaryotic enhancers are able to act over promoters located up to a distance of 100 kb (Zhao and Dean 2005). A possible explanation could be related to the ability of the *APOA5*-promoter to effectively compete for the common

E. A. Ruiz-Narváez (✉) · H. Campos
Department of Nutrition, Room 202, Harvard School of Public Health, 665 Huntington Ave, Boston, MA 02115, USA
e-mail: eruiz@hsph.harvard.edu

E. A. Ruiz-Narváez
Department of Epidemiology, Harvard School of Public Health, Boston, MA, USA

Fig. 1 *APOA1/C3/A4/A5* gene cluster. The human *APOA1/C3/A4/A5* gene cluster spans about 50 kb in the chromosome 11. A common enhancer (*APOC3*-enhancer) regulates transcription of all the genes in the cluster but the *APOA5* gene. Arrows indicate direction of the transcription

enhancer. There are several possible models to explain how enhancers act over large distances (Bondarenko et al. 2003), but regardless of the mechanism the activation signal must travel from the enhancer to the target promoter. This implies that intervening elements between the enhancer and the promoter may interfere with the enhancer action. Because the *APOC3*-enhancer acts over the *APOA4* gene (the last gene of the cluster before *APOA5*), any element interfering with the action over the *APOA5*-promoter must lie in the *APOA5-APOA4* intergenic region.

In humans, the *APOA5-APOA4* intergenic region spans 28 kb and contains 22 full-length *Alu* elements (GenBank accession NT_033899), for an average of one *Alu* element every 1.3 kb. This distance is almost twice as dense as the average in the whole human genome (one *Alu* insert every 2.5 kb) (Mighell et al. 1997). It is noteworthy that in the human chromosome 11 the density of *Alu* elements is one *Alu* insert every 2.7 kb in the whole chromosome, and one *Alu* insert every 3.0 kb in the intergenic regions (Grover et al. 2004). *Alu* elements are repetitive sequences of ≈ 300 nucleotides in length, and they are specific of primate genomes. *Alu* repeats contain a promoter for the RNA-polymerase III (Schmid and Maraia 1992), as well as regulatory elements (for example hormone response elements, HREs) for the RNA-polymerase II (Vansant and Reynolds 1995; Shankar et al. 2004). It has been shown that *Alu* sequences are able to regulate expression of neighboring genes, for example, having enhancer-blocking activity (Willoughby et al. 2000), regulating transcriptional interference (Willoughby et al. 2000), and showing enhancer activity (Landry et al. 2001). We hypothesized that *Alu* elements in the *APOA5-APOA4* intergenic region may affect expression of the *APOA5* gene by interfering with the ability of the *APOC3*-enhancer to act over the *APOA5*-promoter.

If *Alu* sequences in the *APOA5-APOA4* intergenic region have functional importance they will show a slower evolutionary rate of nucleotide change as compared to other *Alu* elements without functional significance. By using a published sequence of the *APOA1/C3/A4/A5* gene cluster of chimpanzee (*Pan troglodytes*) and comparing to the gene cluster of humans, we estimated substitution rates of the *Alu* elements in the *APOA5-APOA4* intergenic region. According to our working hypothesis, significant differences in evolutionary rates will be found if some *Alu* sequences have functional relevance.

If *Alu* repeats were interfering with the action of the *APOC3*-enhancer on the *APOA5*-promoter, insulator-like elements must lie inside those *Alu* sequences. In vertebrates, many insulators require the binding of the zinc-finger protein CTCF to exert their action (Bell et al. 1999; Felsenfeld et al. 2004; Gaszner and Felsenfeld 2006). We scanned the *APOA5-APOA4* intergenic region for the presence of potential CTCF binding sites and assessed whether those sites reside inside *Alu* elements.

Materials and methods

Alu sequences

Nucleotide sequences of the *Alu* elements in the *APOA5-APOA4* intergenic region of humans were obtained from the GenBank accession NT_033899 and from the accession NW_001222312 for chimpanzee. The positions of the *Alu* sequences were determined using the RepeatMasker software (<http://www.repeatmasker.org>) (Smith et al. 1996–2004). We further retrieved the sequences of *Alu* elements 100-kb upstream and 100-kb downstream from the *APOA5* promoter, with the aim of determining how the nucleotide substitution rates of *Alu* inserts within the *APOA4-APOA5* intergenic region compare to the substitution rates of *Alu* elements in the genomic neighborhood.

Statistical analysis

Pairwise alignment of *Alu* sequences and estimation of evolutionary distances were performed using MEGA version 3.1 (Kumar et al. 2004). Substitution rates (R) were estimated as $R = D/2T$, where D is the number of substitutions per site since the time of divergence between humans and chimps. The Tamura–Nei (Tamura and Nei 1993) distance was used to estimate D because it allows for differences of substitution rates between nucleotides and does not assume equality of nucleotide frequencies. The divergence time (T) between humans and chimps was assumed to be 6 million years (Chen et al. 2001). We

scanned the *APOA5-APOA4* intergenic region, using the MotifScanner software of the TOUCAN package (Aerts et al. 2003), for the presence of CTCF binding sites. We used the position weight matrices (PWMs) recently derived by Xie et al. (2007).

The PROC CLUSTER in the Statistical Analysis Systems (SAS) software version 9 (SAS Institute Inc., Cary, NC) was used to identify potential groups of *Alu* elements with similar substitution rates. The PROC GENMOD in the SAS software was used to estimate least square means of explanatory variables on substitution rates. The explanatory variables used were: (1) distance from the transcription-start of the *APOA5* gene [≤ 14 vs. >14 kb] and (2) type of *Alu*-family (oldest *J*-family vs. more recent *S* and *Y* families) and (3) orientation of each *Alu* element relative to the orientation of the *APOA5* transcription.

Results

Out of the 22 full-length *Alu* elements in the human *APOA4-APOA5* intergenic region, 21 were retrieved from the contig NW_001222312 of chimpanzee. The most proximal *Alu* sequence to the transcription-start of the *APOA5* gene was not complete in the contig from chimpanzee; therefore, it was not included in the analysis. Figure 2 shows substitution rates (per site per 10^9 years) of the 21 pairs of *Alu* elements along the *APOA5-APOA4* intergenic region. A great heterogeneity on the substitution rates was observed, from a minimum of 0.59 ± 0.41 to a maximum of 3.56 ± 0.99 . It is noteworthy that *Alu* elements with low substitution rates tended to be located in the first 14 kb upstream of the *APOA5* transcription start site. In particular, *Alu* sequences with the lowest substitution rates are located between 10- and 14-kb upstream from the *APOA5* gene.

Nine CTCF binding sites were predicted in the *APOA4-APOA5* intergenic region (Fig. 2). Seven out of the nine CTCF binding sites were located in the first 14 kb of the intergenic region, and five out of those seven CTCF binding sites (indicated by arrows in Fig. 2) were placed inside *Alu* elements. The remaining two CTCF binding sites in the second half of the intergenic region were located outside of *Alu* repeats.

Cluster analysis showed two clearly defined groups of *Alu* elements (Table 1). The first cluster contains *Alu* sequences with a low substitution rate (mean 0.98 ± 0.18), while the second group includes *Alu* elements with a higher substitution rate (mean 2.74 ± 0.54). Compared to the low substitution rate group, the high substitution rate group had an excess of *Alu* sequences with the same orientation as the *APOA5* gene (5 out of 11 vs. 9 out of 10, respectively, $P = 0.06$, exact test). The oldest *Alu-J* family tended to

Fig. 2 Substitution rates of *Alu* elements and predicted CTCF binding sites along the *APOA5-APOA4* intergenic region. The Tamura–Nei distance was used to estimate the number of substitutions per site since time of divergence between humans and chimps. Time of divergence was assumed to be 6 million years. Note the clustering of slow-evolving *Alu* repeats in between 10- and 14-kb upstream of the *APOA5* gene. Note the clustering of predicted CTCF binding sites in the first half of the intergenic region. Arrows indicate CTCF binding sites placed inside *Alu* elements

occur more frequently in the low substitution rate group (6 out of 11) compared to the high substitution rate group (2 out of 10) ($P = 0.18$, exact test). Finally, slow-evolving *Alu* sequences tended to be located in the first 14-kb upstream of the transcription-start of the *APOA5* gene (8 out of 11) as compared to fast-evolving *Alu* sequences (3 out of 10) ($P = 0.09$, exact test).

The same trends were observed using general linear models. Table 2 shows least square means of the substitution rates by *Alu*-element orientation, *Alu*-family type, and upstream distance from the transcription-start of the *APOA5* gene. In univariate models, *Alu* sequences with the same orientation of the *APOA5* gene had a higher substitution rate as compared to *Alu* sequences with an opposite orientation ($P = 0.02$). The oldest *Alu-J* family showed lower substitution rates as compared to the more recent *Alu-S* and *Alu-Y* families ($P = 0.04$). And *Alu* elements in the first 14-kb upstream of the transcription-start of the *APOA5* gene had lower substitution rates relative to the substitution rates of the *Alu* elements in the second half of the *APOA5-APOA4* intergenic region ($P = 0.06$). To isolate independent effects of each explanatory variable, we used multivariate general linear models. After controlling each variable for the other two variables, the strongest predictor of the substitution rates was the orientation of the *Alu* elements ($P = 0.04$) followed by *Alu*-family type ($P = 0.04$). Effect of position was attenuated after adjusting for orientation and *Alu*-family type ($P = 0.23$).

Table 1 Cluster analysis of the *Alu* elements in the *APOA5-APOA4* intergenic region

Upstream distance (kb) from <i>APOA5</i> gene	Orientation ^a	Family	Substitution rate per site per 10 ⁹ years (SE)
Cluster 1			
6.3	—	Sx	1.09 (0.53)
8.4	—	Jb	1.18 (0.59)
9.8	+	Jb	1.06 (0.52)
10.4	—	Sx	0.84 (0.46)
11.4	+	Jo	1.10 (0.55)
12.6	—	Jo	0.59 (0.41)
13.1	+	Jo	1.07 (0.52)
13.8	—	Sx	1.11 (0.54)
18.0	+	Sx	0.80 (0.45)
18.6	—	Jb	1.10 (0.52)
24.6	+	Sx	0.85 (0.46)
Cluster 2			
4.3	+	Sp	3.17 (0.97)
5.6	+	Sx	2.91 (0.89)
9.2	+	Sx	2.15 (0.74)
14.7	+	Jb	2.16 (0.75)
15.4	—	Sg	2.50 (0.79)
17.6	+	Sx	2.14 (0.75)
19.7	+	Y	3.20 (0.95)
23.0	+	Jb	2.20 (0.78)
23.3	+	Sx	3.41 (1.04)
26.8	+	Sq	3.56 (1.00)

+ Codes for the same direction, — codes for the opposite direction

^a Orientation relative to the transcription direction of the *APOA5* gene

At last, we retrieved all the *Alu* elements in the 200-kb region around the *APOA5* promoter (100-kb downstream and 100-kb upstream from the transcription start of the *APOA5* gene) to estimate the mean substitution rate of *Alu* inserts in the genomic neighborhood of the *APOA5-APOA4* intergenic region. A total of 85 *Alu* elements, 21 inside and 64 outside of the *APOA5-APOA4* intergenic region, were retrieved. Figure 3 shows the distribution of the substitution rates in the 200-kb genomic region that has an overall mean of 1.81 ± 0.12 substitutions per site per 10^9 years. This overall mean is almost twice the mean of the slow-evolving group in the *APOA5-APOA4* intergenic region (0.98 ± 0.18). The lowest substitution rate in the intergenic region (0.59 ± 0.41) is observed around 12.6-kb upstream of the *APOA5* transcription start, and we determined how extreme this value is in the overall distribution. We found that just 6 out of the 85 *Alu* elements (7.0%) in the 200 kb genomic region have substitution rates equal to or lower than 0.59 (Fig. 3).

Table 2 Least square means of the substitution rates of the *Alu* elements in the *APOA5-APOA4* intergenic region

Variable	Substitution rate per site per 10 ⁹ years (SE)	<i>P</i> value
Univariate analysis		
Upstream distance (kb) from <i>APOA5</i> gene		
≤14 kb (<i>n</i> = 11)	1.48 (0.27)	0.06
>14 kb (<i>n</i> = 10)	2.19 (0.27)	
Orientation ^a		
+ (<i>n</i> = 14)	2.13 (0.23)	0.02
− (<i>n</i> = 7)	1.20 (0.32)	
Family		
<i>S</i> and <i>Y</i> (<i>n</i> = 13)	2.13 (0.24)	0.04
<i>J</i> (<i>n</i> = 8)	1.31 (0.30)	
Multivariate analysis ^b		
Upstream distance (kb) from <i>APOA5</i> gene		
≤14 kb (<i>n</i> = 11)	1.41 (0.23)	0.23
>14 kb (<i>n</i> = 10)	1.82 (0.27)	
Orientation ^a		
+ (<i>n</i> = 14)	2.00 (0.21)	0.04
− (<i>n</i> = 7)	1.24 (0.29)	
Family		
<i>S</i> and <i>Y</i> (<i>n</i> = 13)	1.97 (0.22)	0.04
<i>J</i> (<i>n</i> = 8)	1.26 (0.27)	

+ Codes for the same direction, — codes for the opposite direction

^a Orientation relative to the transcription direction of the *APOA5* gene

^b Multivariate models include distance, orientation and family in the same model

Discussion

Even though it is known that expression of the *APOA5* gene is regulated through different mechanisms (Prieur et al. 2003; Genoux et al. 2005; Nowak et al. 2005), it remains to be determined why plasma concentration of apoAV is substantially lower compared to concentrations of other apolipoproteins such as apoAI and apoCIII. A recent study showing that the *APOC3*-enhancer that affects expression of the *APOA1*, *APOC3*, and *APOA4* genes does not up-regulate transcription of the *APOA5* gene (Gao et al. 2005) provided some insight into this issue, but several points remain uncertain. For example, it is unclear why the *APOC3*-enhancer is not able to act upon the *APOA5*-promoter that is just 35 kb away, when it is known that eukaryotic enhancers may exert their effects up to a distance of 100-kb (Zhao and Dean 2005). Therefore, it is possible that the presence of intervening elements hinder the action of the *APOC3*-enhancer over the *APOA5*-promoter, when it is known that enhancer-

Fig. 3 Substitution rates of *Alu* elements in the 200-kb region around the *APOA5* promoter. A total of 85 *Alu* elements were retrieved from the genomic region 100-kb downstream and 100-kb upstream of the *APOA5* promoter. The Tamura–Nei distance was used to estimate the number of substitutions per site since the time of divergence between humans and chimps. Time of divergence was assumed to be 6 million years. The overall mean of the substitution rates is 1.81 substitutions per site per 10⁹ years as compared to just 0.98 substitutions per site per 10⁹ years in the slow-evolving *Alu* group in the *APOA5-APOA4* intergenic region. It is noteworthy that just 6 out of the 85 *Alu* elements showed an equal or lower substitution rate than the lowest substitution rate in the *APOA5-APOA4* intergenic region (0.59 substitutions per site per 10⁹ years) as indicated by the asterisk

action requires travel of the activation signal from the enhancer to the target promoter (Bondarenko et al. 2003).

We examined the *APOA5-APOA4* intergenic region that extends over 28 kb in the *APOA1/C3/A4/A5* gene cluster to identify evidence of *cis*-regulatory elements that may be interfering with the action of the *APOC3*-enhancer. In humans, the *APOA5-APOA4* intergenic region is rich in *Alu* elements, with an average density of one *Alu* insertion every 1.3 kb as compared to a whole-genome average of one *Alu* insertion every 2.5 kb (Mighell et al. 1997). Because *Alu* elements are able to affect gene expression (Vansant and Reynolds 1995; Britten 1996; Willoughby et al. 2000), we hypothesized that some of the *Alu* sequences in the *APOA5-APOA4* intergenic region may affect expression of the *APOA5* gene, perhaps by interfering with the ability of the *APOC3*-enhancer to act over the *APOA5*-promoter. If this hypothesis were true then *Alu* elements with functional relevance will tend to have a smaller nucleotide substitution rate as compared to *Alu* elements with no function. Our data show that *Alu* elements with a lower substitution rate were located in the first 14 kb of the *APOA5-APOA4* intergenic region, were more likely to have an opposite orientation to the *APOA5*-gene transcription and tended to belong to the oldest *Alu-J* family. Although our results do not prove the functionality

of the analyzed *Alu* sequences, they are highly suggestive of such function. Most importantly, they identify a likely location of the *Alu* elements with functional relevance, namely the first 14 kb upstream from the transcription start of the *APOA5* gene. This conclusion was further supported when we retrieved all the *Alu* sequences in the 200-kb genomic region surrounding the *APOA5* promoter. First, the mean substitution rate of the slow-evolving *Alu* group in the *APOA5-APOA4* intergenic region was just half the overall mean in the local genomic region. Also, just 6 out of the 85 *Alu* elements in the 200 kb genomic region showed an equal or lower substitution rate than the lowest value in the *APOA5-APOA4* intergenic region.

To further locate potential regulatory regions in the *APOA5-APOA4* intergenic region, we scanned this region for the presence of binding sites of the insulator protein CTCF. We found nine CTCF binding sites, and most of them (seven out of the nine) were located up to 14 kb from the transcription start of the *APOA5* gene. Also, five out of the nine CTCF binding sites were placed within *Alu* elements. Together, these findings suggest that insulator elements might be located in the first 14 kb of the *APOA5-APOA4* intergenic region. Because slow-evolving *Alu* elements tend to be co-located with the predicted CTCF binding sites, the present results suggest that the same selective constraints are operating in the *APOA5-APOA4* intergenic region in both humans and chimpanzees. More important, these findings suggest that the mode of action of the *APOC3*-enhancer inside the *APOA1/C3/A4/A5* gene cluster is conserved at least in great apes.

Functional experiments are required to determine whether *Alu* elements in the *APOA5-APOA4* intergenic region do affect expression of the *APOA5* gene by interfering with the action of the *APOC3*-enhancer as well as the possible mechanisms for the interference. A likely mechanism is that some of the *Alu* elements in the intergenic region act as silencers of the transcription of the *APOA5* gene. This hypothesis is further supported by the presence of CTCF binding sites inside some *Alu* elements in the intergenic region. In fact, it has been found that *Alu* sequences are able to act as transcriptional silencers in the WT1 (Hewitt et al. 1995), BRCA (Sharan et al. 1999), human epsilon-globin (Wu et al. 1990), and human growth-hormone (Trujillo et al. 2006) genes among others. It is noteworthy that silencers may act over long distances, for example the silencer of the WT1 gene is located about 12 kb from the promoter (Hewitt et al. 1995), and we found that the most conserved *Alu* elements in the *APOA5-APOA4* were located in between 10- and 14-kb upstream of the *APOA5* promoter.

In summary, we found heterogeneity on the substitution rates of the *Alu* elements in the *APOA5-APOA4* intergenic region. The presence of a group of *Alu* sequences with

lower substitution rates (almost threefold lower as compared to the higher substitution group), as well as the occurrence of CTCF binding sites inside *Alu* sequences, suggests that some of the *Alu* elements may have functional relevance. If confirmed by further experiments, our results would have major implications in the study of the effect of the *APOA5* gene on triglyceride metabolism and risk of coronary heart disease. Most of the research on the regulation of expression the *APOA5* has been focused in a few kilobases upstream of the promoter. The present results suggest that elements up to 14 kb from the transcription start of the *APOA5* gene may also regulate *APOA5*-gene expression.

Acknowledgments This study was supported by a grant of the Dean's Initiative Fund for Bioinformatics Research of Harvard School of Public Health. E.R.N. is a Roadmap Fellow in Harvard School of Public Health supported by grants T90 DK070078 and R90 DK071507 from the National Institutes of Health.

References

- Aerts S, Thijs G, Coessens B, Staes M, Moreau Y, De Moor B (2003) Toucan: deciphering the *cis*-regulatory logic of coregulated genes. *Nucleic Acids Res* 31:1753–1764
- Bell AC, West AG, Felsenfeld G (1999) The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* 98:387–396
- Bondarenko VA, Liu YV, Jiang YI, Studitsky VM (2003) Communication over a large distance: enhancers and insulators. *Biochem Cell Biol* 81:241–251
- Britten RJ (1996) DNA sequence insertion and evolutionary variation in gene regulation. *Proc Natl Acad Sci USA* 93:9374–9377
- Chen FC, Vallender EJ, Wang H, Tzeng CS, Li WH (2001) Genomic divergence between human and chimpanzee estimated from large-scale alignments of genomic sequences. *J Hered* 92:481–489
- Felsenfeld G, Burgess-Beusse B, Farrell C, Gaszner M, Ghirlando R, Huang S, Jin C, Litt M, Magdinier F, Mutskov V, Nakatani Y, Tagami H, West A, Yusufzai T (2004) Chromatin boundaries and chromatin domains. *Cold Spring Harb Symp Quant Biol* 69:245–250
- Gao J, Wei Y, Huang Y, Liu D, Liu G, Wu M, Wu L, Zhang Q, Zhang Z, Zhang R, Liang C (2005) The expression of intact and mutant human apoA1/CIII/AIV/AV gene cluster in transgenic mice. *J Biol Chem* 280:12559–12566
- Gaszner M, Felsenfeld G (2006) Insulators: exploiting transcriptional and epigenetic mechanisms. *Nat Rev Genet* 7:703–713
- Genoux A, Dehondt H, Helleboid-Chapman A, Duhem C, Hum DW, Martin G, Pennacchio LA, Staels B, Fruchart-Najib J, Fruchart JC (2005) Transcriptional regulation of apolipoprotein A5 gene expression by the nuclear receptor ROR α . *Arterioscler Thromb Vasc Biol* 25(6):1186–1192
- Grover D, Mukerji M, Bhatnagar P, Kannan K, Brahmachari SK (2004) Alu repeat analysis in the complete human genome: trends and variations with respect to genomic composition. *Bioinformatics* 20:813–817
- Hewitt SM, Fraizer GC, Saunders GF (1995) Transcriptional silencer of the Wilms' tumor gene WT1 contains an Alu repeat. *J Biol Chem* 270:17908–17912
- Kardassis D, Laccotripe M, Talianidis I, Zannis V (1996) Transcriptional regulation of the genes involved in lipoprotein transport. The role of proximal promoters and long-range regulatory elements and factors in apolipoprotein gene regulation. *Hypertension* 27:980–1008
- Ktistaki E, Lacorte JM, Katrakili N, Zannis VI, Talianidis I (1994) Transcriptional regulation of the apolipoprotein A-IV gene involves synergism between a proximal orphan receptor response element and a distant enhancer located in the upstream promoter region of the apolipoprotein C-III gene. *Nucleic Acids Res* 22:4689–4696
- Kumar S, Tamura K, Nei M (2004) MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief Bioinform* 5:150–163
- Lai CQ, Parnell LD, Ordoas JM (2005) The APOA1/C3/A4/A5 gene cluster, lipid metabolism and cardiovascular disease risk. *Curr Opin Lipidol* 16:153–166
- Landry JR, Medstrand P, Mager DL (2001) Repetitive elements in the 5' untranslated region of a human zinc-finger gene modulate transcription and translation efficiency. *Genomics* 76:110–116
- Marz W, Schenk G, Gross W (1987) Apolipoproteins C-II and C-III in serum quantified by zone immunoelectrophoresis. *Clin Chem* 33:664–669
- Mighell AJ, Markham AF, Robinson PA (1997) Alu sequences. *FEBS Lett* 417:1–5
- Nowak M, Helleboid-Chapman A, Jakel H, Martin G, Duran-Sandoval D, Staels B, Rubin EM, Pennacchio LA, Taskinen MR, Fruchart-Najib J, Fruchart JC (2005) Insulin-mediated down-regulation of apolipoprotein A5 gene expression through the phosphatidylinositol 3-kinase pathway: role of upstream stimulatory factor. *Mol Cell Biol* 25:1537–1548
- O'Brien PJ, Alborn WE, Sloan JH, Ulmer M, Boodhoo A, Knierman MD, Schultze AE, Konrad RJ (2005) The novel apolipoprotein A5 is present in human serum, is associated with VLDL, HDL, and chylomicrons, and circulates at very low concentrations compared with other apolipoproteins. *Clin Chem* 51:351–359
- Ogami K, Hadzopoulou-Cladaras M, Cladaras C, Zannis VI (1990) Promoter elements and factors required for hepatic and intestinal transcription of the human ApoCIII gene. *J Biol Chem* 265:9808–9815
- Prieur X, Coste H, Rodriguez JC (2003) The human apolipoprotein AV gene is regulated by peroxisome proliferator-activated receptor- α and contains a novel farnesoid X-activated receptor response element. *J Biol Chem* 278:25468–25480
- Sakurabayashi I, Saito Y, Kita T, Matsuzawa Y, Goto Y (2001) Reference intervals for serum apolipoproteins A-I, A-II, B, C-II, C-III, and E in healthy Japanese determined with a commercial immunoturbidimetric assay and effects of sex, age, smoking, drinking, and Lp(a) level. *Clin Chim Acta* 312:87–95
- Schmid C, Maraia R (1992) Transcriptional regulation and transpositional selection of active SINE sequences. *Curr Opin Genet Dev* 2:874–882
- Shankar R, Grover D, Brahmachari SK, Mukerji M (2004) Evolution and distribution of RNA polymerase II regulatory sites from RNA polymerase III dependant mobile Alu elements. *BMC Evol Biol* 4:37
- Sharan C, Hamilton NM, Parl AK, Singh PK, Chaudhuri G (1999) Identification and characterization of a transcriptional silencer upstream of the human BRCA2 gene. *Biochem Biophys Res Commun* 265:285–290
- Smith AFA, Hubley R, Green P (1996–2004) RepeatMasker Open-3.0
- Talmud PJ, Hawe E, Martin S, Olivier M, Miller GJ, Rubin EM, Pennacchio LA, Humphries SE (2002) Relative contribution of variation within the APOC3/A4/A5 gene cluster in determining plasma triglycerides. *Hum Mol Genet* 11:3039–3046

- Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol* 10:512–526
- Trujillo MA, Sakagashira M, Eberhardt NL (2006) The human growth hormone gene contains a silencer embedded within an Alu repeat in the 3'-flanking region. *Mol Endocrinol* 20(10):2559–2575
- Vansant G, Reynolds WF (1995) The consensus sequence of a major Alu subfamily contains a functional retinoic acid response element. *Proc Natl Acad Sci USA* 92:8229–8233
- Wang QF, Liu X, O'Connell J, Peng Z, Krauss RM, Rainwater DL, VandeBerg JL, Rubin EM, Cheng JF, Pennacchio LA (2004) Haplotypes in the APOA1-C3-A4-A5 gene cluster affect plasma lipids in both humans and baboons. *Hum Mol Genet* 13:1049–1056
- Willoughby DA, Vilalta A, Oshima RG (2000) An Alu element from the K18 gene confers position-independent expression in transgenic mice. *J Biol Chem* 275:759–768
- Wu J, Grindlay GJ, Bushel P, Mendelsohn L, Allan M (1990) Negative regulation of the human epsilon-globin gene by transcriptional interference: role of an Alu repetitive element. *Mol Cell Biol* 10:1209–1216
- Xie X, Mikkelsen TS, Gnirke A, Lindblad-Toh K, Kellis M, Lander ES (2007) Systematic discovery of regulatory motifs in conserved regions of the human genome, including thousands of CTCF insulator sites. *Proc Natl Acad Sci USA* 104:7145–7150
- Zannis VI, Kan HY, Kritis A, Zanni EE, Kardassis D (2001) Transcriptional regulatory mechanisms of the human apolipoprotein genes in vitro and in vivo. *Curr Opin Lipidol* 12:181–207
- Zhao H, Dean A (2005) Organizing the genome: enhancers and insulators. *Biochem Cell Biol* 83:516–524