

RESEARCH NOTE

Open Access



# A human specific Alu DNA cassette is found flanking the genes of transcription factor AP2

Hamdi K. Hamdi<sup>\*</sup> , Siddana Reddy, Nada Laz, Renad Eltaher, Zahraa Kandell, Teif Mahmud, Lamia Alenazi, Basheer Haroun, Mohanad Hassan and Raju Ragavendra

## Abstract

**Objective:** Alu elements are retroposons that invaded the primate genome and shaped its biology. Some Alus inserted recently and are polymorphic in the human population. It is these Alus that are being sought after in disease association studies and regulatory biology. Discovering polymorphic Alus in the human genome can open areas of new research in these fields.

**Results:** Using the polymerase chain reaction on genomic DNA, we identified a polymorphic Alu in the flanking region of the TFAP2B and TFAP2D genes. The new insert was found in higher frequency in Europeans (0.4) and Asians (0.38) and lower frequency in Africans (0.25). We also show this Alu to be part of a 3 Alu cassette that is human specific. The TFAP2B and TFAP2D genes encode members of the transcription factor AP-2, which plays a role in organ development. The insertion of this Alu cassette flanking the transcription factor genes distinguishes humans from the primates. This cassette can possibly affect the regulation of both genes or alternately provoke genomic deletions, which we have shown in this study. Its presence in such a location is intriguing and unquestionably opens an investigational window in disease association studies and in the field of gene regulation.

**Keywords:** Alu, Polymorphic, TFAP2B, TFAP2D, AP2, Dental, Neural, Breast cancer

## Introduction

The Transcription Factor AP2 Beta and Delta (TFAP2B and TFAP2D) genes are located on chromosome 6 and encode members of the AP-2 protein family [1, 2]. AP-2 is a transcription factor that binds directly to DNA either as a homo or hetero dimer regulating gene expression [3]. They display differential binding affinities for different gene promoters and function in both inducing and repressing transcription [4–6]. The AP-2 protein has been shown to be expressed in various tissues including adrenal, esophagus, kidney, placenta, salivary gland, skin and testis [7]. Mutations in these genes have also been discovered in patients with Char syndrome (TFAP2B) [8, 9]. More recently it was shown that mutations in TFAP2B are linked with isolated tooth agenesis, microdontia, supernumerary tooth and tooth root maldevelopment [10]. It has also been shown that the TFAP2B

gene product is tumor promoting in breast cancer and appears to be highly expressed in breast cancer cell lines and in tumor tissues of breast cancer patients [11]. Lobular carcinoma in situ and invasive lobular breast cancer also demonstrated enhanced expression of transcription factor AP-2 [12]. In this paper, we present a polymorphic Alu (*Arthrobacter luteus* restriction enzyme characterized) element insert found flanking both the TFAP2B and TFAP2D genes in humans.

Alu inserts are DNA elements that have invaded the primate genome by retroposition and are now permanent residents in it with over a million interspersed in the human genome [13, 14]. They have contributed to disease by inserting in coding regions, upsetting gene structure [15, 16]. Since coding sequences are less abundant, Alus litter the landscape of non-coding DNA. This contributes to their recruitment for regulatory functions [17, 18] as in the case of the angiotensin converting enzyme (ACE) Alu [19]. Depending on when these Alus have inserted, they can be traced in lineage by their shared descent in the primates at various loci [20]. Alus that are shared

\*Correspondence: 282hamdi@gmail.com  
Basic Medical Sciences Dept., College of Dentistry, Almustqbal University,  
PO Box 156, Buraidah, Qassim 51411, Saudi Arabia



between the primate species are used in the phylogenetic studies of primates [20]. Some Alus are of a more recent origin and can be found to be polymorphic in the human population [14, 21, 22]. It is these Alus that are sought after in disease association studies. Thus, the aim of this study is to identify polymorphic Alu elements in the human population, which can be subsequently used in disease association studies. In this paper, we identified a polymorphic Alu element (HK1) flanking the TFAP gene family.

## Main text

### Methods

#### Genomic DNA

Genomic DNA used in this study was obtained from human primary cell cultures, which have been previously described [19]. The primary cultures of human endothelial cells came from a variety of tissues, which were purchased from the American Type Culture Collection (ATCC). The cells came from the following tissues of origin: umbilical vein; pulmonary vein; iliac vein; abdominal aorta; femoral artery; iliac artery; and pulmonary artery.

#### Polymerase chain reaction (PCR)

Total genomic DNA (0.3 µg) was amplified by polymerase chain reaction (PCR) in a 30 µl volume, which included: 20 mM Tris-HCl (pH 8.3), 50 mM KCl, 2.5 mM MgCl<sub>2</sub>, 250 µM dNTP, 0.2 µM each of forward (5'-GCAGTC CAAACTCGGAAGACTGGG-3') or reverse (5'-GGA CAGAAAGAG GCCTGGACTAC-3') and 5% dimethylsulfoxide (DMSO). Amplification of DNA was performed at 96 °C for 2 min, followed by 35 cycles of denaturation (96 °C for 30 s), annealing (62 °C for 30 s), and elongation (68 °C for 30 s). The last cycle was followed by a final elongation step (65 °C for 5 min). The reactions were stopped by the addition of 3 µl of a 0.25 M EDTA solution at 0 °C, and aliquots were electrophoresed on a 2% (w/v) agarose gel with 0.5 µg/ml ethidium bromide. The resulting bands were scored for the presence of the Alu and sequenced to confirm their identity. The native (empty site) allele was identified by the presence of a single 271 bp PCR product. The insertion allele (Alu) was identified by the presence of a single 582 bp PCR product.

#### Frequency analysis of the HK1 polymorphic Alu in Genbank's genomic assemblies

To determine the Allele frequencies for this locus, we screened Genbank's human genomic database. The Basic local alignment search tool (BLAST) was used to screen the HK1 locus in both complete and partial genomic assemblies [23]. Aligned sequences were included in the data set and filtered to remove redundant alignments from assemblies derived from the same BioSample. Using

the BLAST tool as a virtual polymerase chain reaction (vPCR) we screened Genbank assemblies for the HK1 locus.

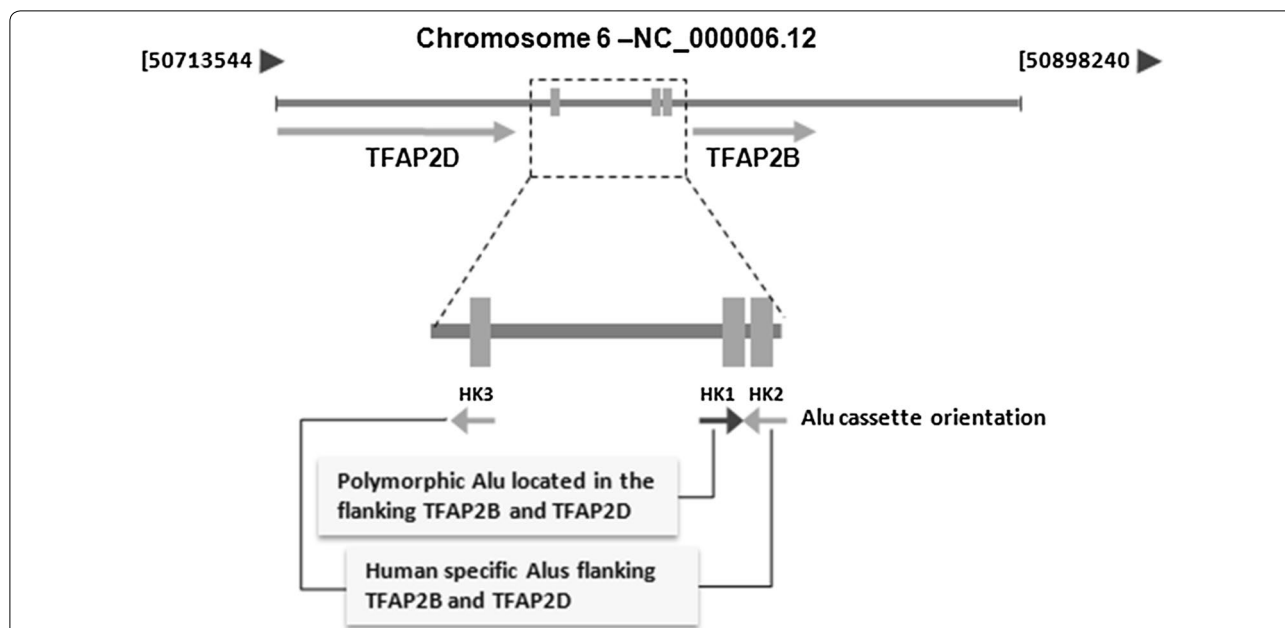
## Results

Using Genbank's map viewer we were able to localize the polymorphic Alu (HK1) in the flanking region of the TFAP2B and TFAP2D genes (Fig. 1). Our PCR amplifications and sequencing yielded 2 alleles for this genomic site. These included the insertion allele and that of the native site. Most Alu elements are retroposons and thus have the hallmark of their RNA origin. Newly inserted Alus have a long poly A tail and are flanked by direct repeats on each end, which is a feature of their insertion at a staggered cut in the DNA. The Alu element we identified was flanked by a poly A tail and direct repeats (Fig. 2). The native site was found in all non-human primates (Fig. 2). The sequences for both alleles were deposited with GenBank under accession numbers: MK574933 and MK574934.

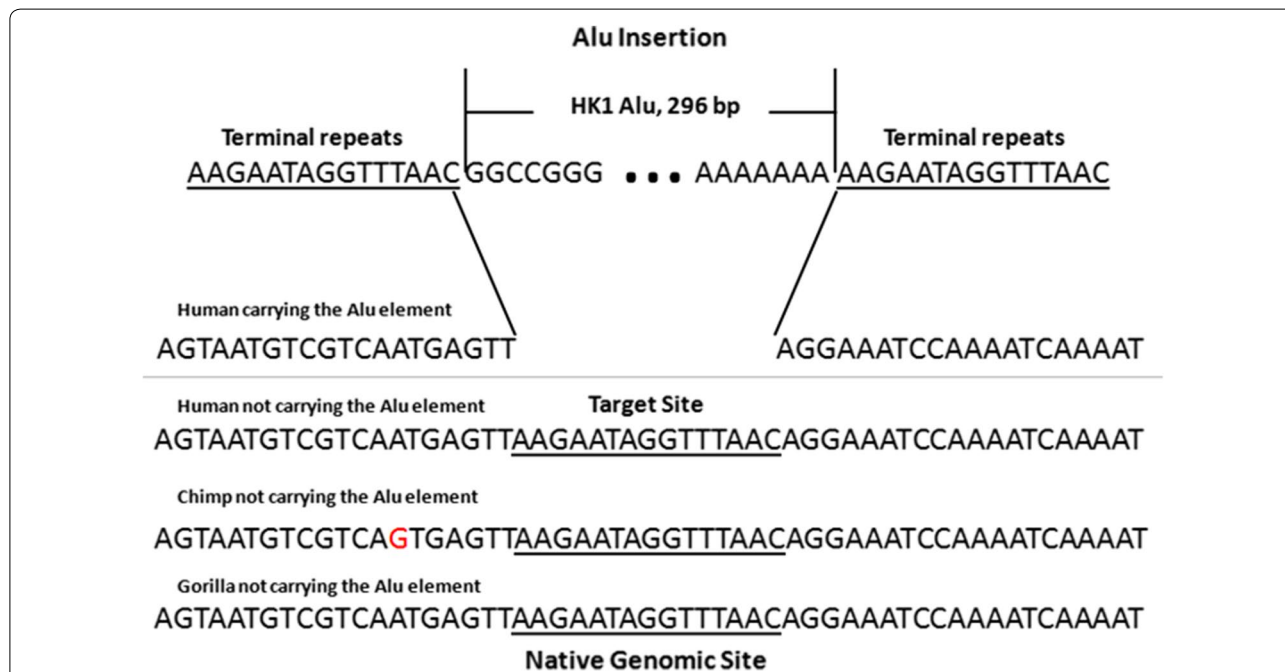
While the goal of this study was not to assess the frequency of the insert in human populations, we conducted a preliminary screen of the polymorphic Alu in Genbank genomic assemblies. It was found in a higher frequency in European (n=10, 0.40) and Asian populations (n=16, 0.38) and at a much lower frequency in Africans (n=16, 0.25). The complete data set is presented as Additional file 1. These numbers, however, should be taken with care due to the limitation in sample size. Although the functional role of this Alu remains to be elucidated, we have found it to be flanked by other Alu elements in reverse orientation forming a cassette of Alus (Fig. 3). Both of the flanking Alus (HK2 and HK3) are human specific and are not found in the primates (Genbank genomic database). The 3 Alus in this cassette are highly homologous, displaying 96% identity and have the potential for intrastrand complementation and secondary structure formation (Fig. 3). Perhaps this configuration can lead to enhancer activity affecting the regulation of TFAP2B and TFAP2D or alternately can play a role in provoking genomic deletions. In fact, we detected such deletions in 2 of Genbank's genomic assemblies (GCA\_000005465.1 and GCA\_003391555.1). The deletion in GCA\_000005465.1 was bracketed and involved the HK1 and HK2 Alus. While in GCA\_003391555.1 a partial deletion of the HK1 Alu was observed.

## Discussion

Alu invasions occurred en masse in the distant past. More recently, however, there is a steady but perhaps slower spread in the collective primate genome [21]. Alus have been studied ever since the advent of genomics, because in fact they make up a significant chunk of the genome.



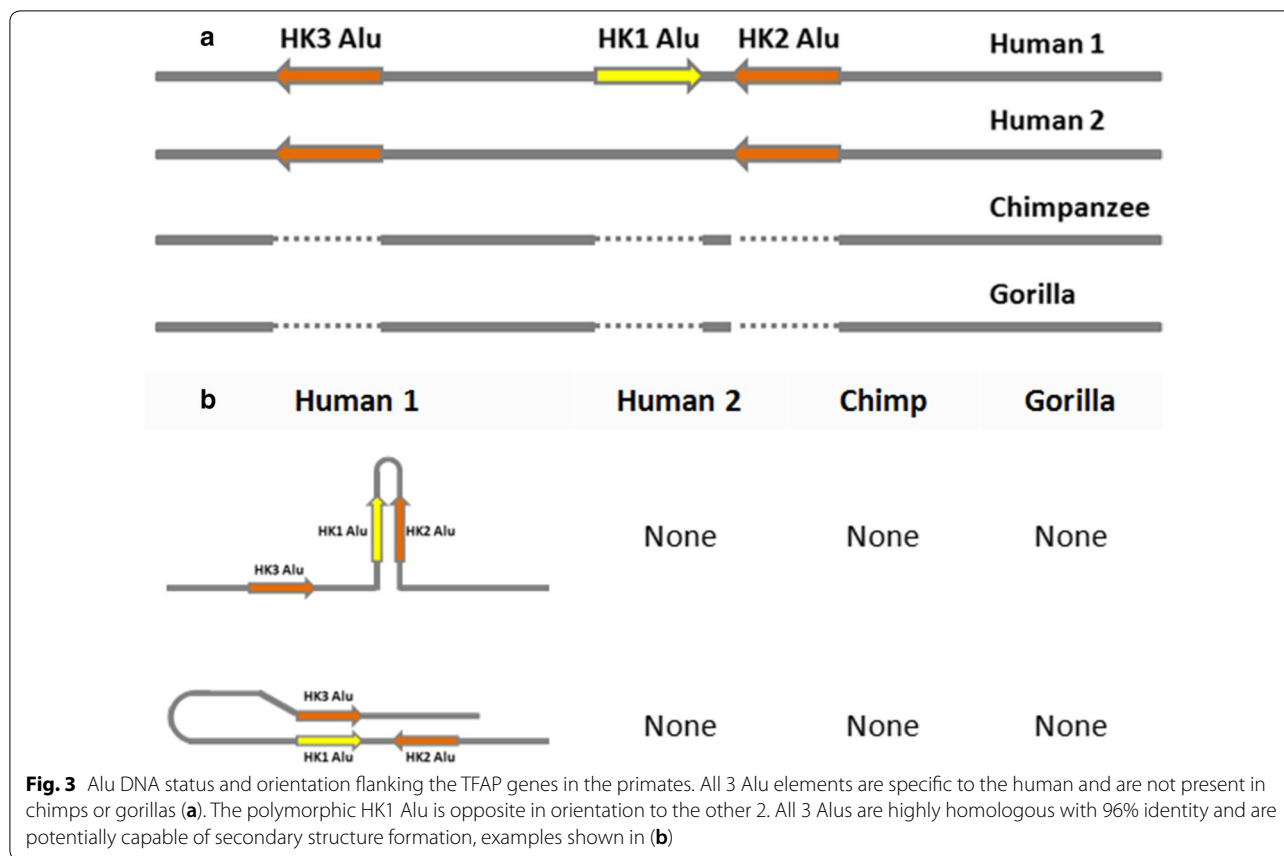
**Fig. 1** Mapping of the polymorphic HK1 Alu. The polymorphic Alu (HK1) is located in the flanking region of TFAP2B and TFAP2D genes. The HK1 Alu is flanked by 2 other Alu elements (HK2 and HK3) that are human specific and in opposite orientation to HK1



**Fig. 2** Sequence analysis of the HK1 polymorphic Alu insertion. Shown are aligned sequences from the region surrounding the unoccupied (native genomic site) in humans, Chimp and Gorilla and occupied (HK1 Alu) Alu insertion site in humans. Terminal repeats and target site are underlined

In the beginning, unfortunately, they were labeled as junk DNA to distinguish them from the genic functional parts of the DNA and this has driven their use simply as time markers for studying primate evolution [14, 20]. Over the

past decade, however, Alu elements have been recognized as genomic engineers in their own right and their study shifted to the realm of biological function [17, 18, 24, 25]. It has also become obvious that evolutionary change



is not limited to the 2% difference in sequence homology between the higher primates but is in fact driven by micro and macro engineering of genomic structure [26]. To distinguish the rapidity of human change from the rest of the primates, it is pertinent to identify other mechanisms of evolutionary change besides that of the single nucleotide polymorphism (SNP) [20]. Here we show a 3 Alu cassette present in the flanking region of TFAP2B and TFAP2D genes. These Alus are all human specific with one being polymorphic in the human population and in opposite orientation to the others. They display high homology to one another having the potential to change local genomic configuration by the formation of DNA secondary structure. The potential for forming this genomic configuration, however, is not available to all humans, since the polymorphic HK1 Alu is a requisite for this formation. Despite the fact that humans carry the HK2 and HK3 Alus, the genomic configuration is not formed unless the cassette is completed with the HK1 Alu, which only some humans carry. In some ways, we can view this genetic system at 3 hierarchical levels. The first level is represented by the primates, which lack all together this cassette and have no potential in accessing this particular genomic configuration. The second

level is represented by those humans that carry the HK2 and HK3 Alus, and we can consider them evolutionarily “primed” for this genomic configuration. Finally, humans that carry the full 3 Alu cassette can be considered evolutionarily in an “active state” fully accessing this genomic configuration with all of its positive and negative outcomes. It is in this category that we have discovered genomic deletion events. This discovery sheds light on the continuing dynamic changes occurring in the human genome. It has been shown previously that Alu elements can influence gene expression by either binding transcription factors, forming chromatin looping or secondary structures and even behaving as enhancer elements inducing or repressing transcription [18, 27]. Since they are widely dispersed in the genome, they have the ability to exert powerful regulation on a local as well as a global level. Perhaps the saltatory argument of evolution comes within reach by considering the changes that might occur in the regulation of genes that code for transcription factors. Since transcription factors have a global effect on gene regulation by impacting the expression of hundreds of genes, a small genomic change in their regulation can have an amplifying effect on biology. The insertion of this Alu cassette flanking the transcription factor genes

produces a change in the human genome distinguishing humans from primates. If this genomic change leads to a specific human change in the regulation of the AP2 transcription factor (regulating the regulator), this can have an amplifying effect on many genes. The flip side of the argument would be the deleterious effect of Alu elements. Perhaps the presence of the Alu cassette we described might lead to genomic rearrangements/deletions as a result of Alu recombination events resulting in disease outcomes. In this paper, we have seen such deletions; their impact on human biology, however, remains to be elucidated. Regardless of the implied biological role for this Alu cassette, this discovery opens research opportunities in various fields including gene regulation as well as disease association studies.

### Conclusion

By identifying strategically located polymorphic Alu sequences in the human genome, we hope to open a doorway to better understand the etiology of human disease and perhaps the driving force of human evolution. After all, one of the most studied genetic polymorphisms in the history of genetic association studies is the ACE ID polymorphism, which is a polymorphic Alu element present in the angiotensin converting enzyme gene [28]. This Alu polymorphism alone has led to over 3700 publications (as of February 2019) and has been shown to associate with varied diseases in the human population [28]. Here we present a newly inserted Alu element flanking the human TFAP2B and TFAP2D genes and like the ACE Alu is also polymorphic in the human population. More interestingly, however, the HK1 Alu is part of a 3 Alu human specific cassette with the potential of reconfiguring chromatin structure.

### Limitations

Although we have found a polymorphic Alu element flanking the genes TFAP2B and 2D, the allelic frequencies were based on a small set of alleles and should be treated with caution. A larger population study of this polymorphic Alu should be conducted across a wide section of the human population to better approximate the allelic frequencies.

### Additional file

**Additional file 1: Table S1.** Distribution of HK1 polymorphic Alu in GenBank Genomic Assemblies.

### Abbreviations

ACE ID: angiotensin converting enzyme insertion-deletion polymorphism; TFAP2B: transcription factor AP2 beta gene; TFAP2D: transcription factor AP2 delta gene; ATCC: American Type Culture Collection.

### Authors' contributions

HKH helped to conceive and design the study, design the primers, interpret data, and write the manuscript. ND, RE, ZK, TM, LA, BH, MH, RR, and SD participated in mapping the polymorphic element, cell culture, genomic DNA extraction, performing PCR and sequencing analyses. All authors read and approved the final manuscript.

### Acknowledgements

We thank Dr. Ali Al-Saif and Dr. Khalid Al-Johani for their support and encouragement in this project. We would also like to thank Almustqbal University-College of Dentistry for supporting this project and providing us with laboratory space and facilities. HKH would also like to extend his eternal gratitude to the late professor Dr. Achilles Dugaiczkyk.

### Competing interests

The authors declare that they have no competing interests.

### Availability of data

The data presented here was deposited with GenBank under the accession numbers: MK574933 and MK574934.

### Consent for publication

Not applicable.

### Ethics approval and consent to participate

Not applicable.

### Funding

Research was partially funded by HKH. The College of Dentistry at Almustqbal University provided the laboratory space and facilities. No external funding was received for this research.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 2 March 2019 Accepted: 1 April 2019

Published online: 11 April 2019

### References

- Eckert D, Buhl S, Weber S, Jäger R, Schorle H. The AP-2 family of transcription factors. *Genome Biol.* 2005;6:246.
- Williamson JA, Boshier JM, Skinner A, Sheer D, Williams T, Hurst HC. Chromosomal mapping of the human and mouse homologues of two new members of the AP-2 family of transcription factors. *Genomics.* 1996;35(1):262–4.
- Moser M, Imhof A, Pscherer A, Bauer R, Amselgruber W, Sinowatz F, et al. Cloning and characterization of a second AP-2 transcription factor: AP-2 beta. *Development.* 1995;121(9):2779–888.
- Duan C, Clemmons DR. Transcription factor AP-2 regulates human insulin-like growth factor binding protein-5 gene expression. *J Biol Chem.* 1995;270(42):24844–511.
- Gaubatz S, Imhof A, Dosch R, Werner O, Mitchell P, Buettner R, et al. Transcriptional activation by Myc is under negative control by the transcription factor AP-2. *EMBO J.* 1995;14(7):1508–19.
- Ikeda K, Maegawa H, Ugi S, Tao Y, Nishio Y, Tsukada S, et al. Transcription Factor Activating Enhancer-binding Protein-2 $\beta$ . *J Biol Chem.* 2006;281(42):31245–53.
- Fagerberg L, Hallström BM, Oksvold P, Kampf C, Djureinovic D, Odeberg J, et al. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol Cell Proteom.* 2014;13(2):397–406.
- Zhao F, Weismann CG, Satoda M, Pierpont MEM, Sweeney E, Thompson EM, et al. Novel TFAP2B mutations that cause char syndrome provide a genotype-phenotype correlation. *Am J Hum Genet.* 2001;69(4):695–703.



9. Nyboe D, Kreiborg S, Darvann T, Dunø M, Nissen KR, Hove HB. A study of familial Char syndrome involving the TFAP2B gene with a focus on facial shape characteristics. *Clin Dysmorphol*. 2018;27(3):1.
10. Tanasubsinn N, Sittiwangkul R, Pongprot Y, Kawasaki K, Ohazama A, Sastraruji T, et al. TFAP2B mutation and dental anomalies. *J Hum Genet*. 2017;62(8):769–75. <https://dx.doi.org/10.1038/jhg.2017.37>.
11. Li Z, Xu X, Luo M, Hao J, Zhao S, Yu W, et al. Activator protein-2 $\beta$  promotes tumor growth and predicts poor prognosis in breast cancer. *Cell Physiol Biochem*. 2018;47(5):1925–35.
12. Raap M, Gronewold M, Christgen H, Glage S, Bentires-Alj M, Koren S, et al. Lobular carcinoma in situ and invasive lobular breast cancer are characterized by enhanced expression of transcription factor AP-2 $\beta$ . *Lab Invest*. 2018;98(1):117–29.
13. Schmid CW, Jelinek WR. The Alu family of dispersed repetitive sequences. *Science*. 1982;216(4550):1065–70.
14. Batzer MA, Deininger PL. Alu repeats and human genomic diversity. *Nat Rev Genet*. 2002;3(5):370–9.
15. Happ H, Weh E, Costakos D, Reis LM, Semina EV. Case report of homozygous deletion involving the first coding exons of GCNT2 isoforms A and B and part of the upstream region of TFAP2A in congenital cataract. *BMC Med Genet*. 2016;17(1):64.
16. Ganguly A, Dunbar T, Chen P, Godmilow L, Ganguly T. Exon skipping caused by an intronic insertion of a young Alu Yb9 element leads to severe hemophilia A. *Hum Genet*. 2003;113(4):348–52.
17. Chen L, Yang L. Review ALU ternative regulation for gene expression. *Trends Cell Biol*. 2017;27(7):480–90.
18. Su M, Han D, Boyd-kirkup J, Yu X, Han JJ. Report evolution of alu elements toward enhancers. *Cell Rep*. 2014;7(2):376–85. <https://doi.org/10.1016/j.celrep.2014.03.011>.
19. Hamdi HK, Castellon R. A genetic variant of ACE increases cell survival: a new paradigm for biology and disease. *Biochem Biophys Res Commun*. 2004;318(1):187–91.
20. Hamdi H, Nishio H, Zielinski R, Dugaiczky A. Origin and phylogenetic distribution of alu dna repeats: irreversible events in the evolution of primates. *J Mol Biol*. 1999;289(4):861–71.
21. Carroll ML, Roy-Engel AM, Nguyen SV, Salem A-H, Vogel E, Vincent B, et al. Large-scale analysis of the Alu Ya5 and Yb8 subfamilies and their contribution to human genomic diversity. *J Mol Biol*. 2001;311(1):17–40.
22. Salem A-H, Kilroy GE, Watkins WS, Jorde LB, Batzer MA. Recently Integrated Alu Elements and Human Genomic Diversity. *Mol Biol Evol*. 2003;20(8):1349–61.
23. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman D. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. 1997;25:3389–402.
24. Hamdi HK, Nishio H, Tavis J, Zielinski R, Dugaiczky A. Alu-mediated phylogenetic novelties in gene regulation and development. *J Mol Biol*. 2000;299(4):931–9.
25. Buckley RM, Kortschak RD, Adelson DL. Divergent genome evolution caused by regional variation in DNA gain and loss between human and mouse. *PLoS Comput Biol*. 2018;14:1–36.
26. Polak P, Domany E. Alu elements contain many binding sites for transcription factors and may play a role in regulation of developmental processes. *BMC Genomics*. 2006;15:1–15.
27. Dridi S. Alu mobile elements : from junk DNA to genomic gems. *Scientifica (Cairo)*. 2012;2012:545328. <https://doi.org/10.6064/2012/545328>.
28. Castellon R, Hamdi HK. Demystifying the ACE polymorphism: From genetics to biology. *Curr Pharm Des*. 2007;13(12):1191–8.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

