

RESEARCH ARTICLE

Open Access

Human-specific protein isoforms produced by novel splice sites in the human genome after the human-chimpanzee divergence

Dong Seon Kim and Yoonsoo Hahn*

Abstract

Background: Evolution of splice sites is a well-known phenomenon that results in transcript diversity during human evolution. Many novel splice sites are derived from repetitive elements and may not contribute to protein products. Here, we analyzed annotated human protein-coding exons and identified human-specific splice sites that arose after the human-chimpanzee divergence.

Results: We analyzed multiple alignments of the annotated human protein-coding exons and their respective orthologous mammalian genome sequences to identify 85 novel splice sites (50 splice acceptors and 35 donors) in the human genome. The novel protein-coding exons, which are expressed either constitutively or alternatively, produce novel protein isoforms by insertion, deletion, or frameshift. We found three cases in which the human-specific isoform conferred novel molecular function in the human cells: the human-specific IMUP protein isoform induces apoptosis of the trophoblast and is implicated in pre-eclampsia; the intronization of a part of *SMOX* gene exon produces inactive spermine oxidase; the human-specific NUB1 isoform shows reduced interaction with ubiquitin-like proteins, possibly affecting ubiquitin pathways.

Conclusions: Although the generation of novel protein isoforms does not equate to adaptive evolution, we propose that these cases are useful candidates for a molecular functional study to identify proteomic changes that might bring about novel phenotypes during human evolution.

Keywords: Splice site, Human genome, Protein variant, Molecular evolution

Background

Humans have many unique traits compared with those of other primates that must be derived from genetic changes acquired during human evolution [1]. These genetic modifications include accelerated amino acid substitutions [2,3], *de novo* origin of protein-coding genes from non-coding sequences [4], formation of novel transcript variants by DNA insertion [5], and in-activation of long-established genes [6,7].

The generation of novel alternative splice sites plays a role in the evolution of gene structure [8,9]. Alternative usage of splice sites often results in insertion or deletion of amino acids and/or frameshift in proteins. Generation of novel splice sites or activation of cryptic splice sites in

transposable elements such as Alu repeats is rather common in human genes [10-13]. These sites are often alternatively spliced, as the novel Alu splice sites are generally weak donors or acceptors and/or the insertion of Alu-derived fragments into the coding region of genes generally results in disruption of the host proteins [14,15]. There are reports on novel exons originating from non-coding intronic sequences in some organisms such as rodents and humans [16,17]. However, these are not species-specific but have been originated during rodent and primate evolution, respectively; the novel "human" exons reported by Zhang and Chasin [17] originated before the human-chimpanzee divergence.

Although many reports are available on novel exons originating during human evolution, most of them have no evidence of protein-coding capability or are not human-specific. In this study, we hypothesized that nucleotide changes in the human genome after the human-

* Correspondence: hahnyc@cau.ac.kr
Department of Life Science, Research Center for Biomolecules and Biosystems, Chung-Ang University, Seoul 156-756, Korea

chimpanzee divergence may have generated novel splice sites and produced novel protein-coding exons. To find such cases, we analyzed annotated human protein coding exons and their orthologous genomic sequences of other primates including chimpanzee and some non-primate mammals. We examined possible changes in the proteins caused by the formation of novel splice sites.

Results and discussion

Human-specific splice sites

We analyzed multiple alignments of human protein-coding exons and their respective orthologous sequences from various primate and mammalian genomes. We identified 50 canonical splice acceptors (AG) and 35 canonical splice donors (GT) that had newly arisen in the human genome after the human-chimpanzee divergence. Lists of the novel splice acceptors and donors reported in this study are presented in Additional file 1: Table S1 and in Additional file 2: Table S2, respectively.

In the present study, we aimed to collect highly plausible cases for generation of novel protein-coding exons induced by human-specific splice sites. To achieve this goal, we used annotated protein-coding exons as the initial data set, excluding exons without annotated coding regions such as alternative exons that were obtained from expressed sequence tag (EST) data analysis. We also employed highly stringent filtering conditions for subsequent analyses. For example, we discarded exons that were supported by only a single transcript record, because these exons are likely to be derived from noisy splicing events [18]. We also excluded exons that were derived from repetitive elements, although many novel exons were reported to be derived from repetitive elements such as Alu or L1 [5,10-13]. Repetitive elements are frequently exapted as coding exons during evolution [19]; however, the vast majority of repeats that appeared within the coding region of mRNAs cause a frameshift or a premature termination codon [15]. Therefore, we only focused on novel coding exons of which their origin was from a non-repetitive genomic region.

In this study, we only collected canonical splice sites (GT-AG) although there are many exons flanked by non-canonical minor splice sites such as GC-AG or AT-AC [18]. In the early phase of this study, we collected all the human splice junction sequences that were different from the chimpanzee sequences as candidates for human-specific splice sites. We found that the most of non-canonical human splice sites were not one of known minor splice sites but random sequences. We assumed this was due to intrinsic errors in the input data such as incomplete genome sequence, incorrect alignment, and non-orthologous alignment. Because we aimed to collect highly accurate cases, we only considered canonical splice sites.

In all the cases except one, the human-specific splice sites reported in this study were associated with genes conserved in mammals. Only one gene (*C14orf182*) appeared to be a *de novo* protein-coding gene derived from a non-coding sequence. Out of 85 derived human-specific exons, 42 were annotated as a part of a RefSeq transcript. Utilization of novel splice sites results in various modifications such as splice site shift of ancestral exons, generation of novel exons (exonization) or intronization of exonic segments. We classified the protein modification types into 19 categories, which are presented in Figure 1. The detailed information on representative cases is shown in Figure 2 and Additional file 3: Figures S1–S9. Multiple sequence alignments of all the novel splice sites and associated exons are presented in Additional file 4: Table S3 (splice acceptor sites) and in Additional file 5: Table S4 (splice donor sites).

In 18 cases, the human-specific splice sites caused a frameshift compared with the ancestral open reading frame. A frameshift mutation often results in premature termination codon and induces nonsense-mediated mRNA decay (NMD) [20]. Generally, a stop codon which is situated at 50–55 nucleotides upstream of the last exon-exon junction can induce NMD. We looked at the stop codon position of the 18 human-specific transcripts and found that in all but one case the stop codon was in the last exon. In one case, the *IMUP* gene transcript variant, the stop codon is in an internal exon. However, it is very close to the last exon-exon junction (2 nucleotides upstream) and would not trigger NMD (see Figure 3A). Therefore, none of the human-specific transcripts in this study is susceptible to NMD, suggesting the human-specific transcripts are stable in the human cells.

In nine cases, we found a splice site-associated polymorphism in the human population, indicating that these human-specific transcripts are differentially expressed in human individuals. The frequencies of the derived allele that generated a human-specific splice site were: *PLP1* (dbSNP accession number rs2233697, AG vs GG), 98.804%; *TXNDC16* (rs28759013, AG vs AA), 96.350%; *KHK* (rs74537742, AG vs GG), 95.256%; *IL12RB1* (rs393548, AG vs GG), 80.378%; *HOXD1* (rs13390932, GT vs GC), 75.443%; *XRCC4* (rs1805377, AG vs AA), 65.843%; *NIPAL2* (rs3735887, GT vs AT), 46.445%; *MRE11A* (rs496797, GT vs AT), 46.374%; and *STXBP4* (rs11658717, AG vs AA), 20.947%.

Splice site shift and type change

The generation of a novel splice site close to the ancestral site of an exon causes a splice site shift that either lengthens or shortens the affected exon (category codes A1–A4 and D1–D4; see Figure 1 for details). Fifty-six of 85 novel splice sites resulted in shifting of splice sites

A. splice acceptors

category		count
ancestral form		
A1	shift; increase; inframe	21
A2	shift; increase; frameshift	3
A3	shift; decrease; inframe	16
A4	shift; decrease; frameshift	0
A5	exonization; inframe	2
A6	exonization; frameshift	4
A7	novel start	2
A8	intronization; inframe	1
A9	intronization; frameshift	1
	<i>total</i>	50

B. splice donors

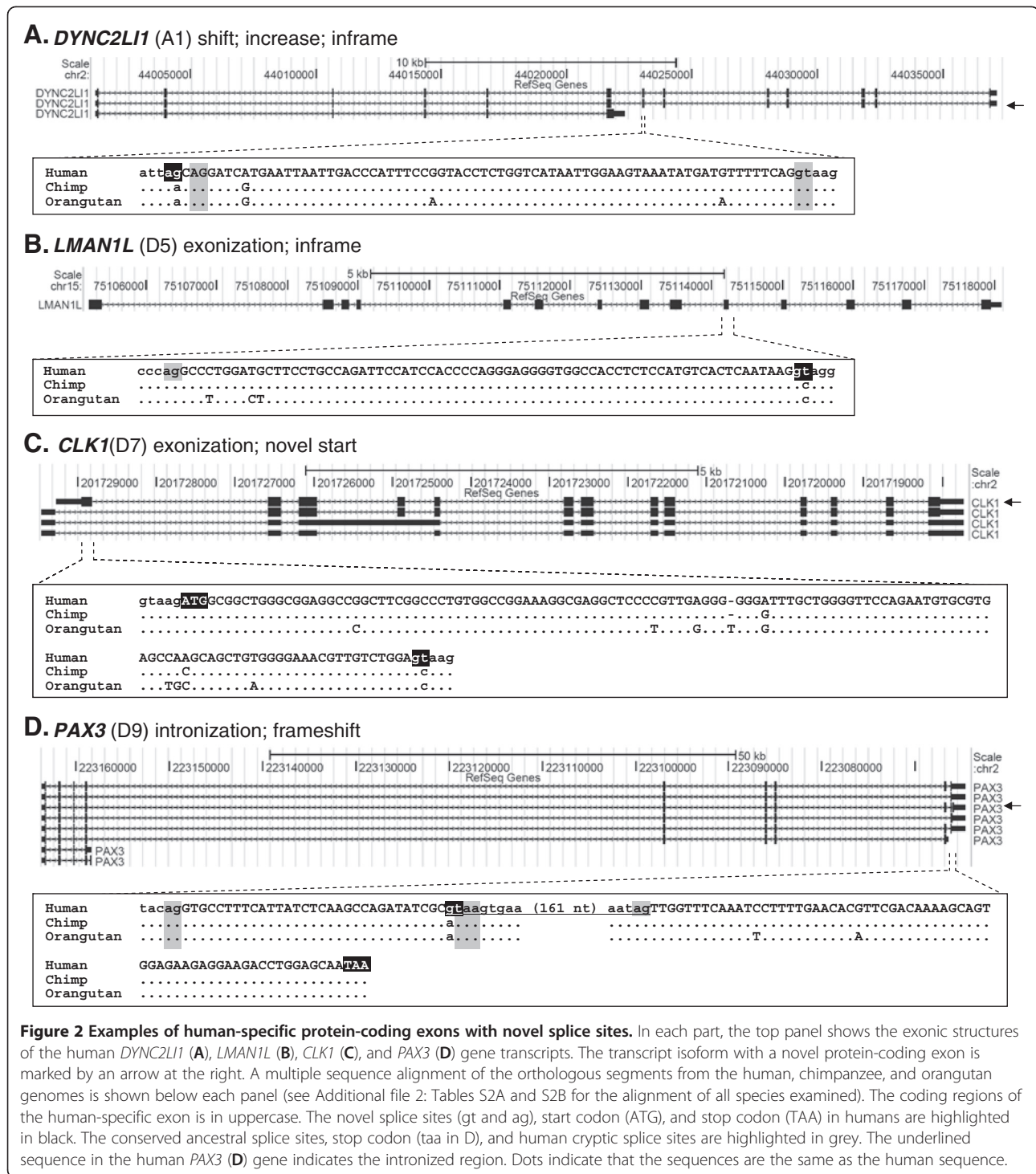
category		count
ancestral form		
D1	shift; increase; inframe	5
D2	shift; increase; frameshift	1
D3	shift; decrease; inframe	7
D4	shift; decrease; frameshift	3
D5	exonization; inframe	3
D6	exonization; frameshift	4
D7	exonization; novel start	8
D8	intronization; inframe	1
D9	intronization; frameshift	2
D10	type change (GC to GT)	1
	<i>total</i>	35

Figure 1 Summary of splice site generation and protein modification. This figure summarizes the categories of transcript structures modified by the generation of novel splice acceptors (A) and donors (B) in human proteins. The left column shows a summary of the modification with the category codes A1–D10. The center column shows a schematic of the transcript and coding region structures. Novel splice sites are marked by asterisks. The ancestral coding regions (black), derived coding regions (grey), alternative coding regions (hatched), and untranslated regions (white) are represented by boxes with different patterns. The last column “count” shows the number of cases reported in this study.

(see Figure 2A and Additional file 3: Figures S1–S4 for representative examples). In 50 cases, the derived exon increased or decreased in size without changing the downstream reading frame.

In 20 cases, the splice site was shifted by three bases, producing tandem splice sites known as NAGNAG acceptors and GTNGTN donors [21–23]. Alternative usage of these splice sites generated single amino acid insertion and deletion isoforms. We found 16 NAGNAG acceptors and four GTNGTN donors (see Figure 2A and Additional file 3: Figures S1 and S2). Usually, the

addition or deletion of a single codon does not affect protein function. However, when a functionally important residue is removed or a nonsense codon is destroyed or added, the result can be substantial and cause diseases in humans [21,24]. There was no premature stop codon in the cases reported in this study. The TassDB2 (<http://www.tassdb.info>) is a comprehensive database of tandem splice sites in human and mouse [25]. Out of 20 tandem splice sites, 18 were recorded in the TassDB2. However, it is not previously reported that these splice sites are human-specific.

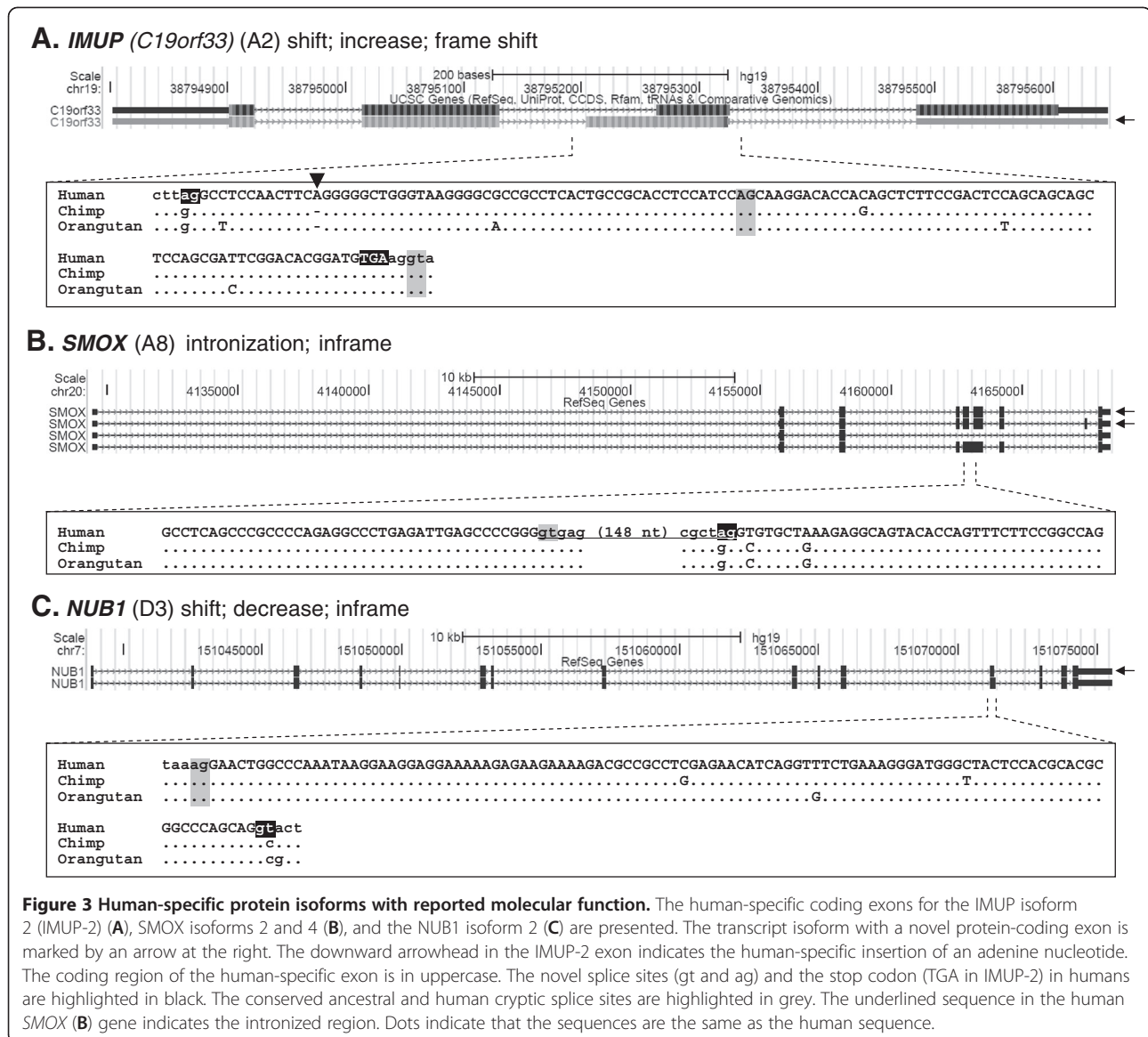


We found a GC-to-GT change in the splice donor of exon 45 of the human *DOCK1* gene (see Additional file 3: Figure S9). However, the orthologous mammalian exon used GC as the splice donor so that the exon boundaries are conserved in human and mammals. Thus, the ancestral GC-AG intron had changed to a canonical GT-AG intron in humans. Although both

the GC-AG and GT-AG splice sites are processed by the standard U2-type spliceosome, the strength of the donor site could be affected [26].

Exonization and intronization

In 21 cases, the novel splice sites (six acceptors and 15 donors) arose in the non-coding region resulting in



exonization of intronic segments (categories A5–A9 and D5–D9; see Figure 1 for details). A cryptic intronic splice donor for a given novel acceptor or a cryptic acceptor for a novel donor was activated in the human genome. The 13 cases generated novel internal exons that added some additional amino acids (five cases) or produced alternative C-termini by frameshift (eight cases) (see Figure 2B and Additional file 3: Figures S5 and S6). In the remaining eight cases, the exons were associated with cryptic or authentic promoters and produced alternative N-termini of the proteins (see Figure 2C and Additional file 3: Figure S7).

Interestingly, in five cases, novel splice acceptors or donors appeared in the exons, and corresponding cryptic exonic donors or acceptors were accordingly activated. As a result, part of the exonic sequence was

spliced out as an intron (see Figure 2D and Additional file 3: Figure S8).

Novel molecular function of human-specific protein isoforms

We performed a literature review to find reported molecular functions of human-specific protein isoforms produced by human-specific splice sites. We found three cases (*IMUP*, *SMOX*, and *NUB1*) of which the molecular function of the human-specific protein isoform has been described.

The human *IMUP* (also known as *C19orf33*) gene encoding immortalization-up-regulated protein produces two protein isoforms reported as *IMUP*-1 and *IMUP*-2, respectively (<http://www.uniprot.org/uniprot/Q9GZP8>). Both isoforms are highly expressed in cancer cells and localized in the nucleus [27,28]. These two protein

variants share 46 N-terminal amino acids but have different C-termini. The IMUP-2 protein (isoform 2 of IMUP) is a human-specific isoform that is produced by using a human-specific splice donor site (Figure 3A). IMUP-2 expression is specifically elevated during preterm pre-eclampsia and under hypoxic conditions, and the IMUP-2 protein induces apoptosis of the trophoblast [29]. Therefore, production of the IMUP-2 human-specific isoform may be functionally involved in placental development and gynecological diseases such as pre-eclampsia. This case clearly shows that a human-specific protein isoform induced by the generation of a novel splice site sequence plays a role in the human cells and is associated with a phenotype.

The human *SMOX* gene encoding spermine oxidase is implicated in various tumors and diseases in humans [30-32]. Alternative splicing of the human *SMOX* gene transcripts results in the production of six protein isoforms (<http://www.uniprot.org/uniprot/Q9NWM0>). A part of exon 5 can act as an intron and is spliced out to produce *SMOX* protein isoforms 2 and 4 [33]. We found that this intronization of an internal part of exon 5 was induced by the generation of a human-specific splice donor sequence (Figure 3B). This intronization removed 53 amino acids from the ancestral form. Many of these residues are highly conserved components of the FAD-binding domain and are essential for catalysis. Therefore, the deletion caused by the intronization seemed to result in the production of catalytically inactive spermine oxidase protein [33,34]. The human-specific *SMOX* protein isoform 4 has been localized in the nucleus [34]. Although the human-specific *SMOX* isoforms are inactive as oxidases, it is possible that they could confer other molecular functions in the human cell in general, or that production of non-functional isoform could decrease the cellular concentration of active spermine oxidase and affect spermine catabolism.

The human *NUB1* gene encodes two isoforms of a negative regulator of ubiquitin-like proteins 1 (originally known as NEDD8 ultimate buster-1) [35]. Isoforms 1 (also known as NUB1L for larger variant) and 2 are 615-aa and 601-aa long, respectively. We identified that the 601-aa-long isoform 2 was produced using a human-specific splice donor site (Figure 3C). The NUB1 protein interacts with ubiquitin-like proteins such as NEDD8 and FAT10 and accelerates their degradation [36-38]. Downregulation of NEDD8 by NUB1 leads to decreased p53 modification, resulting in cytoplasmic localization and inhibition of p53 transcriptional activity [39]. The generation of the novel splice donor site within exon 12 resulted in shortening of the exon and removal of 14 amino acids from the ancestral form. Human-specific isoform 2 has been detected almost equally in all examined tissues and shows a higher expression level than that

of ancestral isoform 1, which is not equally detected in tissues examined [35]. The 14-aa deletion removes one of the ubiquitin associated domains, probably resulting in weakening of the interaction between NUB1 and ubiquitin-like proteins. Actually, isoform 2 is less efficient in promoting degradation of NEDD8 [40]. Therefore, production of the human-specific NUB1 isoform 2 may affect the ubiquitin pathway in human cells.

Conclusions

We identified 85 cases of annotated human protein-coding exons with evidence of substantial expression that had been generated by acquisition of a novel splice site in the human genome. The novel protein-coding exons were expressed either constitutively or alternatively. We found three cases in which the human-specific isoform conferred novel molecular function. We propose that the generation of novel splice sites contributes to the evolution of the human genome and the variety of protein isoforms that may furnish novel functional proteins and novel phenotypes.

Methods

Collection of human-specific splice site candidates

The overall procedure we employed in this study is presented in Figure 4. We downloaded the multiple alignment data of the human protein-coding exon sequences and their mammalian orthologous genome sequences from the University of California Santa Cruz (UCSC) Genome

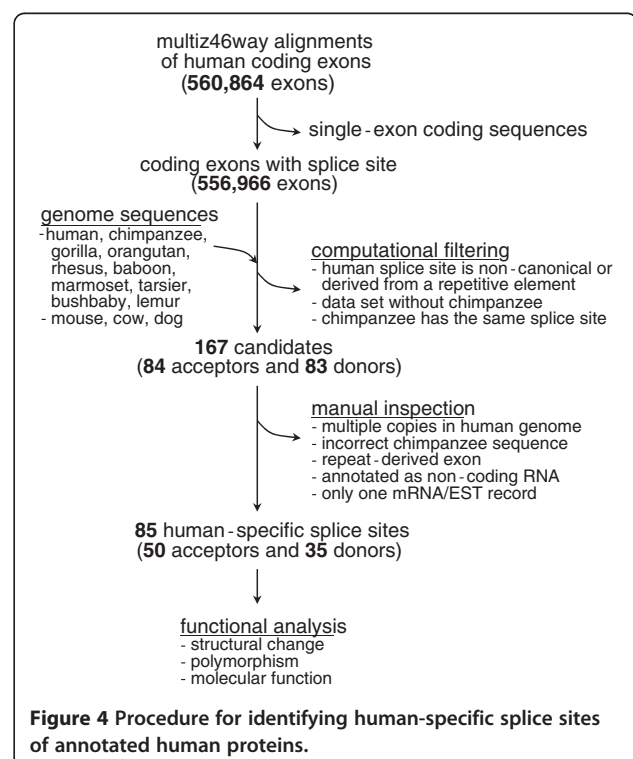


Figure 4 Procedure for identifying human-specific splice sites of annotated human proteins.

Brower Database (<http://genome.ucsc.edu>). The human coding regions were defined in the “knownGene” track, and the multiple sequence alignments of the coding regions were derived from the “multiz46way” alignment data. We extracted the protein-coding exon sequences, padded by 2 bp at both ends, from the respective genome assemblies. The genome assemblies analyzed in the present study included 10 primates and three other representative non-primate mammals: human (hg19), chimpanzee (panTro2), gorilla (gorGor1), orangutan (ponAbe2), rhesus macaque (rheMac2), baboon (papHam1), marmoset (calJac1), tarsier (tarSyr1), bushbaby, also known as galago (otoGar1), mouse lemur (micMur1), mouse (mm9), dog (canFam2), and cow (bosTau4).

We discarded splice sites derived from repetitive elements which were marked in lowercase in the repeat-masked genome sequences. When a human protein-coding exon has a canonical splice sequence (GT as the donor and AG as the acceptor), but the chimpanzee and other primates have a different sequence (other than GT or AG), we considered it to be a case of human-specific acquisition of a novel splice site after the human-chimpanzee divergence. We collected 167 human-specific splice site candidates (84 acceptors and 83 donors) by filtering the multiple alignment data.

Manual inspection of the candidates

We scrutinized each case by visual inspection mainly based on data available in the UCSC Genome Browser. We discarded cases showing one of the following conditions: the novel human exon was derived from a repeat sequence (the exon overlaps a repetitive element in the “RepeatMasker” track); the orthologous chimpanzee genomic segment had a sequencing error (the “Quality Score” track of the chimpanzee genome was reviewed); the orthology relationship of the gene was ambiguous due to the presence of highly similar paralogous copies in the human genome (BLAT was used to map the sequence to the human genome); the corresponding human transcript was annotated as a non-coding RNA (RefSeq record was reviewed); or the human transcript was not supported by two or more transcript records in the database (human mRNA and EST tracks were analyzed). As a result, we identified 50 splice acceptors (AG) and 35 splice donors (GT) that were specific to humans.

We compared the structures of the derived human exons and corresponding ancestral mammalian genome sequences to classify the protein modification types. We examined human mRNA and EST tracks of the UCSC Genome Browser to determine whether a novel splice site is used alternatively or constitutively. For the human simple nucleotide polymorphism (SNP), we analyzed the “Common SNPs(135)” or “Common SNPs(132)” track of the UCSC Genome Browser. We referred to UniProt

database (<http://www.uniprot.org>) for characterized molecular function of the human-specific protein isoforms.

Additional files

Additional file 1: List of the human-specific splice acceptors reported in this study.

Additional file 2: List of the human-specific splice donors reported in this study.

Additional file 3: Examples of the human-specific splice sites.

Additional file 4: Alignments of the exon and splice acceptor sequences.

Additional file 5: Alignments of the exon and splice donor sequences.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

YH conceived of the study, conducted the programming work, and prepared the manuscript. DSK participated in the sequence analysis and the literature survey. Both authors read and approved the final manuscript.

Acknowledgements

This work was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science, and Technology (2012R1A1B3001513) and by the Next-Generation BioGreen 21 Program (SSAC2011-PJ008220), Rural Development Administration, Republic of Korea.

Received: 12 May 2012 Accepted: 9 November 2012

Published: 13 November 2012

References

1. Varki A, Altheide TK: Comparing the human and chimpanzee genomes: searching for needles in a haystack. *Genome Res* 2005, **15**(12):1746–1758.
2. Enard W, Przeworski M, Fisher SE, Lai CS, Wiebe V, Kitano T, Monaco AP, Paabo S: Molecular evolution of FOXP2, a gene involved in speech and language. *Nature* 2002, **418**(6900):869–872.
3. Kim DS, Hahn Y: Identification of novel phosphorylation modification sites in human proteins that originated after the human-chimpanzee divergence. *Bioinformatics* 2011, **27**(18):2494–2501.
4. Knowles DG, McLysaght A: Recent de novo origin of human protein-coding genes. *Genome Res* 2009, **19**(10):1752–1759.
5. Kim DS, Hahn Y: Identification of human-specific transcript variants induced by DNA insertions in the human genome. *Bioinformatics* 2011, **27**(1):14–21.
6. Chou HH, Takematsu H, Diaz S, Iber J, Nickerson E, Wright KL, Muchmore EA, Nelson DL, Warren ST, Varki A: A mutation in human CMP-sialic acid hydroxylase occurred after the Homo-Pan divergence. *Proc Natl Acad Sci USA* 1998, **95**(20):11751–11756.
7. Hahn Y, Jeong S, Lee B: Inactivation of MOXD2 and S100A15A by exon deletion during human evolution. *Mol Biol Evol* 2007, **24**(10):2203–2212.
8. Lev-Maor G, Goren A, Sela N, Kim E, Keren H, Doron-Faigenboim A, Leibman-Barak S, Pupko T, Ast G: The “alternative” choice of constitutive exons throughout evolution. *PLoS Genet* 2007, **3**(11):e203.
9. Mudge JM, Frankish A, Fernandez-Banet J, Alioto T, Derrien T, Howald C, Reymond A, Guigo R, Hubbard T, Harrow J: The origins, evolution, and functional potential of alternative splicing in vertebrates. *Mol Biol Evol* 2011, **28**(10):2949–2959.
10. Corvelo A, Eyraes E: Exon creation and establishment in human genes. *Genome Biol* 2008, **9**(9):R141.
11. Lev-Maor G, Sorek R, Shomron N, Ast G: The birth of an alternatively spliced exon: 3' splice-site selection in Alu exons. *Science* 2003, **300**(5623):1288–1291.
12. Shen S, Lin L, Cai JJ, Jiang P, Kenkel EJ, Stroik MR, Sato S, Davidson BL, Xing Y: Widespread establishment and regulatory impact of Alu exons in human genes. *Proc Natl Acad Sci USA* 2011, **108**(7):2837–2842.

13. Sorek R, Lev-Maor G, Reznik M, Dagan T, Belinky F, Graur D, Ast G: **Minimal conditions for exonization of intronic sequences: 5' splice site formation in alu exons.** *Mol Cell* 2004, **14**(2):221–231.
14. Deininger PL, Batzer MA: **Alu repeats and human disease.** *Mol Genet Metab* 1999, **67**(3):183–193.
15. Sorek R, Ast G, Graur D: **Alu-containing exons are alternatively spliced.** *Genome Res* 2002, **12**(7):1060–1067.
16. Wang W, Zheng H, Yang S, Yu H, Li J, Jiang H, Su J, Yang L, Zhang J, McDermott J, et al: **Origin and evolution of new exons in rodents.** *Genome Res* 2005, **15**(9):1258–1264.
17. Zhang XH, Chasin LA: **Comparison of multiple vertebrate genomes reveals the birth and evolution of human exons.** *Proc Natl Acad Sci USA* 2006, **103**(36):13427–13432.
18. Pickrell JK, Pai AA, Gilad Y, Pritchard JK: **Noisy splicing drives mRNA isoform diversity in human cells.** *PLoS Genet* 2010, **6**(12):e1001236.
19. Schmitz J, Brosius J: **Exonization of transposed elements: a challenge and opportunity for evolution.** *Biochimie* 2011, **93**(11):1928–1934.
20. Amrani N, Sachs MS, Jacobson A: **Early nonsense: mRNA decay solves a translational problem.** *Nat Rev Mol Cell Biol* 2006, **7**(6):415–425.
21. Hiller M, Huse K, Szafranski K, Jahn N, Hampe J, Schreiber S, Backofen R, Platzer M: **Widespread occurrence of alternative splicing at NAGNAG acceptors contributes to proteome plasticity.** *Nat Genet* 2004, **36**(12):1255–1257.
22. Hiller M, Huse K, Szafranski K, Rosenstiel P, Schreiber S, Backofen R, Platzer M: **Phylogenetically widespread alternative splicing at unusual GYNGYN donors.** *Genome Biol* 2006, **7**(7):R65.
23. Tsai KW, Chan WC, Hsu CN, Lin WC: **Sequence features involved in the mechanism of 3' splice junction wobbling.** *BMC Mol Biol* 2010, **11**:34.
24. Hiller M, Huse K, Szafranski K, Jahn N, Hampe J, Schreiber S, Backofen R, Platzer M: **Single-nucleotide polymorphisms in NAGNAG acceptors are highly predictive for variations of alternative splicing.** *Am J Hum Genet* 2006, **78**(2):291–302.
25. Sinha R, Lenser T, Jahn N, Gausmann U, Friedel S, Szafranski K, Huse K, Rosenstiel P, Hampe J, Schuster S, et al: **TassDB2 - A comprehensive database of subtle alternative splicing events.** *BMC Bioinformatics* 2010, **11**:216.
26. Thanaraj TA, Clark F: **Human GC-AG alternative intron isoforms with weak donor sites show enhanced consensus at acceptor exon positions.** *Nucleic Acids Res* 2001, **29**(12):2581–2593.
27. Kim JK, Ryll R, Ishizuka Y, Kato S: **Identification of cDNAs encoding two novel nuclear proteins, IMUP-1 and IMUP-2, upregulated in SV40-immortalized human fibroblasts.** *Gene* 2000, **257**(2):327–334.
28. Kim SJ, An HJ, Kim HJ, Jungs HM, Lee S, Ko JJ, Kim IH, Sakuragi N, Kim JK: **Imup-1 and imup-2 overexpression in endometrial carcinoma in Korean and Japanese populations.** *Anticancer Res* 2008, **28**(2A):865–871.
29. Jeon SY, Lee HJ, Park JM, Jung HM, Yoo JK, Lee JS, Cha DH, Kim JK, Kim GJ: **Increased immortalization-upregulated protein 2 (IMUP-2) by hypoxia induces apoptosis of the trophoblast and pre-eclampsia.** *J Cell Biochem* 2010, **110**(2):522–530.
30. Goodwin AC, Destefano Shields CE, Wu S, Huso DL, Wu X, Murray-Stewart TR, Hacker-Prietz A, Rabizadeh S, Woster PM, Sears CL, et al: **Polyamine catabolism contributes to enterotoxigenic *Bacteroides fragilis*-induced colon tumorigenesis.** *Proc Natl Acad Sci USA* 2011, **108**(37):15354–15359.
31. Chaturvedi R, Asim M, Romero-Gallo J, Barry DP, Hoge S, de Sablet T, Delgado AG, Wroblewski LE, Piazuelo MB, Yan F, et al: **Spermine oxidase mediates the gastric cancer risk associated with *Helicobacter pylori* CagA.** *Gastroenterology* 2011, **14**(5):1696–1708. e1691-1692.
32. Casero RA, Pegg AE: **Polyamine catabolism and disease.** *Biochem J* 2009, **421**(3):323–338.
33. Murray-Stewart T, Wang Y, Devereux W, Casero RA Jr: **Cloning and characterization of multiple human polyamine oxidase splice variants that code for isoenzymes with different biochemical characteristics.** *Biochem J* 2002, **368**(Pt 3):673–677.
34. Murray-Stewart T, Wang Y, Goodwin A, Hacker A, Meeker A, Casero RA Jr: **Nuclear localization of human spermine oxidase isoforms - possible implications in drug response and disease etiology.** *FEBS J* 2008, **275**(11):2795–2806.
35. Tanaka T, Kawashima H, Yeh ET, Kamitani T: **Regulation of the NEDD8 conjugation system by a splicing variant, NUB1L.** *J Biol Chem* 2003, **278**(35):32905–32913.
36. Kito K, Yeh ET, Kamitani T: **NUB1, a NEDD8-interacting protein, is induced by interferon and down-regulates the NEDD8 expression.** *J Biol Chem* 2001, **276**(23):20603–20609.
37. Kamitani T, Kito K, Fukuda-Kamitani T, Yeh ET: **Targeting of NEDD8 and its conjugates for proteasomal degradation by NUB1.** *J Biol Chem* 2001, **276**(49):46655–46660.
38. Hipp MS, Raasi S, Groettrup M, Schmidtke G: **NEDD8 ultimate buster-1L interacts with the ubiquitin-like protein FAT10 and accelerates its degradation.** *J Biol Chem* 2004, **279**(16):16503–16510.
39. Liu G, Xirodimas DP: **NUB1 promotes cytoplasmic localization of p53 through cooperation of the NEDD8 and ubiquitin pathways.** *Oncogene* 2010, **29**(15):2252–2261.
40. Schmidtke G, Kalveram B, Weber E, Bochtler P, Lukasiak S, Hipp MS, Groettrup M: **The UBA domains of NUB1L are required for binding but not for accelerated degradation of the ubiquitin-like modifier FAT10.** *J Biol Chem* 2006, **281**(29):20045–20054.

doi:10.1186/1471-2105-13-299

Cite this article as: Kim and Hahn: Human-specific protein isoforms produced by novel splice sites in the human genome after the human-chimpanzee divergence. *BMC Bioinformatics* 2012 **13**:299.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

