RESEARCH ARTICLE

A Human-Specific Mutation Leads to the Origin of a Novel Splice Form of Neuropsin *(KLK8)*, a Gene Involved in Learning and Memory

Zhi-xiang Lu,[1–3] Jia Peng,[1,2] and Bing Su[1,2]*

[1]*Key Laboratory of Cellular and Molecular Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China;* [2]*Kunming Primate Research Center, Chinese Academy of Sciences, Kunming, China;* [3]*Graduate School, Chinese Academy of Sciences, Beijing, China*

Communicated by David Cooper

Neuropsin (kallikrein 8, KLK8) is a secreted-type serine protease preferentially expressed in the central nervous system and involved in learning and memory. Its splicing pattern is different in human and mouse, with the longer form (type II) only expressed in human. Sequence analysis suggested a recent origin of type II during primate evolution. Here we demonstrate that the type II form is absent in nonhuman primates, and is thus a human-specific splice form. With the use of an in vitro splicing assay, we show that a human-specific T to A mutation (c.71–127T>A) triggers the change of splicing pattern, leading to the origin of a novel splice form in the human brain. Using mutation assay, we prove that this mutation is not only necessary but also sufficient for type II expression. Our results demonstrate a molecular mechanism for the creation of novel proteins through alternative splicing in the central nervous system during human evolution. Hum Mutat 0,1–7, 2007.

KEY WORDS: neuropsin; kallikrein 8; KLK8; alternative splicing; cognition; human evolution

INTRODUCTION

Despite only ~1.2% difference in genomic DNA sequence [Chimpanzee Sequencing Consortium, 2005] human and chimpanzee (*Pan troglodytes*) differ considerably in mental and linguistic capabilities. The genetic changes leading to functional divergence between human and chimpanzee may occur at different levels, including gross alterations in cytogenetic architecture, local chromosomal rearrangements, segmental genomic duplications, single gene creation or loss, and differences in gene transcription and alternative splicing of mRNA [Gagneux and Varki, 2001]. Alternative splicing is one potentially important mechanism for creation of new proteins during evolution [Blencowe, 2006; Xing and Lee, 2005]. This strategy is common in the central nervous system (CNS), where neurons are found to be rich in regulated alternative splicing events [Grabowski, 1998]. Identifying human-specific alternative splice forms in CNS is thus important for understanding the mechanisms of functional evolution of human cognition.

Serine proteases perform a variety of functions under physiological and pathological conditions in the CNS [Yousef et al., 2003]. Neuropsin (also called kallikrein 8, gene symbol KLK8; MIM# 605644) is a member of the kallikrein serine protease family, and was initially cloned from mouse hippocampus [Chen et al., 1995]. Ontogeny studies on mouse suggested that neuropsin functioned in development, neurite outgrowth, fasciculation [Oka et al., 2002], and normality of neuronal synapse [Hirata et al., 2001]. Electrophysiology and behavior studies implied neuropsin's involvement in pathogenesis [Momota et al., 1998] and hippocampal plasticity [Akita et al., 1997; Tamura et al., 2006], associated with learning and memory.

Neuropsin is alternatively spliced in the human brain. Mitsui et al. [1999] reported two types of human neuropsin cDNAs in the brain, with one of them (type I) homologous to the mouse counterpart but the other (type II) absent in mouse. Both transcripts include six exons, and differ only in their exon 3 sequences. Type II includes extra 45 amino acids at the N-terminus of exon 3 due to alternative splicing (Fig. 1a). Expression analysis in humans showed that the expression of type II was developmental-stage related, and preferentially expressed in the adult human brain, suggesting the importance of the type II form to the normal function of human cognition [Mitsui et al., 1999]. Our previous sequence analysis in human and nonhuman primates indicated that only the hominoid species (human and apes) had the intact open reading frame of type II, while the other primate

The Supplementary Material referred to in this article can be accessed at http://www.interscience.wiley.com/jpages/1059-7794/suppmat.

Received 10 February 2007; accepted revised manuscript 26 March 2007.

*Correspondence to: Bing Su, Kunming Institute of Zoology, Chinese Academy of Sciences, 32 East Jiao-Chang Rd., Kunming 650223, Yunnan, China. E-mail: sub@mail.kiz.ac.cn

Grant sponsors: Natural Science Foundation of Yunnan Province of China; Chinese Academy of Sciences; Grant number: KSCX1-YW-R-34; Grant sponsor: National Natural Science Foundation of China; Grant numbers: 30370755, 30525028, and 30630013; Grant sponsor: National 973 Project of China; Grant number: 2006CB701506.

DOI 10.1002/humu.20547

Published online in Wiley InterScience (www.interscience.wiley.com).

species (Old World monkeys and New World monkeys) can not encode type II due to frameshift [Li et al., 2004]. We detected abundant expression of type II in the prefrontal cortex (PFC) of the adult human brain, but no expression was detected in the PFC of lesser apes and Old World monkeys (great apes were not tested), implying that type II originated recently in the primate lineage less than 18 million years ago [Li et al., 2004].

Here we show that the type II form of neuropsin is only expressed in the central nervous system of human, not in nonhuman primates, and its origin is less than 5 million years ago. The occurrence of the type II form in human is caused by a human-specific mutation near the novel splicing site. Our data demonstrate a molecular mechanism for the creation of novel proteins through alternative splicing in the CNS during human evolution.

MATERIALS AND METHODS

Nomenclature and Reference Sequences

The Human Genetic Variation Society (HGVS; www.hgvs.org/mutnomen)–approved guidelines were used in describing the mutations. In order to precisely describe mutant constructs and identify the variants among primates, the human cDNA RefSeq (GenBank accession no. NM_007196.2) was used for numbering the mutant sites in the reporter gene assays. The A of the ATG translation initiation codon was defined as +1. Three RefSeqs (GenBank accession nos. AY563056.1, human; AY563074.1, chimpanzee; and AY563089.1, rhesus macaque) were taken as the wild-type sequences. With the use of 5′RACE, we sequenced the rhesus macaque neuropsin, and the GenBank accession no. is EF407409.

Samples

For gene expression analysis, we collected placenta from human and postmortem brain tissues from human, white-browed gibbon (*Hylobates hoolock*), rhesus macaque (*Macaca mulatta*), and Yunnan golden monkey (*Rhinopithecus bieti*). They were from the collections of the Kunming Primate Research Center (KPRC), Chinese Academy of Sciences (CAS) and local hospitals. The human protocol was approved by the local hospitals. One chimpanzee and one orangutan PFC cDNA samples were kindly provided by Dr. Philipp Khaitovich (CAS-MPG Partner Institute for Computational Biology, Shanghai Institutes for Biological Sciences [SIBS], China). The RNA extraction tests indicated that none of the samples showed visible RNA degradation (data not shown).

For sequencing and in vitro splicing assays, the DNA samples of human, chimpanzees, and rhesus macaque were from the collections in the Kunming Cell Bank of CAS and the Kunming Blood Center. Informed consent was obtained from the human subjects.

RT-PCR

The total RNAs were extracted using TRIzol (Invitrogen, Carlsbad, CA). The RNA samples were treated with DNaseI (Takara, Tokyo, Japan) to remove possible genomic DNA contamination, then subject to reverse transcription using Ominicript Reverse Transcriptase (Qiagen, Valencia, CA). PCR was carried out at 95°C for 5 min, at 94°C for 20 s, and 59°C for 20 s, and 72°C for 30 s for 32 cycles. The primer sequences are included in Supplementary Table S1 (available online at http://www.interscience.wiley.com/jpages/1059-7794/suppmat).

Minigene Constructs

Two hybrid-minigene constructs were kindly provided by Dr. Thomas Cooper (Baylor College of Medicine, Houston, TX). The pSXN minigene reporter system was used to test whether a predicted exonic splicing enhancer (ESE) sequence was functional, and the RTB system was used for testing the effect of ESEs on splicing patterns. The splicing enhancer activities of the predicted sequences were tested in a heterologous system by measuring the rescue of splicing of the 34-nucleotide exon2 of the pSXN minigene reporter construct (4.11.12mus) [Coulter et al., 1997]. Synthesized oligonucleotides were annealed to form double strands, and then inserted into the *Sal*I /*Bam*HI cassette of the pSXN exon 2. The primer sequences are shown in Supplementary Table S1.

To detect the effect of mutation-associated exon skipping, we constructed a series of RTB hybrid-minigene constructs. To clone the three wild-type species exons, they were first PCR-amplified from the genomic DNAs. The PCR primers were designed to amplify exon 3 of type II plus the 5′ and 3′ flanking region (130 bp each side) to cover enough splicing information [Wang and Marin, 2006]. PCR was carried out at 95°C for 4 min, and then at 94°C for 20 s, 59°C for 20 s, and 72°C for 30 s for 35 cycles. The PCR-products contained primer-introduced *Sal*I and *Spe*I sites at the 5′ and 3′ ends, respectively, and they were cloned into the pMD19simple-T-vector (Takara), and then subcloned into the *Sal*I and *Spe*I sites located between the internal exon 2 and the last hybrid exon 4 of the RTB minigene. Other mutation-associated constructs were done by PCR-mediated site-directed mutagenesis. The integrity of the final constructs was confirmed by sequencing (ABI-3130 automatic sequencer, Boston, MA). Refer to Supplementary Table S1 for primer sequences.

In Vitro Functional Splicing Assay

HeLa and SK-N-SH cells were from the Kunming Cell Bank of CAS. Both cell lines were grown in DMEM (Gibco, Rockville, MD) medium supplemented with 10% fetal bovine serum (Hyclone, Logan, UT). Approximately 60 to 80% of confluent cells were transfected by Lipofectamine 2000 (Invitrogen), following the manufacturer's protocol (in 6-well plates, 2 μg vector was mixed with 5 μl lipofectamine). RNA was extracted with TRIzol (Invitrogen), 48 hr after transfection. Transfection experiments were performed in triplicate. For cDNA synthesis, 1 μg of DNase-treated (Fermentas, Hanover, MD) RNA was reverse transcribed using Oligo_d$(T)_{22}$ primer and Omniscript Reverse Transcriptase (Qiagen).

PCR was carried out at 95°C for 4 min, and then at 94°C for 25 s and 61°C for 25 s and 72°C for 30 s for 27 cycles. The PCR products were separated on 1.5% agarose gels containing ethidium bromide (0.5 × TBE) and the band intensities of the PCR products were measured using the Glyko BandScan 4.30 software package (Novato, CA). The primer sequences are included in Supplementary Table S1.

5′RACE of Rhesus Macaque Neuropsin

5′RACE was carried out using a SMART^{TM} RACE cDNA Amplification Kit (Clontech, Palo Alto, CA). The double-stranded cDNA was synthesized from rhesus macaque PFC polyA$^+$ RNA and ligated to a linker adaptor according to the manufacturer's manual. PCR was performed consecutively using the primer combination of NP_R and Primer_A _Mix, then Nest_R and Nest_Primer_A. Primer_A _Mix and Nest_Primer_A were specific for a linker adaptor and included in the kit. About a 450-bp fragment was cloned into pMD19simple-T-vector (Takara) and sequenced. The primer sequences are shown in Supplementary Table S1.

Promoter Analysis Assay

The PCR primers were designed to amplify the fragment covering approximately 1,500 bp upstream and ~600 bp down-

stream of the transcription start site without including the start codon [Heissig et al., 2005]. PCR was carried out at 95°C for 4 min, at 94°C for 30 s and 61°C for 30 s and 72°C for 80 s for 35 cycles. Refer to Supplementary Table S1 for primer sequences.

PCR products were digested by *Mlu*I/*Xho*I, then cloned into pGL3-basic *firefly* luciferase reporter vector (Promega, Madison, WI). In a 24-well cell culture plate, each plasmid (500 ng) was mixed with pRL-TK *renilla* luciferase reporter vector (200 ng) under the control of a low-level constitutive expression, and transfected into HeLa and 293T cells (from Kunming Cell Bank of CAS). Experiments were performed in quintuplicate and the activities of the two luciferases were measured according to the Dual-Luciferase Reporter Assay System Technical Manual (Promega). For transfection efficiency, the measurement of the *firefly* luciferase was normalized to the measurement of the *renilla* luciferase. The 293T cell line experiment was conducted using the same protocol except that transfections were done by the calcium phosphate cell transfection method.

RESULTS

Human-Specific Splice Form of Neuropsin

In this study, we tested the expression of type II in two great ape species, chimpanzee and orangutan (*Pongo pygmaeus*). Although great apes do contain an intact open reading frame [Li et al., 2004], RT-PCR did not detect type II expression in PFC of either species, nor in various lesser apes and Old World monkeys (Fig. 2). Thus, type II is a human-specific isoform which originated less than 5 million years ago [Goodman et al., 1998].

Identification of *cis*-Elements

Sequence comparison identified a human-specific point mutation (c.71–127T>A) in the type II-specific coding region 8 bp from the novel splicing site (Fig. 1b). This mutation is shared among humans (12 Africans, 10 Europeans, and 10 Chinese samples were tested here) and is thus presumably fixed. Sequence analysis using RESCUE-ESE [Fairbrother et al., 2002] predicts that there are exonic splicing enhancer (ESE) sequences (TGTGGA in human, chimpanzee and rhesus, and TGGAAG in human only) 2 to 4 bp from the novel splicing site (Fig. 1b). To test whether the predicted ESEs are functional splicing regulatory sequences, we conducted the in vitro splicing assay using the pSXN minigene reporter system. A 13-bp fragment covering the predicted ESEs was cloned into the pSXN vector and transfected into HeLa cells. The rescue of splicing of the 34-nucleotide pSXN exon 2 was measured by RT-PCR [Coulter et al., 1997] (Fig. 3a) and visualized by agarose gel electrophoresis. We constructed the vectors containing the 13-bp sequences from human, chimpanzee, rhesus macaque, and an artificial sequence with a G mutation at the human specific site (c.71–127A>G) (Fig. 3b). The result reveals that the constructs of human, chimpanzee, and the artificial sequence have similar proportions of exon-inclusion, which are significantly higher than the empty vector and the rhesus macaque construct ($P<0.01$, one-way analysis of variance [ANOVA] test) (Fig. 3c). This implies that this short sequence covering the novel splicing site is indeed a splicing regulatory sequence and the regulatory effects are different between rhesus macaque and human/chimpanzee. The high proportion of exon-inclusion of the artificial sequence is likely due to the regulatory effect of the purine-rich element (GGAGG) [Fairbrother et al., 2002; Tanaka et al., 1994]. A similar result was obtained when using another cell line (SK-N-SH) (data not shown).

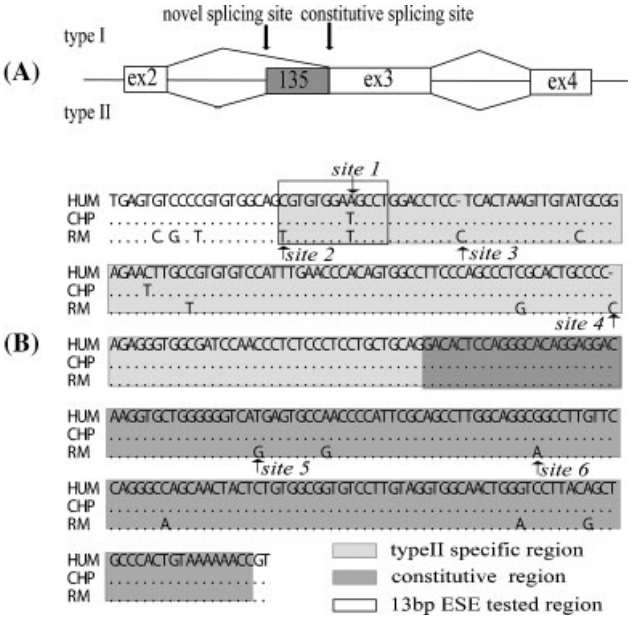

FIGURE 1. **A: The schematic map of the human neuropsin gene structure. B: The aligned neuropsin sequences of the exon 3 region in human and nonhuman primates. Sites 1–6 refer to the sites with sequence variations among primate species which were used in generating mutant constructs in the pSXN and RTB minigene assays (Figs. 3 and 4). Site 1 for RTB-H (c.71–127A>T), RTB-H (c.71–127A>G), and RTB-C (c.71–127T>A); Site2 for RTB-M (c.71–135T>C); Site 3 and Site 4 for RTB-M-del; Site 5 and Site 6 for RTB-M (c.111G>T) and RTB-M (c.144A>G), respectively. The 13-bp sequences used for the pSXN assay (Fig. 3) are highlighted by the square. HUM, human; CHP, chimpanzee; RM, rhesus macaque; ESE, exon splicing enhancer. The GenBank accession numbers of DNA sequences aligned here are AY563056.1, AY563074.1, and AY563089.1.**

Role of the Human-Specific T to A Mutation

Further analysis of splicing patterns in a variety of constructs using the RTB minigene system [Ryan and Cooper, 1996] suggests that the human-specific T-to-A mutation (c.71–127T>A) is both necessary and sufficient for expression of the novel splice variant (Fig. 4). As shown above, although both the human and chimpanzee sequences have ESE activity, the splicing pattern is subject to sequence context of the ESEs. We extensively evaluated the splicing efficiency of the whole exon 3 region and its flanking intron sequences (~130 bp each side) in different primate species (Fig. 4a). The results indicated that only the human construct expressed type II, but not chimpanzee and rhesus macaque (Fig. 4b). When the human-specific site was mutated from A to T (RTB-H (c.71–127A>T)), only the type I product was detected, therefore indicating that the "A" nucleotide was necessary to create the novel splicing site in human (Fig. 4c). We did another construct using the chimpanzee sequence by introducing a point mutation at the human-specific site (RTB-C (c.71–127T>A)). The result clearly showed that the introduced human point mutation in the chimpanzee sequence resulted in the creation of type II (Fig. 4c). Hence, the human-specific mutation is not only necessary but also sufficient in creating the novel splice form. We also tested the artificial construct with the G mutation at the human-specific site (RTB-H (c.71–127A>G)), no type II was detected though the aforementioned data revealed similar proportion of exon-inclusion in the pSXN reporter system. This

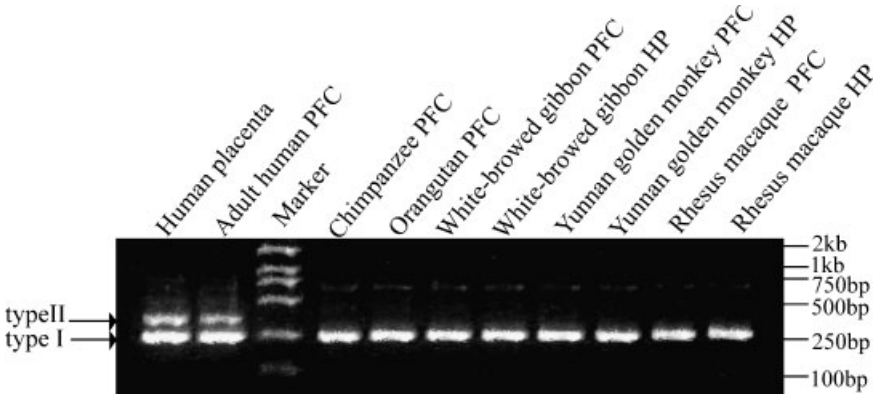

FIGURE 2. **The electromorph of the RT-PCR result. The expected products of type I and type II were confirmed by sequencing. HP, hippocampus; PFC, prefrontal cortex.**

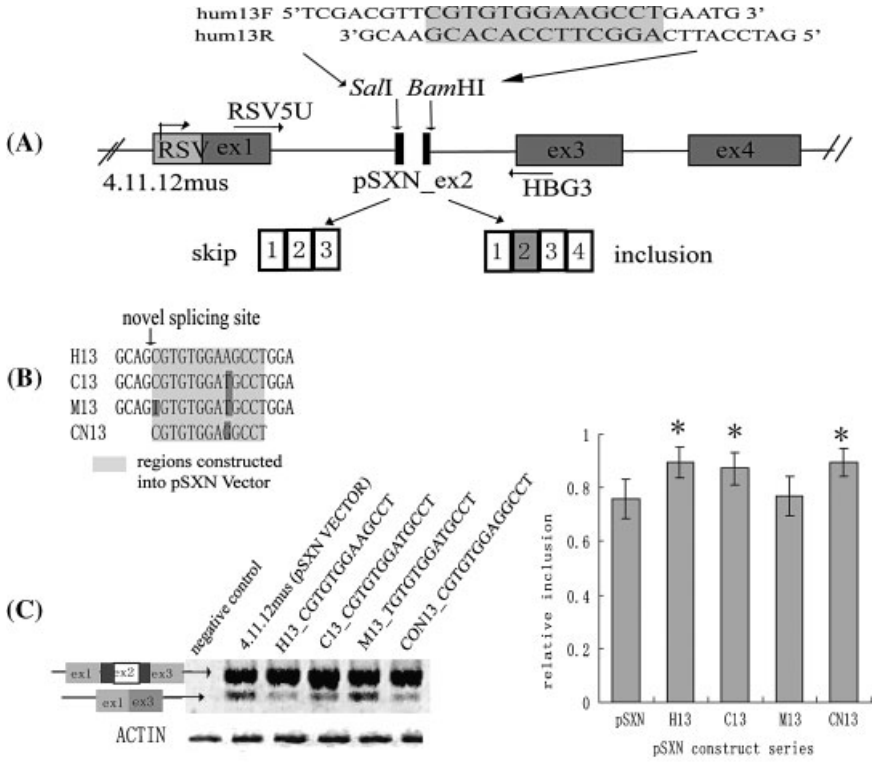

FIGURE 3. **The result of splicing rescue experiments. A: The schematic diagram for pSXN vector construction. B: The 13-bp insert sequences of human, chimpanzee, and rhesus macaque. C: The RT-PCR result of pSXN exon 2 inclusion transcripts from different constructs. The band intensities of the PCR products were quantified by the Glyko BandScan 4.30 software package. *Significant increase of exon2 inclusion as compared to the empty vector and the rhesus macaque construct ($P < 0.01$, one-way ANOVA test).**

result again indicated that ESEs acted in a context-dependent manner in regulating splicing. To eliminate the potential variation of cell types, we did the same test with SK-N-SH cell line, and the result was the same (data not shown).

Weakened Splicing of Constitutive Exon 3 in Great Apes

A weakening effect of the constitutive splicing site of exon 3 was also observed in the RTB minigene experiments, and a significant reduction of type I expression was detected in human (RTB-H) and chimpanzee (RTB-C) when compared with rhesus macaque (RTB-M) (Fig. 4b). This pattern suggests that before the emergence of the type II splice form in human, the weakening of the constitutive splicing site already existed in the common ancestor of human and chimpanzee or even earlier, implying a multistep process leading to the dramatic change of splicing pattern in human. In order to identify the sites responsible for the weakening effect, we constructed four mutants (RTB-M (c.71–135T>C); RTB-M-del; RTB-M (c.111G>T); and RTB-M (c.144A>G)) reflecting the sequence divergence within exon 3 between human/chimpanzee and rhesus macaque, and no visible weakening effects of constitutive splicing were observed (Fig. 4d and e). However, when we replaced the constitutive splicing site

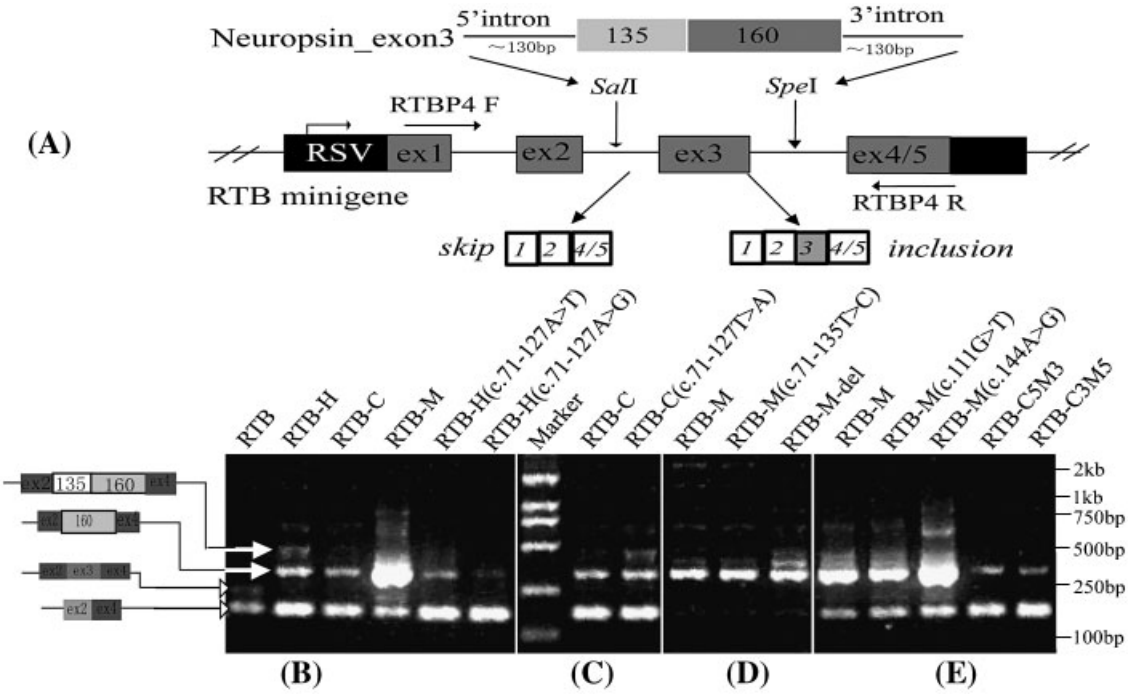

FIGURE 4. **The result of exon skipping tests using the RTB hybrid minigene system. A: The schematic diagram of the RTB minigene vector construction. B–E: The RT-PCR results of different constructs. RTB-H, human wild type; RTB-C, chimpanzee wild type; RTB-M, rhesus macaque wild type; RTB-H (c.71–127A>T), human A to T mutant at the human-specific site; RTB-H (c.71–127A>G), human A to G mutant at the human specific site; RTB-C (c.71–127T>A), chimpanzee T to A mutant at the human specific site; RTB-M (c.71–135T>C), rhesus macaque T to C mutant at the site 1 bp downstream of the novel splicing site (Fig. 1); RTB-M-del, this mutant deletes two insertions (c.[71–37_–38insC;71–114_–115insC]) in rhesus macaque compared to human/chimpanzee, resulting in an intact open reading frame of type II in human and apes; RTB-M (c.111G>T) and RTB-M (c.144A>G), rhesus macaque two ESE mutants, which are the predicted ESE motif sites in rhesus macaque, but not in human and chimpanzee); RTB-C5M3 and RTB-C3M5, two rhesus macaque-chimpanzee hybrid constructs by replacing the rhesus macaque constitutive splicing site (Fig. 1) flanking sequences with the correspondent chimpanzee sequences (266 bp at 5′ end for construct RTB-C5R3, and 264 bp at 3′ for construct RTB-C3M5). The RT-PCR products were cloned into T-vector and confirmed by sequencing.**

(Fig. 1) flanking sequences of the rhesus macaque construct with the correspondent chimpanzee sequences, a significant reduction of constitutive splicing was detected (Fig. 4e). This result suggests that both the 5′ and 3′ regions of the chimpanzee sequence has a weakening effect on the constitutive splicing site, and probably human as well due to sequence divergence during primate evolution. It is likely that both the creation of novel splice form and the weakening of constitutive splicing contribute to the splicing pattern changes during primate evolution, suggesting a multistep process eventually leading to the origin of the type II form in human.

Changes in the Promoter Region

We also found evidence for further changes in transcription regulation during primate evolution. The 5′RACE and dual-luciferase reporter assays indicated different transcription start sites (TSS) between human/chimpanzee and rhesus macaque, and the sequence divergence of 5′UTR of neuropsin between human and chimpanzee caused transcription modification. The 5′RACE revealed that compared with human and chimpanzee [Mitsui et al., 1999], rhesus macaque's TSS in the brain was located 60 bp downstream of the human TSS site (Fig. 5a), where a rhesus macaque specific 8-bp insertion existed. In the promoter region, there are two insertions (119 bp and 10 bp, respectively, Fig. 5a) in chimpanzee, and both inserts are in the core promoter activity region [Cooper et al., 2006]. Using the luciferase reporter system, we tested the effect of the insertions on transcription, and no significant difference was observed in the HeLa cells (Fig. 5b). However, with the 293T cell line, a clear difference was detected, with the human construct showing significant transcription reduction as compared with chimpanzee (Fig. 5b). Hence, besides the splicing pattern changes during primate evolution, the mutations in the promoter region may also alter transcription regulation of neuropsin, which is likely tissue-associated as reflected by the different effect in HeLa and 293T cell lines. Again, this indicates a multistep and complex change of neuropsin expression during primate evolution (Fig. 6).

DISCUSSION

It is common knowledge that alternative splicing of exons requires splicing control elements (SCE). A well-known exonic SCE is a GAA-containing motif. It is usually a functional splicing enhancer located near the splicing site [Louie et al., 2003; Willie and Majewski, 2004]. Our data showed that due to the human-specific T to A (c.71–127T>A) mutation, human gained a novel GAA containing motif and that this splicing enhancer created a novel splicing site 8 bp upstream of the mutation site, indicating that during evolution, the newly occurred enhancers may lead to creation of novel splice forms in a species-specific manner.

Recent reports showed that gene expression can be regulated by the differential use of alternative UTRs [Hughes, 2006; Pesole

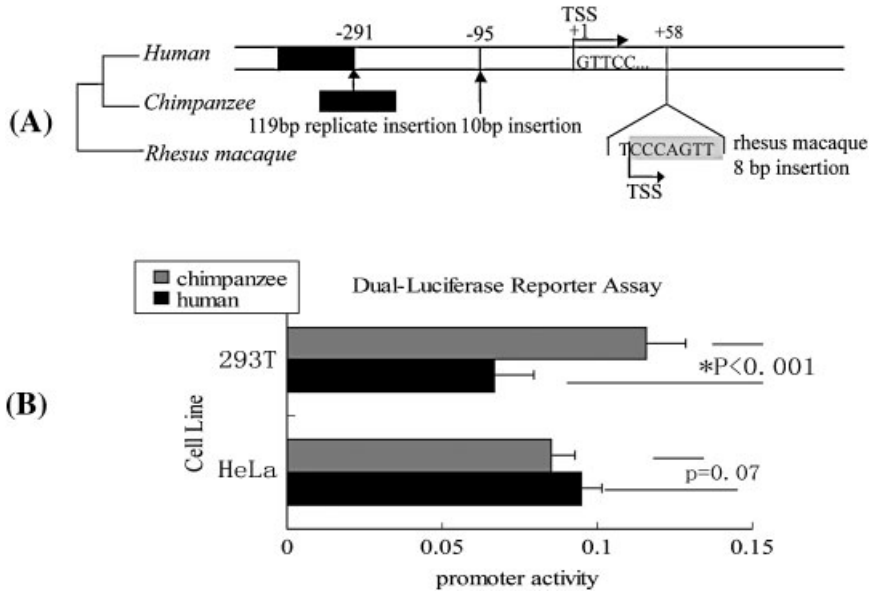

FIGURE 5. **A: The promoter structure comparison among human, chimpanzee, and rhesus macaque. B: The result of dual-luciferase reporter assay. The promoter activity was measured by testing the transformed *firefly* luciferase/*Renilla* luciferase ratio. No difference was observed with the HeLa cell line, but a significant difference was detected between human and chimpanzee with the 293T cell line ($P < 0.001$, Student's *t*-test).**

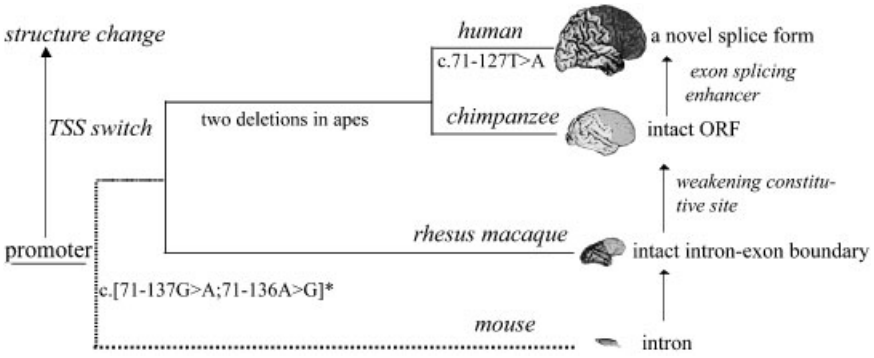

FIGURE 6. **The proposed pathways in creating the novel splicing site in human and expression regulation change of neuropsin during primate evolution. The key mutation events are labeled on the correspondent evolutionary lineages. When compared with mouse, a potential intron–exon boundary occurred due to the GA to AG mutation in primates that created a splicing acceptor site in intron 2–3 [Mitsui et al., 1999]. Two deletions occurred in the ape common ancestor lineage (Site 3 and 4 in Fig. 1), resulting in an intact opening reading frame for type II [Li et al., 2004].**

et al., 2001], which would also affect alternative splicing through promoter identity and occupation [de la Mata et al., 2003; Kornblihtt, 2005; Monsalve et al., 2000]. The dynamic changes in 5′UTR of neuropsin transcripts during primate evolution implies functional divergence of neuropsin in CNS although the coding region is relatively conserved [Li et al., 2004].

The fixation of the T to A mutation in human populations and the relatively abundant expression of type II in both human brain and placenta rule out the possibility that type II may just be a nonfunctional background splice form. Further studies should probe the biological function of type II in human. The extra 45 amino acids of type II may cause protein structure and function modification according to the reported crystal structure of type I neuropsin [Kishi et al., 1999]. The spatial-temporal dependent expression of type II in the human brain [Mitsui et al., 1999] also suggests its role in neurogenesis and cognition.

To understand the genetic basis that underlies the phenotypic traits that set human apart from nonhuman primates, recent studies have been focusing on identifying genes undergone positive Darwinian selection during human evolution [Varki, 2004]. For example, the brain-size–related genes ASPM and MCPH1 have rapid amino acid sequence changes in the human lineage [Evans et al., 2004; Kouprina et al., 2004; Wang and Su, 2004; Woods et al., 2005; Zhang, 2003]. The present results underscore the potential importance of the creation of novel splicing forms in the CNS in the emergence of human cognition.

ACKNOWLEDGMENTS

We thank Dr. Scott W. Roy for critical reading of the manuscript. We are grateful to the technical help of Hui Zhang and Qin Huang (Institute of Biochemistry and Cell Biology, SIBS,

China). We thank Dr. Thomas Cooper (Baylor College of Medicine, Houston, TX), who kindly provided the minigene constructs. We also thank Dr. Philipp Khaitovich (CAS-MPG Partner Institute for Computational Biology, SIBS, China), who provided the great ape cDNA samples.

REFERENCES

Akita H, Matsuyama T, Iso H, Sugita M, Yoshida S. 1997. Effects of oxidative stress on the expression of limbic-specific protease neuropsin and avoidance learning in mice. Brain Res 769:86–96.

Blencowe BJ. 2006. Alternative splicing: new insights from global analyses. Cell 126:37–47.

Chen ZL, Yoshida S, Kato K, Momota Y, Suzuki J, Tanaka T, Ito J, Nishino H, Aimoto S, Kiyama H, Shiosaka S. 1995. Expression and activity-dependent changes of a novel limbic-serine protease gene in the hippocampus. J Neurosci 15(Pt 2):5088–5097.

Chimpanzee Sequencing and Analysis Consortium. 2005. Initial sequence of the chimpanzee genome and comparison with the human genome. Nature 437:69–87.

Cooper SJ, Trinklein ND, Anton ED, Nguyen L, Myers RM. 2006. Comprehensive analysis of transcriptional promoter structure and function in 1% of the human genome. Genome Res 16:1–10.

Coulter LR, Landree MA, Cooper TA. 1997. Identification of a new class of exonic splicing enhancers by in vivo selection. Mol Cell Biol 17:2143–2150.

de la Mata M, Alonso CR, Kadener S, Fededa JP, Blaustein M, Pelisch F, Cramer P, Bentley D, Kornblihtt AR. 2003. A slow RNA polymerase II affects alternative splicing in vivo. Mol Cell 12:525–532.

Evans PD, Anderson JR, Vallender EJ, Gilbert SL, Malcom CM, Dorus S, Lahn BT. 2004. Adaptive evolution of ASPM, a major determinant of cerebral cortical size in humans. Hum Mol Genet 13:489–494.

Fairbrother WG, Yeh RF, Sharp PA, Burge CB. 2002. Predictive identification of exonic splicing enhancers in human genes. Science 297:1007–1013.

Gagneux P, Varki A. 2001. Genetic differences between humans and great apes. Mol Phylogenet Evol 18:2–13.

Goodman M, Porter CA, Czelusniak J, Page SL, Schneider H, Shoshani J, Gunnell G, Groves CP. 1998. Toward a phylogenetic classification of primates based on DNA evidence complemented by fossil evidence. Mol Phylogenet Evol 9:585–598.

Grabowski PJ. 1998. Splicing regulation in neurons: tinkering with cell-specific control. Cell 92:709–712.

Heissig F, Krause J, Bryk J, Khaitovich P, Enard W, Paabo S. 2005. Functional analysis of human and chimpanzee promoters. Genome Biol 6:R57.

Hirata A, Yoshida S, Inoue N, Matsumoto-Miyai K, Ninomiya A, Taniguchi M, Matsuyama T, Kato K, Iizasa H, Kataoka Y, Yoshida N, Shiosaka S. 2001. Abnormalities of synapses and neurons in the hippocampus of neuropsin-deficient mice. Mol Cell Neurosci 17:600–610.

Hughes TA. 2006. Regulation of gene expression by alternative untranslated regions. Trends Genet 22:119–122.

Kishi T, Kato M, Shimizu T, Kato K, Matsumoto K, Yoshida S, Shiosaka S, Hakoshima T. 1999. Crystal structure of neuropsin, a hippocampal protease involved in kindling epileptogenesis. J Biol Chem 274: 4220–4224.

Kornblihtt AR. 2005. Promoter usage and alternative splicing. Curr Opin Cell Biol 17:262–268.

Kouprina N, Pavlicek A, Mochida GH, Solomon G, Gersch W, Yoon YH, Collura R, Ruvolo M, Barrett JC, Woods CG, Walsh CA, Jurka J, Larionov V. 2004. Accelerated evolution of the *ASPM* gene controlling brain size begins prior to human brain expansion. PLoS Biol 2:E126.

Li Y, Qian YP, Yu XJ, Wang YQ, Dong DG, Sun W, Ma RM, Su B. 2004. Recent origin of a hominoid-specific splice form of neuropsin, a gene involved in learning and memory. Mol Biol Evol 21: 2111–2115.

Louie E, Ott J, Majewski J. 2003. Nucleotide frequency variation across human genes. Genome Res 13:2594–2601.

Mitsui S, Tsuruoka N, Yamashiro K, Nakazato H, Yamaguchi N. 1999. A novel form of human neuropsin, a brain-related serine protease, is generated by alternative splicing and is expressed preferentially in human adult brain. Eur J Biochem 260:627–634.

Momota Y, Yoshida S, Ito J, Shibata M, Kato K, Sakurai K, Matsumoto K, Shiosaka S. 1998. Blockade of neuropsin, a serine protease, ameliorates kindling epilepsy. Eur J Neurosci 10:760–764.

Monsalve M, Wu Z, Adelmant G, Puigserver P, Fan M, Spiegelman BM. 2000. Direct coupling of transcription and mRNA processing through the thermogenic coactivator PGC-1. Mol Cell 6:307–316.

Oka T, Akisada M, Okabe A, Sakurai K, Shiosaka S, Kato K. 2002. Extracellular serine protease neuropsin (KLK8) modulates neurite outgrowth and fasciculation of mouse hippocampal neurons in culture. Neurosci Lett 321:141–144.

Pesole G, Mignone F, Gissi C, Grillo G, Licciulli F, Liuni S. 2001. Structural and functional features of eukaryotic mRNA untranslated regions. Gene 276:73–81.

Ryan KJ, Cooper TA. 1996. Muscle-specific splicing enhancers regulate inclusion of the cardiac troponin T alternative exon in embryonic skeletal muscle. Mol Cell Biol 16:4014–4023.

Tamura H, Ishikawa Y, Hino N, Maeda M, Yoshida S, Kaku S, Shiosaka S. 2006. Neuropsin is essential for early processes of memory acquisition and Schaffer collateral long-term potentiation in adult mouse hippocampus in vivo. J Physiol 570(Pt 3):541–551.

Tanaka K, Watakabe A, Shimura Y. 1994. Polypurine sequences within a downstream exon function as a splicing enhancer. Mol Cell Biol 14: 1347–1354.

Varki A. 2004. How to make an ape brain. Nat Genet 36:1034–1036.

Wang YQ, Su B. 2004. Molecular evolution of microcephalin, a gene determining human brain size. Hum Mol Genet 13:1131–1137.

Wang M, Marin A. 2006. Characterization and prediction of alternative splice sites. Gene 366:219–227.

Willie E, Majewski J. 2004. Evidence for codon bias selection at the pre-mRNA level in eukaryotes. Trends Genet 20:534–538.

Woods CG, Bond J, Enard W. 2005. Autosomal recessive primary microcephaly (MCPH): a review of clinical, molecular, and evolutionary findings. Am J Hum Genet 76:717–728.

Xing Y, Lee C. 2005. Evidence of functional selection pressure for alternative splicing events that accelerate evolution of protein subsequences. Proc Natl Acad Sci USA 102:13526–13531.

Yousef GM, Kishi T, Diamandis EP. 2003. Role of kallikrein enzymes in the central nervous system. Clin Chim Acta 329:1–8.

Zhang J. 2003. Evolution of the human *ASPM* gene, a major determinant of brain size. Genetics 165:2063–2070.