Periodic Explosive Expansion of Human Retroelements Associated with the Evolution of the Hominoid Primate

Five retroelement families, L1 and L2 (long interspersed nuclear element, LINE), Alu and MIR (short interspersed nuclear element, SINE), and LTR (long terminal repeat), comprise almost half of the human genome. This genome-wide analysis on the time-scaled expansion of retroelements sheds light on the chronologically synchronous amplification peaks of each retroelement family in variable heights across human chromosomes. Especially, L1s and LTRs in the highest density on sex chromosomes Xq and Y, respectively, disclose peak activities that are obscured in autosomes. The periods of young L1, Alu, LTR, and old L1 peak activities calibrated based on sequence divergence coincide with the divergence of the three major hominoid divergence as well as early eutherian radiation while the amplification peaks of old MIR and L2 account for the marsupial-placental split. Overall, the peaks of autonomous LINE (young and old L1s and L2s) peaks and non-autonomous SINE (Alus and MIRs) have alternated repeatedly for 150 million years. In addition, a single burst of LTR parallels the Cretaceous-Tertiary (K-T) boundary, an exceptional global event. These findings suggest that the periodic explosive expansions of LINEs and SINEs and an exceptional burst of LTR comprise the genome dynamics underlying the macroevolution of the hominoid primate lineage.

Key Words : *Retroelements; Primate Evolution; Sex Chromosomes*

Tae-Min Kim, Seung-Jin Hong, Mun-Gan Rhyu

Department of Microbiology, College of Medicine, The Catholic University of Korea, Seoul, Korea

Address for correspondence
Mun-Gan Rhyu, M.D.
Department of Microbiology, College of Medicine
The Catholic University of Korea, 505 Banpo-dong,
Socho-gu, Seoul 137-701, Korea
Tel : +82.2-590-1220, Fax : +82.2-596-8969
E-mail : rhyumung@catholic.ac.kr

INTRODUCTION

Eukaryotic genomes contain variable amounts of retroelements, which propagate within host genome using RNA intermediates (1). Approximately half of the human genome is composed of five major retroelement families that are classified into three groups, termed as the long interspersed nuclear element (LINE; L1 and L2), short interspersed nuclear element (SINE; Alu and MIR), and long terminal repeat (LTR) (2). The extreme abundance of retroelements in the modern human genome reflects the possibility of the explosive outbreaks that resulted in the massive propagation and fixation of the highly repetitive retroelement copies. When neutral or at least tolerable retroelements do not come up with certain threshold copy numbers, most of them are eliminated from a breeding population through random drift (3). A massive number of new retroelement copies is fixed within a population shortly after amplification (4), and the ancient retroelement's activity can be traced based on the wave of neutral mutations increased with time (1). The genome-wide analysis of Alu sequence divergence demonstrates a sharp amplification peak occurring 40 million years ago (MYRA), which corresponds to the expansion of the primate lineage (2, 5). This finding strongly supports the view that the explosive expansion of retroelements increased the genomic diversity

responsible for the speciation events of host (6-8).

In spite of phylogenetic significances of individual retroelement families, the evidence of explosive expansions for retroelement other than Alus is considerably limited due to ambiguous peak activities. An approximate timescale for the entire retroelement formation roughly revealed the continued flourishing of autonomous L1s and LTRs for 150 million years (2). Since self-preserving retroelements cause harmful effects on the host genome by direct insertional mutagenesis or by altering genomic stability (9), the intragenomic propagation of retroelements has been selectively restrained depending on their effects on the host. Therefore, the pattern of retroelement amplification has been obscured in the modern genome, which makes it difficult to retrace when and how retroelement peak activities have appeared and influenced the evolution of the host genome.

Wide variations in the retroelement density across chromosomes have been observed previously (10, 11). Among all human chromosomes, chromosome 19, X, and Y contain Alu, L1, and LTR elements in the highest density, respectively. Because Alu elements have been supposed to promote protein expression and selectively accumulated nearby genes (12), it is reasonable to assume that the gene-rich chromosome 19 preferentially favored Alu insertion. The overwhelming frequency of LTRs in chromosome Y can be explained by

the suppression of recombination between the heteromorphic sex chromosomes, which would otherwise remove potentially harmful genetic elements selectively (13, 14). Overrepresented L1s on chromosome X have been suggested to function as waystations for the spreading of heterochromatization; thus, playing a role in the inactivation of chromosome X (15). Basically, the variable composition of retroelements in human chromosomes may be a testament to the exapted role of self-preserving retroelements for the symbiotic advantage of both the host genome and the parasitic retroelement (16). Given that the initial integration of retroelements occur randomly (17, 18), each chromosome is believed to have formulated a unique composition of retroelements through the process of post-insertion selection. Therefore, propagation of retroelements is expected to leave expansion curves similar in the time of amplification and variable in the height of peaks across chromosomes.

We assumed that the comparative analysis of retroelement substitutions on individual chromosomes, including sex chromosomes, would unveil the most prominent expansion curves, which most likely represent the period of peak activity associated with explosive amplification. In this study, LINEs, SINEs, and LTRs on each chromosome were partitioned according to sequence divergence to trace time-scaled retroelement activities. Overall, the most prominent expansion curves

revealed five reciprocal peaks of autonomous LINE and non-autonomous SINE amplification in a timescale of 150 million years. The time of an exceptional LTR burst was calibrated comparably with the Cretaceous-Tertiary (K-T) boundary between the Mesozoic and Cenozoic periods approximately 65 MYRA (19). In particular, the periodic peak activities of retroelements correspond temporally with speciation events in the hominoid primate lineage, suggesting the programming of genome dynamics by retroelements for the hominoid evolution.

MATERIALS AND METHODS

Data sources of human retroelements

The retroelement data for all figures and tables were retrieved from the December 2001 draft human genome assembly at http://genome.ucsc.edu. The genomic locations and annotations of retroelements were derived from the Repeat-Masker program output (http://ftp.genome.washington.edu/RM/RepeatMasker.html) based on RepBase (20). The sequence divergence of retroelements was directly taken from RepeatMasker results based on the algorithm of Waterman and Smith (21) in sensitive setting. For Table 1, the retroele-

Table 1. Chromosomal proportion of major retroelement subfamilies in the human genome (%)

	AluY	AluS	AluJ	*FAM	MIR	L1P	L1M	L2	‡ERV class I	ERV (K) class II	ERV (L) class III	MaLR
Chr1	1.6	7.3	2.9	0.5	2.9	6.1	9.3	4.4	2.9	0.3	1.5	3.5
Chr2	1.3	5.3	2.0	0.3	1.9	6.1	9.9	2.9	2.5	0.2	1.4	3.6
Chr3	1.3	5.3	2.0	0.3	2.3	7.3	9.8	3.2	3.0	0.3	1.5	3.8
Chr4	1.1	4.1	1.6	0.3	1.8	8.2	8.8	3.1	3.1	0.3	1.8	4.1
Chr5	1.1	4.6	1.6	0.3	1.9	6.5	8.4	2.8	2.3	0.2	1.4	3.6
Chr6	1.2	5.0	1.8	0.3	1.6	5.9	8.4	2.8	2.6	0.3	1.4	3.2
Chr7	1.6	7.2	2.8	0.4	1.7	6.2	10.3	2.9	3.0	0.3	1.4	3.4
Chr8	1.2	5.3	2.1	0.3	2.2	6.7	9.4	3.2	2.9	0.4	1.5	3.8
Chr9	1.6	6.7	2.4	0.4	2.6	6.3	10.6	3.3	2.6	0.2	1.4	3.8
Chr10	1.5	6.8	2.7	0.4	2.2	5.6	10.4	2.8	2.7	0.3	1.4	3.4
Chr11	1.5	6.4	2.4	0.4	3.2	7.6	9.7	4.1	3.1	0.3	1.5	3.7
Chr12	1.6	7.3	3.0	0.5	2.7	6.7	9.2	4.1	3.0	0.3	1.5	4.0
Chr13	1.3	4.8	1.8	0.3	1.6	6.6	10.4	3.2	3.3	0.2	1.7	4.3
Chr14	1.5	6.5	2.6	0.4	2.2	6.6	9.9	3.2	3.0	0.3	1.6	4.1
Chr15	1.7	7.3	2.8	0.4	2.2	4.8	10.5	2.8	2.3	0.2	1.2	2.8
Chr16	2.0	10.3	4.3	0.6	3.1	3.5	8.1	3.1	2.4	0.3	1.2	3.6
Chr17	2.5	12.0	4.1	0.6	2.4	2.9	7.7	3.2	2.1	0.3	1.0	2.7
Chr18	1.3	5.2	2.1	0.3	1.7	5.7	10.6	2.9	2.9	0.1	1.5	3.7
Chr19	3.3	17.2	5.9	0.9	1.9	3.4	6.9	2.8	4.6	1.2	1.2	2.3
Chr20	1.5	8.0	3.1	0.5	3.2	3.8	9.4	4.3	2.2	0.2	1.7	4.2
Chr21	1.8	6.4	2.2	0.3	1.5	5.4	10.7	2.6	3.6	0.3	2.1	6.1
Chr22	2.7	11.3	3.9	0.6	3.2	2.2	8.2	3.5	2.1	0.3	1.0	2.8
†Xp	1.1	6.0	2.6	0.3	1.6	8.7	11.2	2.7	3.6	0.3	1.8	4.4
Xq	1.0	4.1	1.5	0.2	1.9	16.9	16.3	3.3	4.3	0.4	1.9	4.3
ChrY	1.4	4.2	1.8	0.3	0.5	16.0	8.8	0.9	12.0	2.5	1.3	2.8
Average	1.5	6.4	2.4	0.4	2.2	6.6	9.7	3.2	2.9	0.3	1.5	3.7

*FAM includes both FLAM/FRAM (free left/right Alu monomer); †Short arm (Xp) and long arm (Xq) of chromosome X were separately analyzed; ‡Heterogeneous LTR retroelements were categorized into three autonomous (ERV class I, II, and III) and one non-autonomous (MaLR) subfamilies.

ments were further categorized into major subfamilies according to the similarities with the consensus sequences and the chromosomal fraction of each subfamily was calculated by dividing the total retroelement size by gap-excluded chromosomal size. To analyze the time-coursed pattern of retroelement expansion using sequence divergence, the chromosomal fraction was plotted at each bin of divergence level in 1% scale. The expansion curves of retroelements on individual chromosomes were compared for the highest peak made among chromosomes. For chromosome X, its long arm (Xq) was separately analyzed from the short arm (Xp), because Xq and Xp are composed of different anciently conserved regions (XCR) and recently added regions from autosomes (XAR), respectivley (22). The cytogenetic bands of sex chromosomes were redrawn according to bands information of the same genomic version (Dec. 2001) with the analysis of retroelements.

Age estimation using sequence divergence

Since sequence divergence from the master copy consensus is proportional to the integration age, the divergence range of retroelements from 1% to 38% was converted into the timescale of 150 million years. The estimated ages, using the level of sequence divergence, were compared with the timescale drawn previously using the approximate molecular clock (2). Since a high mutation rate of CpG transition leads to an overestimated age for Alus containing high CpG contents (23), we used the Alu-specific age estimated from previous literature (2). The age of LTRs on chromosome Y was recalibrated using the relative divergence ratio of 1.57 to compensate for the rapid divergence rate of chromosome Y (2, 24).

RESULTS

Composition of retroelements in the human chromosomes

The five families of retroelements in the human genome were collectively retrieved from the public database. The composition of retroelement subfamilies in each chromosome is listed in Table 1. For chromosome X, we separately analyzed the long (Xq) and short (Xp) arms, each of which derived from different origins (22). Chromosome 19 contains Alus in the highest density, comprising 27.2% of chromosome. The major subfamilies, Alu J (3.26%), Alu S (17.2%), Alu Y (5.9%), and free Alu monomers (0.85%) are also most frequent on chromosome 19. The density of Alu distribution is lowest on chromosome 4 (7.1%) and Xq (6.9%). Two major L1 subfamilies, L1P (primate, 16.9%) and L1M (mammal, 16.3%), commonly occupy chromosome Xq in the highest density. On chromosome Y, class I endogenous retrovirus (ERV) elements are overrepresented at 12% - four times the genomic average (2.9%) - while class III ERVs (1.3%) and non-autonomous mammalian apparent LTR retroposons (MaLR, 2.8%) are slightly underrepresented compared to the autosomal averages of 1.5% and 3.7%, respectively. Two ancient retroelements, L2s and MIRs, have considerably low average densities, comprising 3% and 2.5% of each chromosome.

The comparison of L1, Alu and LTR contents in each chromosome is shown in Fig. 1. The gene-rich chromosome 19, 17 and 22 harbor non-autonomous Alus in the highest density, whereas the sex chromsomose Xq and Y reveal autonomous L1s and LTRs in the highest density. In autosomes, Alus are variably distributed according to gene density, and L1s and LTRs are rather uniformly distributed.
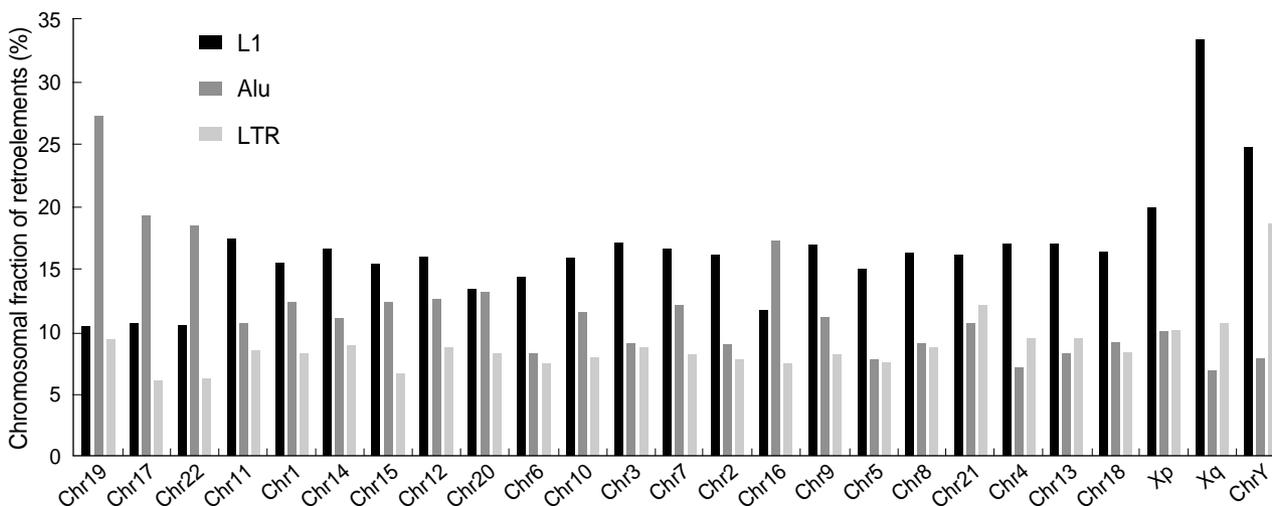
Fig. 1. Chromosomal distribution of three major retroelements, L1, Alu, and LTR, in the human genome. Individual chromosomes are aligned in the order of gene-rich chromosomes in the left and gene-poor and sex chromosomes in the right. Gene-rich and -poor chromosomes demonstrate the opposite retroelement compositions, Alu-rich and L1-poor versus Alu-poor and L1-rich, respectively. The long arm of chromosome X (Xq) and chromosome Y contain L1s and LTRs in the highest density, respectively.

Fig. 2. Time-scaled expansion curves of L1, Alu, and LTR retroele-ments. The chromosomal fractions of retroelement family are plot-ted against percent of sequence divergence. The expansion curves of L1s most prominent on chromosome Xq, Alus on chro-mosome 19, and LTRs on chromosome Y are compared with those of autosomal ave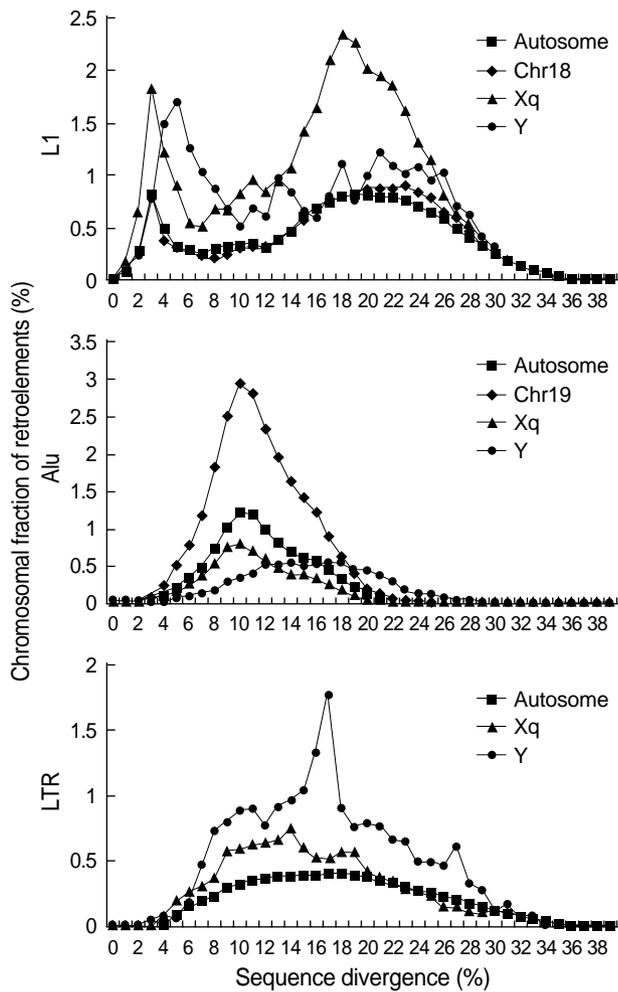rage and counterpart sex chromosomes. The sex chromosomes disclose the L1s and LTRs peak activi-ties, both of which are obscured in autosomes. The peak curves of L1s and Alus on chromosome Y tend to be skewed toward the higher divergence than those on other chromosomes.

Substitution-scaled distribution pattern of retroelements on individual chromosomes

The chromosomal fractions of five retroelement families were plotted according to the degree of sequence divergence for the analysis of time-scaled accumulation (Fig. 2, 3). L1s in the highest density on chromosome Xq make the most prominent expansion curve with two amplification peaks at the sequence divergences of 3% and 18% (Fig. 2). The domi-nant amplification peak at 18% is unique on chromosome Xq, and the height is three times higher than that of the average amplification peak on autosomes. The over-accumu-lation of L1s on Xq continues throughout the lifetime of
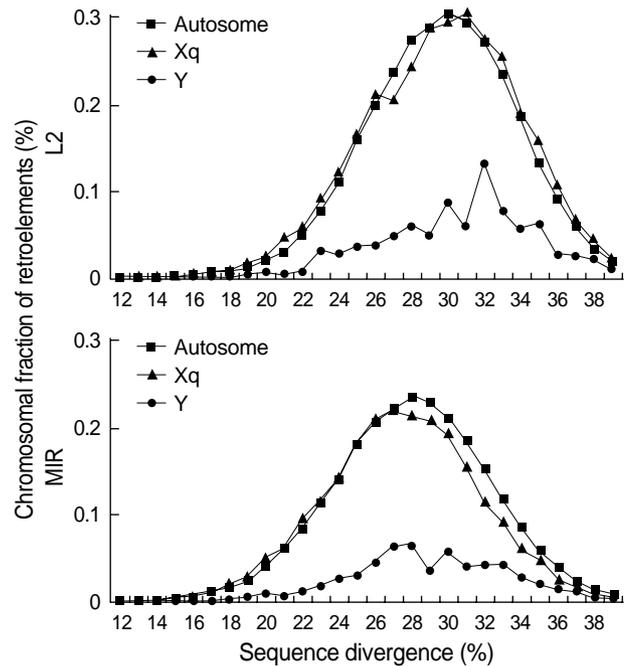

Fig. 3. Evolutionary waves of L2s and MIRs belonging to ancient LINE and SINE, respectively. These old retroelements encompass the range of high divergence from 20% to 36% and their peak fractions at 28-31% divergence occupy only less than 0.3% of chromosome. Chromosome Y contains an immaterial level of L2s and MIRs.

L1s, encompassing the divergence range from 1% to 30%. Chromosome Y contains an L1 amplification peak at 5% divergence, which is as prominent as that on chromosome Xq, but has no L1 peak at 18% divergence.

A single amplification peak of Alus at 10% sequence diver-gence is most prominent on gene-rich chromosome 19, and this over-dominance is sustained in the range of sequence divergence from 4% to 21% (Fig. 2). Other autosomes and chromosome Xq have the reduced peaks of Alu amplification in the divergence range similar with chromosome 19. Alus on chromosome Y display a widened expansion curve shifted toward the higher divergence.

LTRs in the highest density on chromosome Y make a single amplification peak at 17% divergence, which is four times the height of the autosomal average amplification (Fig. 2). The LTR peak abruptly appears in the middle of the wave showing a rapid rise and slow fall with increased divergence. The over-representation of LTRs on chromosome Y is main-tained from 7% to 27% divergence. The amplification curve of LTRs on chromosome X is more prominent than the auto-somal amplification at less than 20% divergence.

Autosomes and chromosome X show similar L2 and MIR amplification rate and chromosomal fraction (Table 1), and make the amplification peaks at the divergence rates of 31% and 28%, respectively (Fig. 3). The expansion curves of L2s and MIRs similarly initiate at 38% divergence rate and ter-minate at 20% divergence rate, before the thriving of the
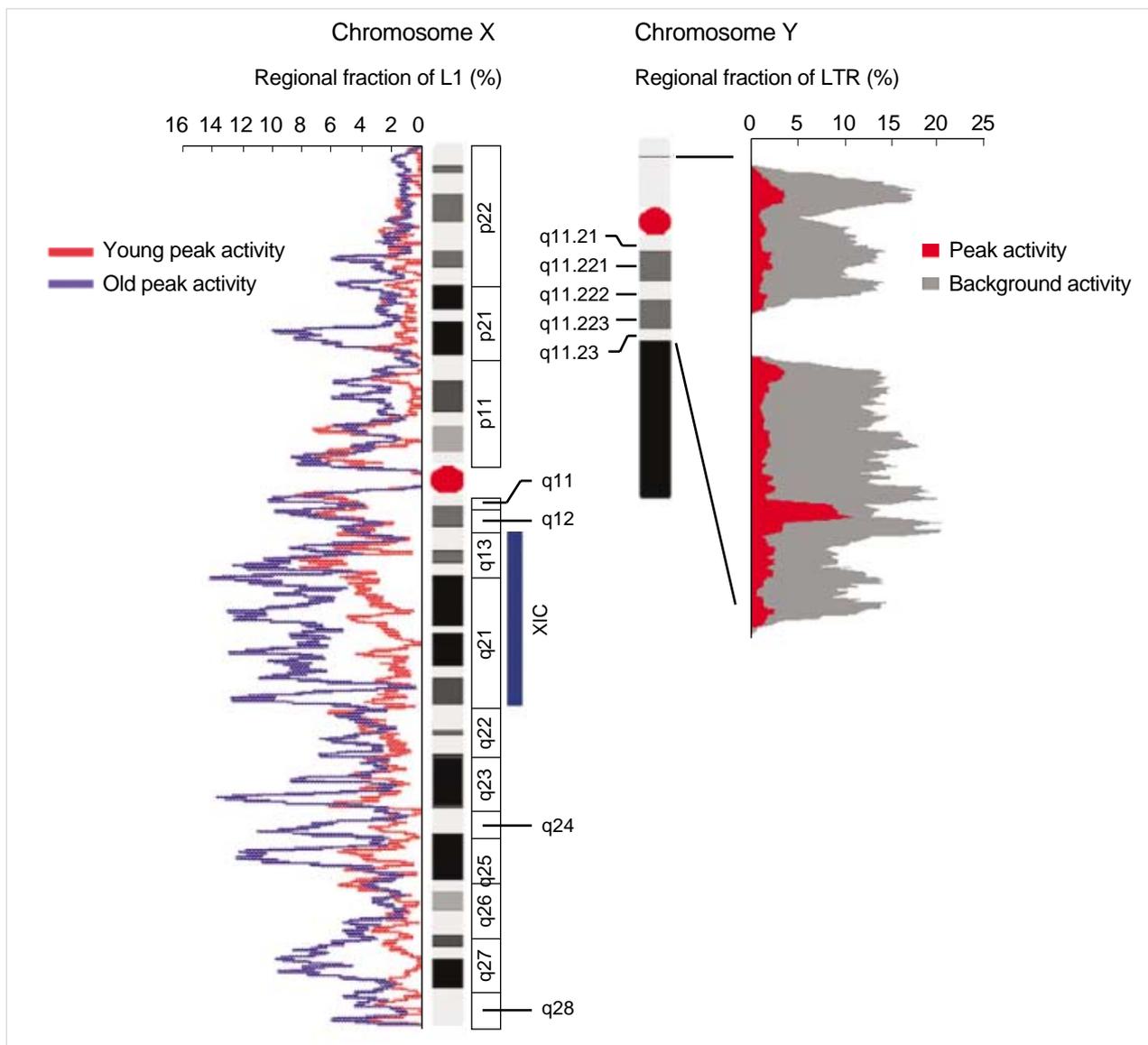
Fig. 4. Distribution of autonomous retroelements in the highest density on the sex chromosome X (L1s) and Y (LTRs). The L1 old peak activity at 3-5% divergence (red) and young peak activity at 17-20% divergence (blue), and the LTR peak activity at 15-17% divergence (red) against the background activity (gray) are separately plotted for the accumulation pattern of retroelements at the different periods. The proportions of retroelements are demonstrated in 1 mb-sized windows with a 100 kb step. Old peak L1s extensively cluster at Xq13-Xq21 that include inactivation center (XIC) and peak LTRs are exceptionally concentrated on Yq11.

young LINE (L1) and SINE (Alu) families. The peak heights of L2 and MIR amplifications, especially an immaterial peak on chromosome Y, are considerably low compared with those of L1 and Alu amplifications.

Overall, three peak activities of LINEs (young L1s, old L1s, and L2s) and two peak activities of SINEs (Alus and MIRs) make the seesaw-like transition of amplifications indicating the reciprocal relationship between autonomous LINEs and non-autonomous SINEs. Retroelements on autosomes and chromosome X tend to peak synchronously at similar divergence levels while varying in extent of fixation. The amplification peaks of retroelements on chromosome Y tend to be skewed toward higher sequence divergence as compared with those on other chromosomes.

Regional distribution pattern of retroelements on sex chromosomes

We analyzed the regional distribution of L1 and LTR on sex chromosomes X and Y, respectively (Fig. 4). The distribution of L1s comprising amplification peaks at 3-5% (young copies) and 15-17% (old copies) sequence divergence was separately mapped on chromosome X. We found prominent and extensive accumulation of old L1 copies in the cytoge-
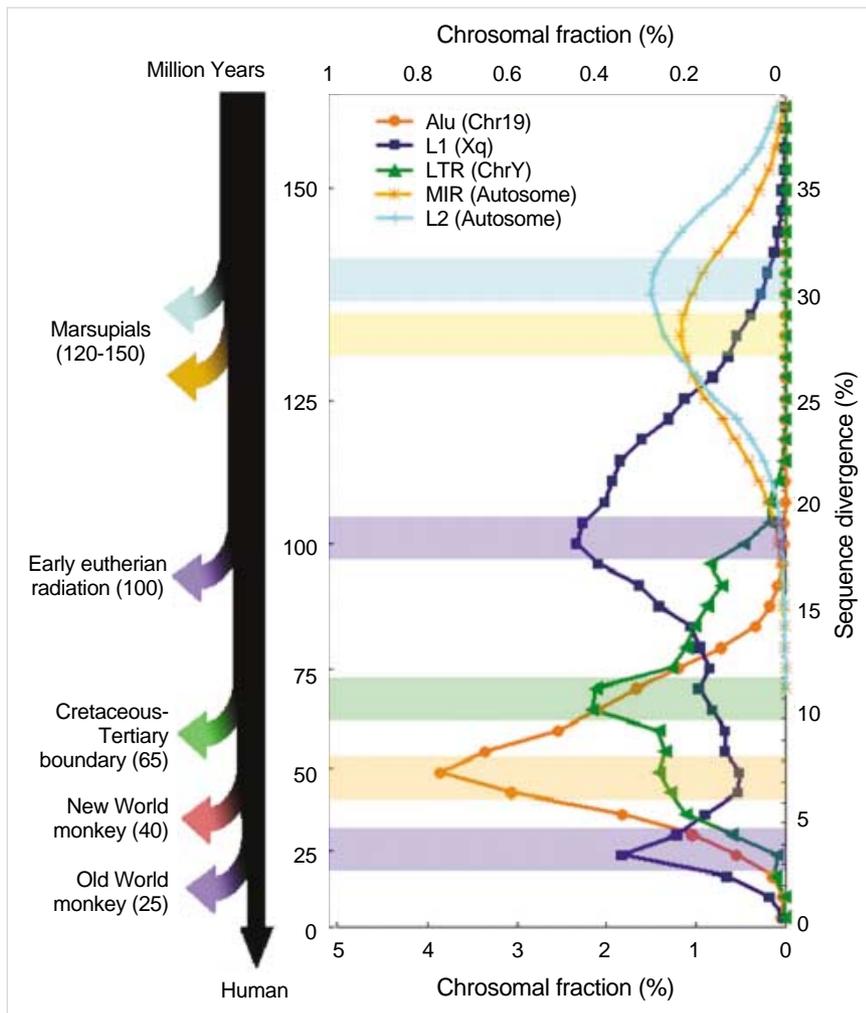
Fig. 5. Correlation between the retroelement peak activity and the branching time of hominoid lineage. The evolutionary waves of Alus, L1s, and LTRs are representatively displayed by the most prominent peak curves on chromosome 19, Xq, and Y, respectively. The faster divergence of CpG-rich Alus and LTRs on recombination-suppressed chromosome Y was taken into consideration for the age recalibration (see Method). The average chromosomal fractions of L2 and MIR are plotted in the scale $5 \times$ magnified, because the peak curves are too small to be compared. The timescale of mammalian evolution previously proposed on the basis of fossil and molecular data is indicated in the vertical axis at the left side. The periods of retroelement peak activity are marked by shadow bar across the schematic diagram.

netic bands of Xq13-Xq21, which contain XIC (X chromosome inactivation center) - known as the main regulatory center of chromosome X inactivation (25). Other spots of old L1 accumulation are associated with dark G-bands and sporadically dispersed along chromosome Xq. The short arm of chromosome X shows a single old L1 cluster at Xp21 while generally suppressed of L1 accumulation. The young L1 copies reveal a relatively random distribution compared to old L1 copies over chromosome X except for XIC.

LTRs on chromosome Y were categorized into the peak (15-17% divergence) and background activities (Fig. 4). LTRs categorized into a peak activity group are specifically concentrated in the middle of the euchromatic portion of the long arm, whereas LTRs at the background activity are uniformly distributed along chromosome Y.

Relationship between retroelement peaking and hominoid primate evolution

The most prominent expansion curves of five retroelement families were collectively plotted in the evolutionary timetable of the primate lineage (Fig. 5). Chromosome 19, X and Y were chosen to represent the peak curves of Alu, L1 and LTR amplification, respectively. All L2s and MIRs in autosomes were analyzed for an unbiased estimate of the peak age, because the number of these old copies on a single chromosome was not enough to represent a normal distribution. Overall, the expansion curves of the five retroelement families display peak activity a total of six times, including two peaks of L1s, each of which is obviously distinguished from the dormant or less active periods of other retroelements. Although the expansion curves of L2s and MIRs are small, a time interval exists between a pair of LINE and SINE activities.

The amplification peaks of ancient L2s and MIRs roughly coincide with the marsupial-placental split about 120-150 MYRA (26). The peak activity of old L1s was found to chronologically correspond to the early eutherian radiation around 100 MYRA (27). An exceptional burst of LTRs on chromosome Y was calibrated at 70 MYRA, closely corresponding to the K-T boundary where extraterritorial impacts were thought to have been accompanied by ecological maelstroms at 65 MYRA (19). The amplification peaks of Alus and young

L1s, both of which are restricted to the primate lineages, coincide with the expansions of the New World monkey and Old World monkey at 40 and 25 MYRA, respectively (28).

DISCUSSION

The evolutionary history of human retroelements has been described as successive amplification of subfamilies during the steady rise and fall of retroelement families (29). Previous genome-wide analysis of LTR and L1 substitutions revealed slow expansion curves with no prominent peaks over the long evolutionary period (2). In this comparative analysis of retroelement divergence on individual chromosomes describes retroelement activity synchronized in an entire chromosomal set. Given that potentially harmful autonomous retroelements have been selectively deleted through recombination processes (14), the synchronous amplification peaks of retroelements in variable densities across chromosomes could be a consequence of post-insertion selection for evolutionary advantage. Recombination-suppressed chromosome Y appears to retain the expansion pattern of LTRs with the least selection. Chromosome X, which requires a high density of L1s to mediate heterochromatization for dosage compensation in response to the small-sized chromosome Y, is thought to have conserved the native expansion pattern of L1s with the least elimination. Therefore, retroelements in individual chromosomes might have amplified concurrently while being fixed differently, suggesting that chromosomal retroelement dynamics is similar in when they thrive and dissimilar in expansion magnitude. The retroelement peak acitivity supports the so-called massive bombardment of retroelements that facilitates the divergence of the host genome (6, 8).

In a breeding population, multicopy retroelements propagate depending on the amplification rates that increase the probability of fixation within a population. Only a significantly large number of newly amplified copies can counter random drift and breed throughout the population. Theoretically, 2N or more insertions are required to allow one insertion to get fixed in a breeding population of N individuals (30). Thus, extremely large numbers of Alus and L1s appeared to have amplified explosively at the time of peak activity that created an excessive number of copies over a population. The expansion curves of the old retroelements such as MIRs and L2s are less prominent as compared with their young counterparts, Alus and L1s and it must be taken into account that some old elements are too diverged to be recognized. However, considering that the normal distribution of neutral substitutions is considerably dispersed around a peak fraction with time, such small amplification peaks are likely to reflect the peak activity of retroelements for the corresponding genome.

Studies have suggested a competitive relationship between autonomous L1 and non-autonomous Alu elements based on the parasitic behavior of Alus that exploit the protein machinery of L1s for retrotransposition (31, 32). This study on the evolutionary wave of retroelements demonstrates that the different retroelement families have never explosively expanded at the same period; thereby coping with the competitive conflicts between selfish DNAs. In addition, it is likely that evolutionary events, such as speciation facilitated by retroelement peaks, provided new niches for the amplification of the other silenced retroelements. Consequently, LINEs and SINEs appeared to reciprocally promote the amplification of each other through repeated reshaping of the host genome.

The dormant periods of retroelements between amplification peaks is noted in terms of the accommodation through the presumptive functionality on host fitness (33) or the stability of retroelements themselves (17). It is believed that the initial integration of L1s and Alus are random while they become differentially redistributed with time (3). It is possible that selfish autonomous L1s are more enriched in the gene-poor regions for lesser detrimental influence, whereas Alus promoting gene expression are more concentrated in the gene-rich region through degeneration of useless contents in the gene-poor region (34). Alternately, the host genome positively selects Alu insertion in the gene-rich regions for adaptive advantage due to its role in promoting protein expression (12). This reciprocal redistribution of LINEs and SINEs in a dormant period would reorganize the explosive insertions for the beneficial and stable relationship between host and retroelement, even for symbiotic co-evolution. Therefore, LINEs and SINEs are likely to have expanded in the host genome as temporally and spatially reciprocal partners.

LTRs were found to have over-accumulated on chromosome Y throughout the period of active amplification. This unique overrepresentation of LTRs can be explained by the lack of recombination in chromosome Y, because other recombination-competent chromosomes are able to eliminate potentially harmful LTRs (14). In fact, we found no LTR peaks on chromosomes other than chromosome Y. The large fraction of amplified LTR copies seems to have been removed through selective recombination due to its adverse influence on the host genome. According to Muller's ratchet and hitchhiking theory for the degeneration of chromosome Y, the non-recombining region is unable to remove the stochastic insertion of retroelements, while tolerating mutational deletion (35, 36). In support of Muller's theory, the evolutionary waves of L1s and LTRs on chromosome Y reveal a skewed expansion curve with a high density of young copies and a low density of old copies. These findings suggest that the rapid gain and loss of genetic contents, rather than recombination, is a main dynamic underlying the unique retroelement composition of chromosome Y.

Of the two amplification peaks made by L1s, the older one was found to largely constitute multiple L1 hot spots in the long arm of chromosome X. A single hot spot of old L1 copies

in the short arm of chromosome X coincides with the ancestral conserved region (XCR) on chromosome Xp described in a previous report (37). The most extensive hot spot is located in the proximal part of chromosome Xq that contains XIC. The high density of L1s in XIC was previously suggested to promote the spread of the X inactivation signal by XIST (X inactive-specific transcript) RNA in support of the Lyon repeat hypothesis (38). Alternately, because the two sex chromosomes have evolved differently in terms of predominant retroelement content, the decreased sequence homology in highly repetitive copies takes advantage for minimizing the undesirable recombination between sex chromosomes (39). The recombination suppression caused by the inversion of chromosome Y (40) is likely to have been further solidified due to different repeated sequences.

A single LTR burst is noted in association with the ancient outbreaks of retroviruses accompanied by the concurrent activation of endogenous retroelements. Exogenous retroviral infection is thought to have activated in trans endogenous retroviruses, which are defected in infectivity due to lack of appropriate viral components (41). The massive horizontal transmission of active exogenous retroviruses was most likely to be critically detrimental for host individuals and populations. The estimated age of an endogenous LTR peak at the K-T boundary implies a sudden change in the global environment, which facilitated evolution and conferred a niche for the LTR expansion in the primate ancestral genome. The Earth's ecosystem and climate are believed to have been severely disturbed at the K-T boundary at 65 MYRA (42, 43). These changes are likely to have modified habitats and genomic DNA through large quantities of dust and radioactive gas, consequently promoting the speciation of mammals. Therefore, a LTR amplification peak is possible in association with such a special global event, but not an accidental event.

The global changes at the K-T boundary of 65 MYRA are believed to have had important consequences for diversification of terrestrial organisms. Some investigators have used a placental mammalian evolution in a molecular clock to suggest the inter- and intra-ordinal divergence of placental mammals before and after the K-T boundary, respectively (27). In this respect, the human LTR explosion at 70 MYRA is timely compatible with the emergence of prosimians at the K-T boundary. The peak of old L1 copies appearing before the LTR peak is correlated with the inter-ordinal divergence of early placental mammals about 100 MYRA (44). The ancient L2s and MIRs appeared to expand explosively during the inter-ordinal divergence of marsupial and eutherian lineages about 120-150 MYRA. The peaks of Alus (40 MYRA) and young L1s (25 MYRA) following LTR peaking could be responsible for the active intra-ordinal diversification of the Old World monkey and New World money, respectively, in the hominoid lineage. The LTR explosion independent of LINE/SINE expansion is thought to have converted the evolutionary impact of reciprocal LINE/SINE expansion from inter- to intra-ordinal eutherian divergence.

In conclusion, this comparative analysis of retroelement divergence among individual chromosomes sheds light on the genome dynamics driven by the peak activity of two monophyletic selfish DNAs, SINEs and LINEs. The ebb-and-flow-like expansion curves of LINEs and SINEs suggest that the human genome has evolved through repeated SINE/LINE amplification and exaptation for 150 million years. An exceptional expansion peak of LTRs associated with the K-T boundary is considered a pivotal event in the inter- and intra-ordinal transition of eutherian diversification that reshuffled the programmed genome dynamics.

ACKNOWLEDGEMENTS

REFERENCES

1. Smit AF. *The origin of interspersed repeats in the human genome. Curr Opin Genet Dev 1996; 6: 743-8.*

2. International Human Genome Sequencing Consortium. *Initial sequencing and analysis of the human genome. Nature 2001; 409: 860-921.*

3. Deininger PL, Batzer MA. *Mammalian retroelements. Genome Res 2002; 12: 1455-65.*

4. Brookfield JF. *Selection on Alu sequences? Curr Biol 2001; 11: R900-1.*

5. Mighell AJ, Markham AF, Robinson PA. *Alu sequences. FEBS Lett 1997; 417: 1-5.*

6. Brosius J. *Genomes were forged by massive bombardments with retroelements and retrosequences. Genetica 1999; 107: 209-38.*

7. Hamdi HK, Nishio H, Tavis J, Zielinski R, Dugaiczyk A. *Alu-mediated phylogenetic novelties in gene regulation and development. J Mol Biol 2000; 299: 931-9.*

8. Sverdlov ED. *Retroviruses and primate evolution. Bioessays 2000; 22: 161-71.*

9. Ostertag EM, Kazazian HH Jr. *Biology of mammalian L1 retrotransposons. Annu Rev Genet 2001; 35: 501-38.*

10. Gu Z, Wang H, Nekrutenko A, Li WH. *Densities, length proportions, and other distributional features of repetitive sequences in the human genome estimated from 430 megabases of genomic sequence. Gene 2000; 259: 81-8.*

11. Jurka J, Kapitonov VV. *Sectorial mutagenesis by transposable elements. Genetica 1999; 107: 239-48.*

12. Chu WM, Ballard R, Carpick BW, Williams BR, Schmid CW. *Potential Alu function: regulation of the activity of double-stranded RNA-activated kinase PKR. Mol Cell Biol 1998; 18: 58-68.*

13. Erlandsson R, Wilson JF, Paabo S. *Sex chromosomal transposable element accumulation and male-driven substitutional evolution in humans. Mol Biol Evol 2000; 17: 804-12.*

14. Boissinot S, Entezam A, Furano AV. *Selection against deleterious LINE-1-containing loci in the human lineage. Mol Biol Evol 2001;*

18: 926-35.

15. Bailey JA, Carrel L, Chakravarti A, Eichler EE. *Molecular evidence for a relationship between LINE-1 elements and X chromosome inactivation: the Lyon repeat hypothesis. Proc Natl Acad Sci USA 2000; 97: 6634-9.*

16. Jurka J. *Repeats in genomic DNA: mining and meaning. Curr Opin Struct Biol 1998; 8: 333-7.*

17. Pavlicek A, Jabbari K, Paces J, Paces V, Hejnar JV, Bernardi G. *Similar integration but different stability of Alus and LINEs in the human genome. Gene 2001; 276: 39-45.*

18. Ovchinnikov I, Troxel AB, Swergold GD. *Genomic characterization of recent human LINE-1 insertions: evidence supporting random insertion. Genome Res 2001; 11: 2050-8.*

19. Shukolyukov A, Lugmair GW. *Isotopic evidence for the Cretaceous-Tertiary impactor and its type. Science 1998; 282: 927-9.*

20. Jurka J. *Repbase update: a database and an electronic journal of repetitive elements. Trends Genet 2000; 16: 418-20.*

21. Smith TF, Waterman MS. *Identification of common molecular subsequences. J Mol Biol 1981; 147: 195-7.*

22. Graves JA. *The origin and function of the mammalian Y chromosome and Y-borne genes--an evolving understanding. Bioessays 1995; 17: 311-20.*

23. Kapitonov V, Jurka J. *The age of Alu subfamilies. J Mol Evol 1996; 42: 59-65.*

24. Kjellman C, Sjogren HO, Widegren B. *The Y chromosome: a graveyard for endogenous retroviruses. Gene 1995; 161: 163-70.*

25. Rastan S. *Non-random X-chromosome inactivation in mouse X-autosome translocation embryos--location of the inactivation centre. J Embryol Exp Morphol 1983; 78: 1-22.*

26. Hope RM, Cooper S, Wainwright B. *Globin macromolecular sequences in marsupials and monotremes. In: Graves JA, Hope RM, Cooper DW, editors, Mammals from Pouches and Eggs: Genetics, Breeding and Evolution of Marsupials and Monotremes. Melbourne: CSIRO, 1990: 147-72.*

27. Springer MS, Murphy WJ, Eizirik E, O'Brien SJ. *Placental mammal diversification and the Cretaceous-Tertiary boundary. Proc Natl Acad Sci USA 2003; 100: 1056-61.*

28. Kumar S, Hedges SB. *A molecular timescale for vertebrate evolution. Nature 1998; 392: 917-20.*

29. Smit AF. *Interspersed repeats and other momentos of transposable elements in mammalian genomes. Curr Opin Genet Dev 1999; 9: 657-63.*

30. Hartl DL, Clark AG. *Principles of population genetics, 2nd ed. 1989.*

31. Ohshima K, Hamada M, Terai Y, Okada N. *The 3´ ends of tRNA-derived short interspersed repetitive elements are derived from the 3´ ends of long interspersed repetitive elements. Mol Cell Biol 1996; 16: 3756-64.*

32. Feng Q, Moran JV, Kazazian HH Jr, Boeke JD. *Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. Cell 1996; 87: 905-16.*

33. Makalowski W. *Genomic scrap yard: how genomes utilize all that junk. Gene 2000; 259: 61-7.*

34. Jabbari K, Bernardi G. *CpG doublets, CpG islands and Alu repeats in long human DNA sequences from different isochore families. Gene 1998; 224: 123-7.*

35. Charlesworth B. *The evolution of chromosomal sex determination and dosage compensation. Curr Biol 1996; 6: 149-62.*

36. Bachtrog D, Charlesworth B. *Reduced adaptation of a non-recombining neo-Y chromosome. Nature 2002; 416: 323-6.*

37. Graves JA. *The rise and fall of SRY. Trends Genet 2002; 18: 259-64.*

38. Lyon MF. *X-chromosome inactivation: a repeat hypothesis. Cytogenet Cell Genet 1998; 80: 133-7.*

39. Hsu SJ, Erickson RP, Zhang J, Garver WS, Heidenreich RA. *Fine linkage and physical mapping suggests cross-over suppression with a retroposon insertion at the npc1 mutation. Mamm Genome 2000; 11: 774-8.*

40. Lahn BT, Page DC. *Four evolutionary strata on the human X chromosome. Science 1999; 286: 964-7.*

41. Walsh CP, Chaillet JR, Bestor TH. *Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. Nat Genet 1998; 20: 116-7.*

42. Alvareaz L, Alvarez W, Asaro F, Michel H. *Extraterrestrial Cause for the Cretaceous-Tertiary Extinction. Science 1980; 208: 1095-180.*

43. Kerr R. *Testing an Ancient Impact's Punch. Science 1994; 263: 1371-2.*

44. Easteal S. *Molecular evidence for the early divergence of placental mammals. Bioessays 1999; 21: 1052-8.*