Transposable element insertions have strongly affected human evolution

Roy J. Britten[1]

Division of Biology, California Institute of Technology, Pasadena, CA 91125; and Department of Ecology and Evolution, University of California, Irvine, CA 92697

Contributed by Roy J. Britten, October 4, 2010 (sent for review June 7, 2010)

Comparison of a full collection of the transposable element (TE) sequences of vertebrates with genome sequences shows that the human genome makes 655 perfect full-length matches. The cause is that the human genome contains many active TEs that have caused TE inserts in relatively recent times. These TE inserts in the human genome are several types of young Alus (AluYa5, AluYb8, AluYc1, etc.). Work in many laboratories has shown that such inserts have many effects including changes in gene expression, increases in recombination, and unequal crossover. The time of these very effective changes in the human lineage genome extends back about 4 million years according to these data and very likely much earlier. Rapid human lineage-specific evolution, including brain size is known to have also occurred in the last few million years. Alu insertions likely underlie rapid human lineage evolution. They are known to have many effects. Examples are listed in which TE sequences have influenced human-specific genes. The proposed model is that the many TE insertions created many potentially effective changes and those selected were responsible for a part of the striking human lineage evolution. The combination of the results of these events that were selected during human lineage evolution was apparently effective in producing a successful and rapidly evolving species.

Alu sequences | speed of evolution | last 3 myr

The aim of this paper is an explanation for the high speed of evolution of the human lineage, which has been exceptional compared with other animals. The high speed of evolution of human lineage brain size is recognized by comparison of fossil brain sizes (1, 2). Evolution of the lineage leading to humans during the last several million years was striking. In this period the brain in our lineage tripled in mass (1, 2). The function of the brain also changed rapidly but there are few useful fossils. What we know is that the result was the modern human brain, which has been called the most complex thing in the universe. We believe the brain evolution was due to natural selection and genomic variation.

A major source of variation has been the insertion of transposable elements (TEs). They can be identified as catalysts of evolution because their contribution to variation increased the speed of evolution (3–8). TE element insertions increase the rate of recombination (3) and when there are already many copies present nearby as there are for Alu elements the new ones increase the rate of unequal crossover. The insertions affect genes and their expression (8). Humans stand alone in two respects: the speed of evolution and the large number and activity of TEs. This recognition leads directly to the proposal that they are functionally connected. In other words the high frequency of TE insertions is responsible at least in part for the rapid human evolution. It might be due to the increased variation and recombination that certain specific sets of genes were activated or suppressed or the increased total number of opportunities for useful variation.

Results

The first stage of this work is the examination of recent inserts of TEs in the human genome. A collection (3) of the sequences of all of the TEs of vertebrates is here called vertte and includes 2,732 TEs of which about 1,700 match human DNA in a Wublast comparison. Vertte was compared with human DNA sequences (*Methods*). The results were filtered to include only perfect full-length matches and are shown in Table 1, where column 1 gives the number of perfect matches of each element for which perfect matches occur. It is remarkable that only Alus make numbers of perfect matches although all human TEs are in the vertte collection. Three of the TEs made only single perfect matches, indicating their presence in the human genome but not suggesting great insertion activity. The descriptions are from GIRI (3), also the source of the vertte list. All of the examples with multiple copies in Human DNA are AluY families, including AluYa5, AluYb8, etc., as shown in Table 1 and these are known as young Alus (3). There is a total of 655 such recent inserts in human DNA.

Table 2 reports the results of the same assay of recent TE inserts done on chimpanzee DNA using the same vertte collection. The results are very different from those for human DNA. The total number of perfect inserts is 283 and only one principal known element has apparently been active in producing a large number of inserts. Jerzy Jurka in conversation stated that a reason for this difference could be that the TEs of the human genome have been more studied. However Mills et al. (4) have studied the new insertions of mobile elements since the split between the lineages leading to chimpanzee and human. Their method was to align the *Pan* and *Homo sapiens* genomes and identify the gaps in each, further requiring that the apparent insert in the other species be bounded by target site duplications (TSDs). Thus the number was independent of the species-specific knowledge of the transposable elements and avoids Jurka's criticism. They observed 5,530 new Alu TEs in human and 1,642 new Alu TEs in chimpanzee in ~6 million years. This will be further examined below when considering the older copies we observe.

Vertte was also compared with eight other vertebrate genomic DNAs (orangutan, macaque, mouse, zebrafish, pufferfish, opossum, and platypus). None show many perfect matches and thus no apparent large amounts of TE activity. These observations are limited by the absence of full studies of the TEs of these species. Thus we cannot safely estimate the amount of TE activity, The situation at present is that aside from human and chimp no highly active TEs have been identified that have been inserting many copies, but the studies are incomplete and will be reported in a future paper.

Perfect copies last for a limited time due to the rate of mutation. The DNA sequence difference between chimp and human bulk DNA sequences is 1.2% (5), which has occurred in about 6 Myr and thus 1.2/12 = 0.1% per million years in each lineage. I am assuming that the rate of mutation of Alu sequences is about the same as the neutral drift of the human genome. In other words there are not positive or negative selective forces affecting Alu inserts unless they are incorporated into a gene. This implies

Author contributions: R.J.B. designed research, performed research, contributed new reagents/analytic tools, analyzed data, and wrote the paper.

The author declares no conflict of interest.

[1]E-mail: r.britten@ca.rr.com.

Table 1. Perfect multiple matches of TEs to human DNA

Number	Description of TE	
14 > AluY	SINE1/7SL	Primates
14 > AluYa4	SINE1/7SL	Primates
268 > AluYa5	SINE1/7SL	*Homo sapiens*
9 > AluYa8	SINE1/7SL	Primates
187 > AluYb8	SINE1/7SL	*Homo sapiens*
32 > AluYb9	SINE1/7SL	*Homo sapiens*
84 > AluYc1	SINE1/7SL	Primates
21 > AluYc2	SINE1/7SL	Primates
17 > AluYd8	SINE1/7SL	Primates
4 > AluYf2	SINE1/7SL	Primates
2 > AluYh9	SINE1/7SL	Primates

the occurrence of a mutation in a 280 long Alu about once per 3 Myr but allowing for the greater rate of mutations in CpGs the time estimate is about 1.5 Myr per mutation per Alu. Thus the best estimate of recent (perfect) Alu insertions is from the present back to around 1.5 Mya.

Discussion

This is an extraordinary correlation. Human evolution has been rapid, particularly brain evolution in the last several million years. It is the only species known to make such rapid evolutionary progress. Now it is shown that human is the only species studied to have so many TE insertions. Recognition of this correlation leads to the concept that Alu insertions underlie rapid human evolution. Human DNA also shows a large amount of TE polymorphism (6).

The hominin fossil record documents a history of evolutionary events including the origins of bipedalism; the emergence of our genus *Homo*; the first manufacture and use of flaked stone tools; increases in brain size and our complex culture, much of this change occurring in the last 3 Myr. The molecular evidence of recent 100% matching TE inserts applies to the last 1.5 Myr. The insertion rates were apparently faster in earlier times as is shown by counting the number of insertions that have had mutations. The results for zero, one, and two mutations are shown in Table 3. The second column is a copy of Table 1. It is clear that this "zero" column represents events near to the present, and the "one" column events farther in the past, and the "two" column events still farther in the past. There are more inserts in the "two mutation" column so we can conclude that the Alus were more active in earlier times. Precise times cannot be calculated from so few mutations but because the rate of mutation is about one per Alu per 1.5 Myr, two-mutation examples go back to about 3–4.5 Myr on average.

It appears from Table 3 that the rate of human lineage Alu insertion was higher in earlier times. This conclusion is supported by the Mills et al. (4) observation of 5,530 Alu insertions fixed in the human lineage in the ~6 Myr since the split of human and chimpanzee lineages.

Alu inserts increase the rate of recombination (3, 7). And many inserts affect the regulation of genetic activity of nearby genes (7, 8). Because there are many Alus present, unequal crossover can occur, opening the possibility of duplications and deletions of inter-Alu regions. All these processes together are sources of variation. There are at present many Alu insert polymorphisms. One hundred examples of polymorphic Alu insertions in proteins have been used for human population genetics studies (6). These are the result of relatively recent insertions.

Table 2. Perfect multiple matches of TEs to chimpanzee DNA

Number	Description of TE	
7 > AluY	SINE1/7SL	Primates
268 > AluYc1	SINE1/7SL	Primates
4 > PTERV1c_LTR	ERV1	Pan troglody
2 > AluYi6	SINE1/7SL	Primates

Table 3. TE insert count human DNA with zero, one, and two mutations

Id no.	Zero	One	Two			
21	14	118	216	>AluY	SINE1/7SL	
22	0	18	132	>AluYa1	SINE1/7SL	
23	14	316	616	>AluYa4	SINE1/7SL	
24	268	613	540	>AluYa5	SINE1/7SL	*H. sapiens*
25	9	13	7	>AluYa8	SINE1/7SL	Primates
26	0	0	2	>AluYb3a1	SINE1/7SL	Primates
28	187	441	402	>AluYb8	SINE1/7SL	*H. sapiens*
29	32	220	458	>AluYb9	SINE1/7SL	*H. sapiens*
31	84	187	160	>AluYc1	SINE1/7SL	Primates
32	21	105	179	>AluYc2	SINE1/7SL	Primates
33	0	0	7	>AluYc5	SINE1/7SL	Primates
37	17	23	29	>AluYd8	SINE1/7SL	Primates
39	0	0	33	>AluYe5	SINE1/7SL	Primates
40	0	14	126	>AluYf1	SINE1/7SL	Primates
41	4	3	8	>AluYf2	SINE1/7SL	Primates
42	0	18	43	>AluYg6	SINE1/7SL	Primates
43	2	0	9	>AluYh9	SINE1/7SL	Primates
44	1	4	16	>AluYi6	SINE1/7SL	Primates
715	0	1	4	>MADE1	Mariner/Tc1	*H. sapiens*
Sum	655	2,093	2,987			

In a paper entitled "Evolutionary impact of human Alu repetitive elements" (9) the focus is on Alu element history with comments on recombination and regional duplication. One paper describes the structural variations caused by TE insertions (10) and another identifies TEs as the drivers of evolution (11). Taken together the content of these papers shows that the Alu insertions summarized here have had enormous effect on human evolution. There are many publications on the effect of TEs on the genome (e.g., refs. 12–16). The effect of Alu insertions on specific genes has been studied (16–19). Further studies have been made of the effect of TEs on the human genome and its expression (20–29).

To further explore the effects of these Alu inserts Wublast comparison was made with the set of refmRNA sequences, which are principally coding sequences. It is found that the eight principal Alus in Table 1 make significant matches to 4,000 of the nearly 28,000 messenger RNAs in this collection. Only one match was 100%—a recent insertion in the *MBOAT-1* gene. Most of the matches are quite divergent and indicate the presence of Alu sequences in general, rather than the specific named Alus used as probes.

The evolution of humans is exceptional among all of the millions of animals. The lineage leading to evolution of humans must have branched as our ancestors became able to produce advanced stone tools, then continued to advance through language to modern society. It is hard to give a precise date when the lineage leading to humans first advanced beyond what any other animal has ever achieved, but it probably happened about the time the growth in brain size really got under way. It is a fair guess that it was about the time of *Homo habilis* (1, 2) or about 2 Mya. The evidence presented here in Table 3 indicates that the large number of insertions of Alus was underway at that time, supporting the concept that the Alu insertions played a major role in the human lineage evolution. The values in column 3 have been corrected because 28 (AluYb8) and 29 (AluYb9) differ in sequence by only one nucleotide and thus give column 3 (one mutation) values that are too large. The same is true for 31 (AluYc1) and 32 (AluYc2), which were also corrected.

The human genome contains about 1.8 million recognizable Alu sequence residues, which have been inserted over many tens of millions of years of primate evolution. This larger than usual estimate is based on Wublast comparisons with an expectation of 1e-3 or less. The large number opens the question of what was special about the last few million years that allowed the insertion of a few thousand new Alus to have such a striking effect. The best that can be said is that all of the aspects of human lineage biology had evolved to a stage of readiness. The role of Alu sequences has been discussed for many years, initially being considered junk, then selfish, and then as more was learned of actual gene sequences, Alus were recognized as contributing to gene regulation. They are now seen as having much structural and evolutionary significance. It has been a bumpy road but there is no doubt about the validity of the current view, based, as it is, on clear evidence and many publications.

The location of the 655 perfect Alu copies resulting from recent insertions was examined. They are more or less evenly distributed from end to end of the genome in what appears a random pattern. There is some clustering but no more than was observed with a model calculation with a comparable number of random number generated locations.

Because conserved noncoding sequences (CNS) are often involved in the regulation of gene activity, it seemed worthwhile to find whether they occur near Alu inserts. Briefly, I found no evidence of such a correlation. The test was to identify human sequences with better than 75% match to *Takifugu rubripes* and rice genomes. About one out of five of the 655 10,000-nt-long human sequences containing perfect Alu inserts have such apparent CNS. In addition I found that the great majority are members of low-frequency sequence families in both Fugu and rice. The sequences that I looked at were all simple sequences, reducing my confidence that they were really active CNS. More study is required to resolve the original question of a correlation between CNS and TE recent inserts.

The evolutionary events such as development of bipedalism and brain growth depend on natural selection and various structural and biochemical opportunities that are part of the evolutionary processes. The TE insertions principally cause variation through increase in the rate of recombination (3, 7) and affecting the regulation of genetic activity of nearby genes (8–10) as well as opening new possibilities for unequal crossover. I assume that the insertions also cause variation by mechanisms yet to be explored. Taken together these effects speed up the process of evolution in the human lineage.

It is not likely that this speculative model will be immediately accepted because the habit of considering Alu elements as junk DNA is widespread, whereas this model places them in a central position in human evolution. The following paragraphs describe direct evidence of the action of Alu sequence in some evolutionarily important examples that support this model.

An example is described in the work of A. Varki (16) as follows from the abstract: "Specific events include Alu-mediated inactivation of the *CMAH* gene, resulting in loss of synthesis of the Sia *N*-glycolylneuraminic acid (Neu5Gc) and increase in expression of the precursor *N*-acetylneuraminic acid (Neu5Ac); increased expression of alpha2-6–linked Sias (likely because of changed expression of *ST6GALI*); and multiple changes in *SIGLEC* genes encoding Sia-recognizing Ig-like lectins (Siglecs)." Hayakawa, et al. (17) state "We have found that, although a region containing a 92-bp exon and an AluSq element in the hydroxylase gene is intact in all nonhuman primates examined, the same region in the human genome is replaced by an AluY element that was disseminated at least one million years ago. We propose a mechanistic model for this Alu-mediated replacement event, which deleted the 92-bp exon and thus inactivated the human hydroxylase gene. It is suggested that Alu elements have played potentially important roles in genotypic and phenotypic evolution in the hominid lineage."

Li, et al. (18) write in their abstract: "To understand whether any human-specific new genes may be associated with human brain functions, we computationally screened the genetic vulnerable factors identified through genome-wide association studies and linkage analyses of nicotine addiction and found one human-specific de novo protein-coding gene, *FLJ33706* (alternative gene symbol *C20*or*f203*). Cross-species analysis revealed interesting evolutionary paths of how this gene had originated from noncoding DNA sequences: insertion of repeat elements especially Alu contributed to the formation of the first coding exon and six standard splice junctions on the branch leading to humans and chimpanzees, and two subsequent substitutions in the human lineage escaped two stop codons and created an open reading frame of 194 amino acids. We experimentally verified FLJ33706's mRNA and protein expression in the brain."

Zhang and Chasin (19) in a paper titled "Comparison of multiple vertebrate genomes reveals the birth and evolution of human exons," state in the abstract: "Remarkably, the great majority of new cassette exons are composed of highly repeated sequences, especially Alu."

Microcephalin (*MCPH1*) is an example of a different kind (20). MCPH1 is one of a group of genes involved in the establishment of human brain size and has a history of insertions by TEs. It has 14 exons and many of the introns include TE sequences. Curiously one of the exons (last or no. 14) includes an AluY sequence at 88% precision. It includes nucleotides 50–283 of the Alu and occupies positions 301–534 of that exon, which is 625 nt long. The total number of TEs recognized in the gene is in the thousands. If limits are set to longer than 70 nt and better than 70% accuracy, 450 fragment matches are seen. The total length of these matches is 24% of the length of this gene. The total without limits indicates that 57% of the *MCPH1* gene length is TEs. These intron insertions have been occurring for a long period in the past as implied by the large amount of divergence leading to large mismatches. Each of these many inserts in the introns could have had gene regulatory effects, affecting brain growth. There are many human genes that include TEs in introns or nearby but this one is of importance because of its effect on brain growth.

A major issue is the way TEs contribute variation and affect genome function. Each insertion has great potential power due to a wide range of possible mechanisms. The publication of the chimp genome (5) shows that chimp DNA contains many fewer Alus than human. Comparing Tables 1 and 2 shows that human DNA includes 12 new Alus, whereas chimp DNA contains only 5, only 1 of which is active and has many copies. The last common ancestor of chimp and human existed about 6 Mya and presumably had no more Alus than chimp. The implication is that during the last 6 Myr processes in the human genome generated at least seven new families of Alus including master genes responsible for the many insertions now observed and shown in Table 1. Many of these events affected the evolution of the human genome and many of them must have occurred during the few million year period considered in this paper. The simple insertion of an Alu may increase the rate of recombination or by interacting with Alus already present cause unequal crossover. It may change the regulation of genes by inserting in an intron or nearby a gene (12–14). Another possible affect is the production of RNA interference molecules (21) and other regulatory RNAs that have not been fully explored yet.

There is much evidence (22–26) of the effect of Alu insertions on the human genome and its function. Recent work (27) has shown that noncoding (regulatory) changes have dominated human evolution. Work from Eviator Nevo's group (28) focuses on editing of transcripts and because Alu TEs are among the high rate targets for this process, they find that the editing level in the transcripts analyzed is higher in human brain compared with nonhuman primates. Also the editing level is higher in human compared with chimpanzee. Moreover, new editable species-specific Alu insertions,

subsequent to the human–chimpanzee split, are significantly enriched in genes related to neuronal functions and neurological diseases. The authors conclude that editing of transcripts has a strong influence on human evolution, specifically on brain evolution.

Whereas the evidence presented relates entirely to Alu sequences, it is likely that other TEs are also involved as indicated by the data of Lowe et al. (30) who show that in human DNA many fragments of TEs have contributed to conserved noncoding elements that are specifically near genes. They conclude by writing "Indeed, as our appreciation for the contributions of repeats to different aspects of genome evolution continues to grow, it now seems that these unwanted, and often ignored, children of the genome played multiple crucial roles during the evolution of the human lineage."

Conclusion

TE insertions occur frequently during human lineage evolution. Table 3 shows that the rate of human insertion was larger in the past back to about 3–4.5 Mya and probably much earlier. This observation has led me to speculate that Alu insertions underlie rapid human evolution. There is no doubt about the occurrence of many TE insertions and some of these affected human gene expression. A question is whether there were as many as required and this issue suggests the model is speculative. The proposed model is that the many TE insertions created many potentially effective changes and those changes selected were responsible for the striking human evolution. The number of potential events was effective in the combination selected during human lineage evolution. TE activity has likely been responsible for the speed and exceptional character of human lineage evolution.

Methods

All comparisons were made with Wublast because it reports divergent matches at preferred sensitivity. A collection of all of the listed vertebrate mobile elements was downloaded from GIRI (3). This collection identified here as vertte includes 2,732 TEs of which about 1,700 match human DNA with any divergence that has a Wublast expectation of 1e-3 or less. The human genome version 36 was converted into 10,000-nt segments after removal of all NNNs. Comparisons testing for the occurrence of TE inserts were made between the vertte collection and the 10,000-nt segments of human DNA version 36. Similar 10,000-nt segments were prepared for each of the species studied. Fortran software was written to interpret the results.

ACKNOWLEDGMENTS. Louis Bouwer did much of the computer work supported by Eric H. Davidson.

1. Rogers J, et al. (2010) On the genetic architecture of cortical folding and brain volume in primates. *Neuroimage* 53:1103–1108.
2. Park MS, et al. (2007) Evolution of the human brain: Changing brain size and the fossil record. *Neurosurgery* 60:555–562, discussion 562.
3. Jurka J (2000) Repbase update: A database and an electronic journal of repetitive elements. *Trends Genet* 16:418–420.
4. Mills RE, et al. (2006) Recently mobilized transposons in the human and chimpanzee genomes. *Am J Hum Genet* 78:671–679.
5. Chimpanzee Sequencing and Analysis Consortium (2005) Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* 437:69–87.
6. Watkins WS, et al. (2003) Genetic variation among world populations: Inferences from 100 Alu insertion polymorphisms. *Genome Res* 13:1607–1618.
7. Witherspoon DJ, et al. (2009) Alu repeats increase local recombination rates. *BMC Genomics* 10:530.
8. Ponicsan SL, Kugel JF, Goodrich JA (2010) Genomic gems: SINE RNAs regulate mRNA production. *Curr Opin Genet Dev* 20:149–155.
9. Jurka J (2004) Evolutionary impact of human Alu repetitive elements. *Curr Opin Genet Dev* 14:603–608.
10. Xing J, Witherspoon DJ, Ray DA, Batzer MA, Jorde LB (2007) Mobile DNA elements in primate and human evolution. *Am J Phys Anthropol* 45(Suppl 45):2–19.
11. Kazazian HH, Jr. (2004) Mobile elements: Drivers of genome evolution. *Science* 303: 1626–1632.
12. Medstrand P, et al. (2005) Impact of transposable elements on the evolution of mammalian gene regulation. *Cytogenet Genome Res* 110:342–352.
13. van de Lagemaat LN, Landry JR, Mager DL, Medstrand P (2003) Transposable elements in mammals promote regulatory variation and diversification of genes with specialized functions. *Trends Genet* 19:530–536.
14. Landry JR, Medstrand P, Mager DL (2001) Repetitive elements in the 5′ untranslated region of a human zinc-finger gene modulate transcription and translation efficiency. *Genomics* 76:110–116.
15. Bennett EA, Coleman LE, Tsui C, Pittard WS, Devine SE (2004) Natural genetic variation caused by transposable elements in humans. *Genetics* 168:933–951.
16. Varki A (2010) Colloquium paper: Uniquely human evolution of sialic acid genetics and biology. *Proc Natl Acad Sci USA* 107(Suppl 2):8939–8946.
17. Hayakawa T, Satta Y, Gagneux P, Varki A, Takahata N (2001) Alu-mediated inactivation of the human CMP- N-acetylneuraminic acid hydroxylase gene. *Proc Natl Acad Sci USA* 98:11399–11404.
18. Li CY, et al. (2010) A human-specific de novo protein-coding gene associated with human brain functions. *PLoS Comput Biol* 6(3):e1000734.
19. Zhang XH, Chasin LA (2006) Comparison of multiple vertebrate genomes reveals the birth and evolution of human exons. *Proc Natl Acad Sci USA* 103:13427–13432.
20. Rimol LM, et al. (2010) Sex-dependent association of common variants of microcephaly genes with brain structure. *Proc Natl Acad Sci USA* 107:384–388.
21. Gasior SL, Palmisano M, Deininger PL (2006) Alu-linked hairpins efficiently mediate RNA interference with less toxicity than do H1-expressed short hairpin RNAs. *Anal Biochem* 349:41–48.
22. Konkel MK, Batzer MA (2010) A mobile threat to genome stability: The impact of non-LTR retrotransposons upon the human genome. *Semin Cancer Biol* 20:211–221.
23. Cordaux R, Udit S, Batzer MA, Feschotte C (2006) Birth of a chimeric primate gene by capture of the transposase gene from a mobile element. *Proc Natl Acad Sci USA* 103: 8101–8106.
24. Deininger P (1993) Induction of DNA rearrangement and transposition. *Proc Natl Acad Sci USA* 90:3780–3781.
25. Salem AH, et al. (2003) Alu elements and hominid phylogenetics. *Proc Natl Acad Sci USA* 100(22):12787–12791.
26. Jurka J (1997) Sequence patterns indicate an enzymatic involvement in integration of mammalian retroposons. *Proc Natl Acad Sci USA* 94:1872–1877.
27. Haygood R, Babbitt CC, Fedrigo O, Wray GA (2010) Contrasts between adaptive coding and noncoding changes during human evolution. *Proc Natl Acad Sci USA* 107: 7853–7857.
28. Paz-Yaacov N, et al. (2010) Adenosine-to-inosine RNA editing shapes transcriptome diversity in primates. *Proc Natl Acad Sci USA* 107:12174–12179.
29. Bennett EA, et al. (2008) Active Alu retrotransposons in the human genome. *Genome Res* 18:1875–1883.
30. Lowe CB, Bejerano G, Haussler D (2007) Thousands of human mobile element fragments undergo purifying selection near developmental genes. *Proc Natl Acad Sci USA* 104:8005–8010.