Mobile elements create structural variation: Analysis of a complete human genome

Jinchuan Xing,[1] Yuhua Zhang,[1] Kyudong Han,[2] Abdel Halim Salem,[2,3,5] Shurjo K. Sen,[2,6] Chad D. Huff,[1] Qiong Zhou,[1] Ewen F. Kirkness,[4] Samuel Levy,[4] Mark A. Batzer,[2] and Lynn B. Jorde[1,7]

[1]Department of Human Genetics, Eccles Institute of Human Genetics, University of Utah, Salt Lake City, Utah 84109, USA; [2]Department of Biological Sciences, Louisiana State University, Baton Rouge, Louisiana 70803, USA; [3]Department of Anatomy, Faculty of Medicine, Suez Canal University, Ismailia 41111, Egypt; [4]J. Craig Venter Institute, Rockville, Maryland 20850, USA

Structural variants (SVs) are common in the human genome. Because approximately half of the human genome consists of repetitive, transposable DNA sequences, it is plausible that these elements play an important role in generating SVs in humans. Sequencing of the diploid genome of one individual human (HuRef) affords us the opportunity to assess, for the first time, the impact of mobile elements on SVs in an individual in a thorough and unbiased fashion. In this study, we systematically evaluated more than 8000 SVs to identify mobile element-associated SVs as small as 100 bp and specific to the HuRef genome. Combining computational and experimental analyses, we identified and validated 706 mobile element insertion events (including *Alu*, L1, SVA elements, and nonclassical insertions), which added more than 305 kb of new DNA sequence to the HuRef genome compared with the Human Genome Project (HGP) reference sequence (hg18). We also identified 140 mobile element-associated deletions, which removed ~126 kb of sequence from the HuRef genome. Overall, ~10% of the HuRef-specific indels larger than 100 bp are caused by mobile element-associated events. More than one-third of the insertion/deletion events occurred in genic regions, and new *Alu* insertions occurred in exons of three human genes. Based on the number of insertions and the estimated time to the most recent common ancestor of HuRef and the HGP reference genome, we estimated the *Alu*, L1, and SVA retrotransposition rates to be one in 21 births, 212 births, and 916 births, respectively. This study presents the first comprehensive analysis of mobile element-related structural variants in the complete DNA sequence of an individual and demonstrates that mobile elements play an important role in generating inter-individual structural variation.

[Supplemental material is available online at http://www.genome.org. The sequence data from this study have been submitted to GenBank (http://www.ncbi.nlm.nih.gov/Genbank/) under accession nos. FI569689–FI569698.]