REFERENCES

Chang, P.Y., Hom, R.A., Musselman, C.A., Zhu, L., Kuo, A., Gozani, O., Kutateladze, T.G., and Cleary, M.L. (2010). J. Mol. Biol. Published online May 7, 2010. 10.106/j.jmb.2010.05.005.

Chen, J., Santillan, D.A., Koonce, M., Wei, W., Luo, R., Thirman, M.J., Zeleznik-Le, N.J., and Diaz, M.O. (2008). Cancer Res. *68*, 6199–6207.

Hom, R.A., Chang, P.Y., Roy, S., Musselman, C.A., Glass, K.C., Selezneva, A.I., Gozani, O., Ismagilov, R.F., Cleary, M.L., and Kutateladze, T.G. (2010). J. Mol. Biol. Published online May 8, 2010. 10.1016/j.jmb.2010.04.067.

Ruthenburg, A.J., Allis, C.D., and Wysocka, J. (2007a). Mol. Cell *25*, 15–30.

Ruthenburg, A.J., Li, H., Patel, D.J., and Allis, C.D. (2007b). Nat. Rev. Mol. Cell Biol. *12*, 983–994.

Slany, R. (2009). Haematologica *94*, 984–993.

Wang, Y., Han, R., Zhang, W., Yuan, Y., Zhang, X., Long, Y., and Mi, H. (2008). FEBS Lett. *582*, 835–839.

Wang, Z., Song, J., Milne, T.A., Wang, G.G., Haitao, L., Allis, C.D., and Patel, D.J. (2010). Cell, this issue.

Xia, Z.B., Anderson, M., Diaz, M.O., and Zeleznik-Le, N.J. (2003). Proc. Natl. Acad. Sci. USA *100*, 8342–8347.

Retrotransposition and Structural Variation in the Human Genome

James R. Lupski[1,*]

[1]Departments of Molecular and Human Genetics and Pediatrics, Baylor College of Medicine and Texas Children's Hospital, Houston, TX 77030, USA

*Correspondence: jlupski@bcm.edu

DOI 10.1016/j.cell.2010.06.014

New assays are revealing that the diploid human genome contains extensive amounts of structural variation. Genome-wide approaches described in three papers in this issue (Beck et al., 2010; Huang et al., 2010; Iskow et al., 2010) paint a dynamic portrait of our genome, revealing a prominent role for repetitive sequences in shaping its structural variation.

The amount of variation in the human genome is tremendous! Such variation consists primarily of single nucleotide variations (SNVs) and copy number variations (CNVs). CNVs are extensive in the human genome, but their importance has only recently been appreciated with genome-wide technologies that can assay for CNVs and other structural changes such as inversions that deviate from the "normal" diploid state. Personal genome sequences have revealed that each one of us differs from the reference haploid human genome by about 3.0–3.5 $\times$ 10^6 single nucleotide variants (Lupski et al., 2010). Regarding the 6 $\times$ 10^9 base pairs (bp) of the diploid human genome, recent studies reveal that each of our genomes has an average of ~1000 CNVs ranging from about 500 bp to 1.3 Mb in size (median is 2.9 kb) (Conrad et al., 2010). The mechanisms underlying the genomic rearrangements that result in CNVs are diverse (Hastings et al., 2009). This issue of *Cell* reports three studies from different groups, which independently demonstrate that highly repetitive sequences in the human genome such as the LINE-1 and *Alu* elements contribute significantly to structural variations (Beck et al., 2010; Huang et al., 2010; Iskow et al., 2010) (Figure 1). Furthermore, they show that many endogenous LINE-1 sequences undergo active transposition in both germline and somatic cells to a much greater degree than previously thought. Complementing this trio of papers is a recent study in *Genome Research* that further substantiates the contention that repetitive elements contribute to structural variation (Ewing and Kazazian, 2010).

Huang and colleagues perform transposon insertion profiling by microarray (TIP-chip) to make a genome-wide map of human L1(Ta) retrotransposons, a younger class of LINE-1 repetitive elements. They identified numerous new human L1(Ta) insertional polymorphisms. Their data suggest that the occurrence of new insertions is twice as high as previously estimated; they calculate one insertion in every 108 births. Interestingly, they explored L1(Ta)s in 69 unrelated male probands with X-linked intellectual disability and identify an intronic insertion in the *DACH2* gene, the ortholog of the *dachshund* gene that regulates neuronal differentiation in fruit flies. Mammalian *Dach2* is highly expressed in fetal brain relative to other tissues, and mapping studies have implicated it as a potential locus for intellectual disability. Obviously more work needs to be done to prove causation, but it is important to note that no current diagnostic assays are able to evaluate either de novo repetitive sequence insertions in introns or CNVs of repetitive sequences (i.e., dimorphic polymorphism) given that interrogating oligonucleotides are unique sequence probes. Recent mutation surveys of >200 unrelated males from families with X-linked intellectual disability that screened for either exonic coding region sequence variations (SNVs) or CNVs identified apparent causative alleles in less than half of the families; clearly

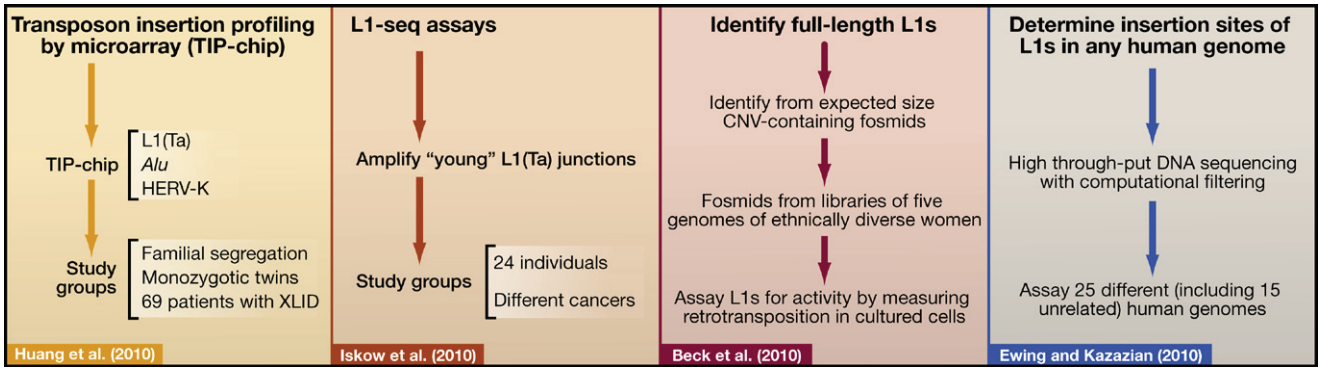

Figure 1. Identifying Repetitive Sequences and Structural Variations in the Human Genome

Four different studies use diverse genome-wide assays of personal genomes with or without next-generation massively parallel DNA sequencing to identify repetitive sequences and structural variations in the human genome (Beck et al., 2010; Huang et al., 2010; Iskow et al., 2010; Ewing and Kazazian, 2010). The L1(Ta) repeat sequences represent relatively young retrotransposition events of LINE elements. *Alu* is the most frequent repetitive sequence class in the human genome, originally identified more than 30 years ago by reassociation techniques. All four studies show that LINE and Alu elements contribute to structural variations in the human genome. XLID, X-linked intellectual disability.

other mutational mechanisms must be involved. An important aspect of the TIP-chip approach is that it can be generalized to other repetitive sequences such as HERV-K endogenous retroviral elements and even *Alu* elements. It will be interesting to see how many dimorphisms relate to a gain versus a loss of an *Alu* element (Edwards and Gibbs, 1992).

In their study, Iskow and colleagues combine element/locus junction-specific PCR with next-generation 454 DNA sequencing in 76 individuals to demonstrate that *Alu* and L1 insertions are abundant in human populations. Furthermore, they document that de novo somatic L1 insertions occur at high frequencies in human lung cancer genomes. Genome-wide analyses of methylation status suggest that altered DNA methylation is associated with and may be responsible for the high levels of L1 mobilization observed in these tumors. Their data suggest that transposon-mediated mutagenesis is extensive in human genomes of both germline and somatic cells and that "natural mutagenesis" of human genomes can occur by endogenous transposons. This might be particularly interesting to think of in the context of mouse phenotypes and the mutant strains in which the identified mutations involve retrotransposons.

Iskow et al. document insertions into human genes, and as might be anticipated from target size, intronic insertions outnumber exonic insertions. Many of the repetitive sequence-related dimorphisms have minor allele frequencies (MAFs) of <5%, suggesting more recent events. This is in contrast to a database of human retrotransposon insertion polymorphisms (dbRIP; http://dbrip.brocku.ca/), where the average MAF is 18%. In fact, 9 out of 47 (19.1%) of the rare insertions identified were found in only a single human (MAF < 1.1%). Consistent with their idea of a great deal of structural variation related to the activity of repetitive elements, these authors point to the results of personal genome sequences, which show a peak frequency for structural variations at around the 300 bp size correlating with the size of *Alu*.

Meanwhile, in the third study of the trio, Beck and colleagues describe the prevalence of active L1s in the human population. The authors used a fosmid-based, paired-end DNA sequencing strategy to identify 68 "dimorphic" full-length L1s, from five individuals from different parts of the world, that are absent from the haploid reference human genome sequence. Incredibly, more than half of these L1s (37 of 68) are highly active in an experimental assay, parenthetically pioneered by the Moran laboratory, that directly measures transposition by a cultured cell retrotransposition method. Their studies empirically document that "hot" L1s are much more abundant in the human population than previously appreciated. Furthermore, ongoing L1 retrotranspositions may be a major source of interindividual genetic structural variation as well as a means of shaping the human genome during evolution.

Finally, in their study, Ewing and Kazazian (2010) use high-throughput DNA sequencing to determine insertion sites for virtually all members of the human-specific L1 retrotransposon family in 25 individuals. They find that any two individuals differ on average at 285 sites with respect to L1 insertion presence or absence, that is, structural dimorphism. Ewing and Kazazian estimate that the rate of L1 retrotransposition in humans is between 1/95 and 1/270 births. They further propose that the number of dimorphic L1 elements in the human population with gene frequencies >0.05 is between 3,000 and 10,000; dbRIP currently contains 577 in total!

These four studies each demonstrate that personal genomes differ tremendously in the specific locus position of individual repetitive sequence elements and that both germline and somatic de novo transposition events may be occurring much more frequently than thought. This likely reflects an endogenous property of genome structure, as the strain specificity of repetitive sequence-based PCR fingerprinting (so-called "rep-PCR") in bacteria reflects the different positioning of repetitive elements within bacterial genomes of individual (i.e., personal) strains (Versalovic et al., 1991). We are now faced with multiple lines of experimental evidence that all show that repetitive sequences figure prominently

in the structural variations of the human genome. This evidence includes the sheer number of repetitive elements, the de novo transposition/insertion/mutation rates, the fact that similar elements (e.g., *Alu-Alu*, LINE-LINE) can act as substrates for homologous recombination and can generate CNVs by nonallelic homologous recombination (NAHR), and that such elements can provide microhomology for priming polymerase extension during template switching in replication-based mechanisms (Zhang et al., 2009).

The junk DNA has come out of the closet or garage (depending on your favorite euphemism) and begun to reveal its immense value to genome evolution and human biology. It has certainly been fascinating to witness the transformation of human repetitive sequences from a nuisance that needs to be quenched during hybridizations to a central player in structural variation, arguably the most common form of genetic variation in humans and one that figures prominently in evolution and disease.

REFERENCES

Beck, C.R., Collier, P., Macfarlane, C., Malig, M., Kidd, J.M., Eichler, E.E., Badge, R.M., and Moran, J.V. (2010). Cell, this issue.

Conrad, D.F., Pinto, D., Redon, R., Feuk, L., Gokcumen, O., Zhang, Y., Aerts, J., Andrews, T.D., Barnes, C., Campbell, P., et al. (2010). Nature *464*, 704–712.

Edwards, M.C., and Gibbs, R.A. (1992). Genomics *14*, 590–597.

Ewing, A.D., and Kazazian, H.H. (2010). Genome Res. Published online May 20, 2010. 10.1101/gr.106419.110.

Hastings, P.J., Lupski, J.R., Rosenberg, S.M., and Ira, G. (2009). Nat. Rev. Genet. *10*, 551–564.

Huang, C.R.L., Schneider, A.M., Lu, Y., Niranjan, T., Shen, P., Robinson, M.A., Steranka, J.P., Valle, D., Civin, C.I., Wang, T., et al. (2010). Cell, this issue.

Iskow, R.C., McCabe, M.T., Mills, R.E., Torene, S., Pittard, W.S., Neuwald, A.F., Van Meir, E.G., Vertino, P.M., and Devine, S.E. (2010). Cell, this issue.

Lupski, J.R., Reid, J.G., Gonzaga-Jauregui, C., Rio Deiros, D., Chen, D.C., Nazareth, L., Bainbridge, M., Dinh, H., Jing, C., Wheeler, D.A., et al. (2010). N. Engl. J. Med. *362*, 1181–1191.

Versalovic, J., Koeuth, T., and Lupski, J.R. (1991). Nucleic Acids Res. *19*, 6823–6831.

Zhang, F., Carvalho, C.M., and Lupski, J.R. (2009). Trends Genet. *25*, 298–307.

Alzheimer's Disease Neurons Fail the Acid Test

Wim Annaert[1,*] and Bart De Strooper[1,*]

[1]Center for Human Genetics, KULeuven, and VIB Department for Molecular and Developmental Genetics, Leuven B-3000, Belgium

*Correspondence: wim.annaert@cme.vib-kuleuven.be (W.A.), bart.destrooper@cme.vib-kuleuven.be (B.D.S.)

DOI 10.1016/j.cell.2010.06.009

Mutations in the *presenilin* genes are the most common cause of familial forms of Alzheimer's disease. Although it is well known for its role in the generation of amyloid peptide, Lee et al. (2010) now report that presenilin 1 deficiency also impacts maturation of the lysosomal proton pump, affecting autophagocytosis and protein turnover.

Although amyloid plaques and neurofibrillary tangles are the classic hallmarks of Alzheimer's disease (AD) pathology, AD is also characterized by endosomal/lysosomal abnormalities including the aberrant accumulation of subcellular structures involved in autophagy (the degradation of long-lived organelles and macromolecules), such as autophagosomes, autolysosomes, and lysosomal dense bodies. Multiple lines of evidence now suggest a potential link between AD and autophagy. These include reports that *presenilin 1*, the gene most frequently mutated in familial forms of AD, has a role in autophagocytosis or may have a more general function in subcellular membrane trafficking (Esselens et al., 2004; Wilson et al., 2004; Sannerud and Annaert, 2009). In addition, disruption of the autophagy pathway promotes the deposition of amyloid plaques and accelerates neuronal loss (Pickford et al., 2008; Tooze and Schiavo, 2008).

In findings presented in this issue, Lee et al. (2010) strengthen this emerging connection between AD and autophagy and propose a mechanism that contributes to autophagosome accumulation observed in the disease. In normal cells, autophagosomes fuse with acidified lysosomes to promote the degradation of the autophagosome's contents. Lee et al. now report that disease-causing mutations in presenilin 1 impair the acidification of lysosomes, thereby interfering with the subsequent autophagosome clearance (Figure 1). Given that one of the signatures of neurodegenerative disorders is accumulation of aggregates of misfolded protein, decreased turnover of protein (as observed in the current experiments) might contribute to the disease process. These provocative and potentially controversial observations come amidst growing interest in molecular pathways that act in parallel (or possibly intersect) with the classical amyloid cascade.