Cell Cycle 12:1, 5–6; January 1, 2013; © 2013 Landes Bioscience

The variable somatic genome

Maeve O'Huallachain,[1,2] Sherman M. Weissman[1,3,*] and Michael P. Snyder[1,*]

[1]Department of Genetics; Stanford University School of Medicine; Stanford, CA USA; [2]Department of Molecular, Cellular and Developmental Biology; Yale University; New Haven, CT USA; [3]Department of Genetics; Yale University; New Haven, CT USA

Many genetic variations in human populations are known, but genetic variability within healthy individuals is less familiar. Somatic mosaicism is the occurrence of genetically distinct cells within the same organism or tissue. Somatic mosaicism is well-known to occur in healthy cells of the immune system, and it has been frequently identified in disease, particularly cancer. A multi-step theory of cancer proposed in the 1970s[1] and continually developed[2] posits that genetic instability in precancerous cells leads to an accumulation of mutations that eventually give rise to tumorigenesis. Our results reveal new characteristics of intra-individual genetic instability.[3]

It remains widely assumed that healthy cells that arise from the same zygote contain identical genomes. This assumption underlies the research and diagnostic practice of using easily accessible blood or saliva DNA to perform genetic tests. A growing body of evidence counters that somatic genome rearrangements are relatively common. Age-related copy number variations (CNVs) in human blood cells were identified using SNP arrays.[4] Retrotransposition[5] was discovered in the human brain, although the extent of retrotransposition may have initially been overestimated.[6] CNVs in DNA from somatic tissues[7] were found using aCGH with low-resolution BAC arrays. BAC arrays suffer from reproducibility concerns and can detect variations only above ~50 kb. Structural variations starting at ~1 kb are thought to account for most of the variation in human genomes. Consequently, a great deal of somatic genomic variation likely remains understudied.

We analyzed DNA from diverse tissues from six unrelated individuals for CNVs using high-resolution aCGH.[3] The tissues were obtained during routine autopsy of subjects without known hereditary disorders or cancer. We identified 178 tissue-specific CNVs across 32 tissues; 73 were validated by a secondary method. The CNVs ranged from 2–184 kb, with the majority of events below 50 kb. Seventy-nine percent of the validated CNVs intersected genes. The mechanisms of somatic CNV formation cannot be definitively determined by aCGH. However, our breakpoint analyses revealed that somatic CNV breakpoints are enriched near microsatellite repeats.

Commonly occurring somatic CNVs have implications for many aspects of biology. Seven somatic CNVs were identified in the same genomic region in more than one individual. We suspect that these locations are hotspots for somatic rearrangement. Liver, small intestine and pancreas were among the tissues with potential hotspot events. These tissues are derived from the endoderm germ layer. The hotspots perhaps play a role in the differentiation process. These same tissues also exhibited the most somatic CNVs overall. Compared with other tissues, such as brain, that displayed fewer somatic CNVs, these endoderm-derived tissues are known to experience greater cell proliferation and turnover. Conceivably, cell types with higher rates of division have more occasions to undergo genomic rearrangements. The subjects analyzed in our study were of middle to advanced age, and their dividing tissues are expected to have undergone many cell divisions. A comparison of DNA from somatic tissues of younger subjects with that from older subjects could reveal a relationship between a subject's age and the degree of somatic variation. However, somatic variation may occur early in development. This would explain the occurrence of the CNV in a substantial fraction of the cells in a tissue sample.

The aCGH signals suggest that there are not only genomic differences between tissues, but also within tissues. We attribute this, in part, to a portion of the tissues containing nonparenchymal cells from blood vessels and connective tissue that do not contain the particular CNVs. However, the signals indicate that the CNVs are heterogeneous throughout the parenchymal cells. Somatic CNVs potentially occur in many or all cells but are below the level of detection of aCGH. This hypothesis is supported by a recent study[8] showing that many of the CNVs detected in iPS cells derived from human skin fibroblasts were also present at a low level in the primary fibroblast cultures. We speculate that highly deleterious CNVs make cells unviable and are never detected, while cells with neutral or advantageous CNVs persist and divide into clonal populations that can be detected. As technologies to analyze single cells improve, somatic genomic variation may be detected at the single-cell level.

Our results indicate that somatic variations should be considered in biological sampling practices. "Normal" tissues from distant lineages, which are used for comparison, may identify somatic events that are not relevant to the disease state. Clinicians using blood and saliva DNA for diagnosis of genetic disorders may miss relevant somatic mutations in other tissues.

Further analyses are necessary to determine the contribution of these somatic variations to biological variability and disease predisposition. Some of the somatic CNVs contain genes that have been shown in the literature to be expressed in the tissues we tested. Somatic CNVs that intersect tumor suppressors and/or

*Correspondence to: Sherman M. Weissman and Michael P. Snyder; Email: sherman.weissman@yale.edu and mpsnyder@stanford.edu
Submitted: 11/09/12; Accepted: 11/11/12
http://dx.doi.org/10.4161/cc.23069
Comment on: O'Huallachain M, et al. Proc Natl Acad Sci USA 2012; 109:18018-23; PMID:23043118; http://dx.doi.org/10.1073/pnas.1213736109

oncogenes may directly induce tumorigenesis. Alternatively, somatic cells may acquire CNVs over many cell divisions, and the somatic CNV load creates genome instability or predisposition to malignant changes. Perhaps certain stages of development have greater protection from somatic variation or somatic variation provides variability that confers an advantage under certain environmental conditions. It remains to be seen how wide-reaching the implications of somatic genomic variation are on human biology.

References

1. Nowell PC. Science 1976; 194:23-8; PMID:959840; http://dx.doi.org/10.1126/science.959840
2. Croce CM. N Engl J Med 2008; 358:502-11; PMID:18234754; http://dx.doi.org/10.1056/NEJMra072367
3. O'Huallachain M, et al. Proc Natl Acad Sci USA 2012; 109:18018-23; PMID:23043118; http://dx.doi.org/10.1073/pnas.1213736109
4. Forsberg LA, et al. Am J Hum Genet 2012; 90:217-28; PMID:22305530; http://dx.doi.org/10.1016/j.ajhg.2011.12.009
5. Baillie JK, et al. Nature 2011; 479:534-7; PMID:22037309; http://dx.doi.org/10.1038/nature10531
6. Evrony GD, et al. Cell 2012; 151:483-96; PMID:23101622; http://dx.doi.org/10.1016/j.cell.2012.09.035
7. Piotrowski A, et al. Hum Mutat 2008; 29:1118-24; PMID:18570184; http://dx.doi.org/10.1002/humu.20815
8. Abyzov A, et al. Nature 2012; In press; PMID:23160490; http://dx.doi.org/10.1038/nature11629