Exploring the Non-coding Genome with Chromosomal Structural Rearrangements

Cynthia C. Morton, Ph.D.
Mini-Symposium on Precision Medicine
CityU of Hong Kong
April 16, 2024

Since recognition of the correct number of chromosomes in humans in 1956, chromosomal abnormalities have been foundational in mapping the human genome. Various technologies including banded chromosomes, FISH and microarrays have been invaluable in increasing the resolution of structural rearrangements followed by nucleotide resolution in the early 2000's. The Developmental Genome Anatomy Project (DGAP) has championed nucleotide precision analyses of apparently balanced chromosomal rearrangements in the setting of clinical phenotypes and accelerated dramatically with next-generation sequencing. Initially focused on gene disruptions, it became increasingly clear that 3D organization of chromatin underlies position effects on protein-coding genes. Despite impressive successes in identifying genetic etiologies in protein-coding genes in individuals with rare diseases, the lack of diagnoses in many still begs the question of whether additional etiologies may be detected increasingly in the noncoding genome. Evolving knowledge of the human genome reveals elements of the noncoding genome such as micro RNAs and long non-coding RNAs (lncRNAs) that will doubtless contribute to a more precise understanding of gene expression in human biology. Chromosomal structural rearrangements can also provide signposts for their investigation.

A mother-daughter pair (DGAP353) with a 46,XX,t(14;17)(q24.3;q23) karyotype were referred to DGAP with nonsyndromic mild-to-moderate sensorineural hearing loss. The translocation was detected by amniocentesis following a positive maternal serum screen with an increased risk for trisomy 21 in the female fetus and was subsequently determined to be maternally inherited. Other than deafness, the mother and daughter are clinically normal and had a normal exome analysis. A next-generation cytogenetic nucleotide level research nomenclature made possible description of the aberration in a single line consistent with conventional cytogenetic nomenclature (maternally inherited by the daughter): 46,XX,t(14;17)(q24.3;q23)mat.seq[GRCh38] t(14;17)(14pter→14q23.3(+)(65,855,3{58-60})::17q23.2(+)(61,393,84{1-3})→17qter;17pter→17q23.2(+)(61,393,812)::TATATACC::14q23.3(+)(65,855,359)→14qter)mat.

Interpretation of the t(14;17) was initially focused on *TBX2* at 17q23.2 as a CNV encompassing a genomic region including *TBX2* had been reported in an individual with hearing loss among other clinical findings. Subsequent genome annotation in DGAP353 revealed disruption of *TBX2-AS1* at 17q23.2 on the negative strand 290 base pairs upstream of *TBX2* on the positive strand. *TBX2* was recently reported in a KO mouse model as a master regulator of hair cell fate. Upon ablation of *TBX2*, inner hair cells are replaced with outer hair cells. *TBX2-AS1* disruption with potential loss of function predicted from the translocated allele may result in dysregulated *TBX2* expression. DGAP353 became a potential model for identification of lncRNAs as etiologic in n-of-one clinical disorders and has led to an investigation of additional partially solved and unsolved DGAP cases that might be attributed to genomic disruptions involving lncRNAs. Additional DGAP cases will be presented substantiating this hypothesis.

lncRNAs may be significant contributions to the etiology of rare diseases and have yet to be included in gene panels or exome analyses. Single nucleotide variants in lncRNAs for interpretation of pathogenic changes is challenging given that lncRNAs do not encode proteins and are known to be highly tissue specific. Reassessment of chromosomal structural rearrangements may provide evolving insight into genetic etiologies of clinical disorders as the human and mouse genomes become increasingly annotated for lncRNAs.