

Genes on or off! How regulatory DNA variations disrupt the balance of P53 and cMYC binding

Identification of causative DNA variants in common complex diseases is very important to screen individuals with high-risk for diseases and for developing therapies to target the genetic cause of diseases, in particular cancer. DNA variations that lie outside the protein-coding regions can alter gene expression without affecting its protein function. This research area is critical in the context of disease susceptibility and severity. Basic and translational research has been previously directed towards DNA variations located within sequences that code for protein products due to their obvious effect on the function of the corresponding protein. However, recent genome-wide association studies demonstrate that the majority of DNA variants associated with cancer and other common complex diseases are located in outside the protein coding-regions, while few causative regulatory variants have been identified so far. The specific function of sequences that lie outside the protein-coding regions is not known for the majority of the loci compared to those within the coding regions, which makes prediction of the effect of non-coding variants challenging.

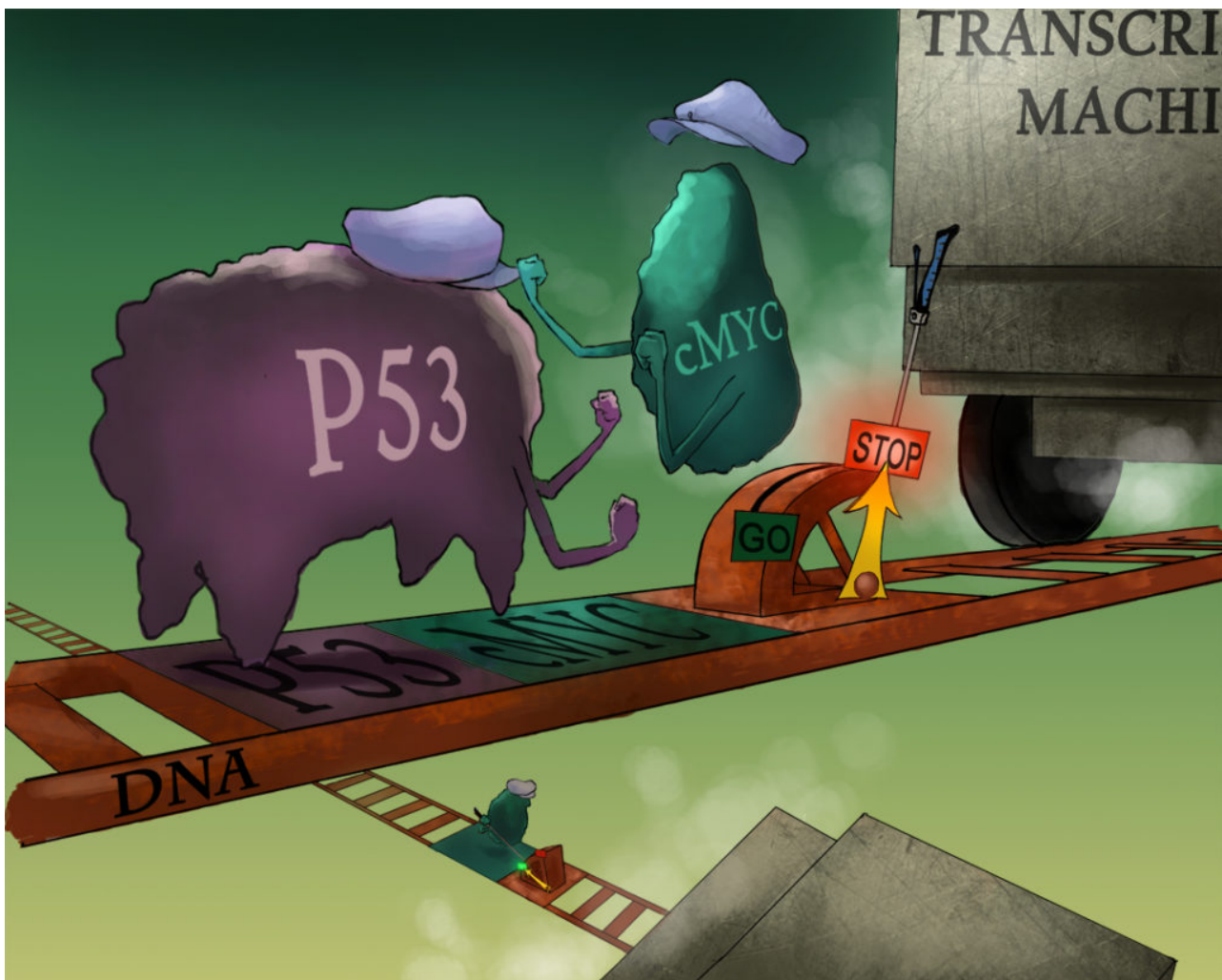


Fig. 1. An illustrative model of the P53 and cMYC competitive binding at overlapping binding sites in regulatory regions, which affects target gene expression through interaction with the transcriptional machinery (Transcri Machi).

To understand the underlying mechanism by which non-coding DNA variations alter the expression pattern of target genes, it is very important to identify the molecular regulatory grammar that dictates the readout by the transcriptional machinery. In our study, we identified a novel mechanism by which P53 and cMYC, two well-studied transcription factors that play a major role in cancer development, compete for binding at the same DNA site to regulate expression of target genes. Genome-wide analysis of the binding profiles for P53 and cMYC in human cancer and mouse embryonic cells mapped a significant number of putative regulatory elements with occupancy signals for both factors. Bioinformatic analysis of the co-occupied regions by P53 and cMYC identified on average two cis-overlapping motifs (CisOMs). At these CisOMs, where both P53 and cMYC may bind, we hypothesized that DNA variation of a single nucleotide can favor binding of one factor over the other and therefore affect gene expression. Furthermore, within every two CisOM identified, one contains a single nucleotide polymorphism (SNP). Gene ontology of predicted target genes for CisOMs showed that the majority are involved in cellular functions that control cell division, differentiation and death. Molecular evidence of enhancer activity at six co-occupied regions and P53 and cMYC binding to a CisOM within one of the tested co-occupied regions was verified via biochemical assays. These results suggest an important molecular mechanism occurring at non-coding regions which can be used to identify causative regulatory SNPs and variants associated with common complex diseases.

Katherine Kin and Walid D. Fakhouri

*Department of Diagnostic and Biomedical Science, Center for Craniofacial Research,
University of Texas Health Science Center at Houston School of Dentistry, Houston, TX, USA*

Publication

[The effect of non-coding DNA variations on P53 and cMYC competitive inhibition at cis-overlapping motifs.](#)

Kin K, Chen X, Gonzalez-Garay M, Fakhouri WD
Hum Mol Genet. 2016 Apr 15