C3 / Padres Pedal the Cause 2017

"Decoding Colon Cancers Using Boolean Principles"

Principal Investigators:
Pradipta Ghosh, MD (Moores Cancer Center)
Debashis Sahoo, PhD (Moores Cancer Center)
Manuel Perucho, PhD (Sanford Burnham Prebys)

SCIENTIFIC ABSTRACT
Hierarchical organization is common to all organs and tissues where diverse cell types coordinately perform their functions. Disruption of such hierarchical organization is often the first step in neoplastic transformation, and during oncogenic progression. Diversity of cell types within the normal hierarchical organization is characterized by a gene expression signature; this 'normal' signature is fundamentally distinct from the disrupted hierarchical organization that is encountered in pre- neoplastic states and in cancers. Gene expression relationship is fundamental to our understanding of biological systems at all levels, and drives most, if not all, techniques for detecting, diagnosing, and treating a disease. Among many such relationships, Boolean principle is one criteria that has tremendous potential to unravel the underlying complexity of normal tissues and cancers. Published work using Boolean logic has successfully helped chart the changing landscape of gene expression signature during B cell differentiation, and during the initiation and progression of bladder and colon cancers. The complete potential of this approach, however, remains unrealized. This application proposes to use the precision of mathematics [Boolean logic] to analyze noisy "big-data" [RNA Seq datasets from patient-derived samples], to unravel new mechanisms/biomarkers/biology in the pathogenesis of colorectal cancers (CRCs). The first aim is to build a virtual colon crypt, i.e., map the transcriptome and the proteome of a colon crypt in an unbiased manner by applying Boolean logic/network and single-cell gene expression dataset from normal and CRCs. The second aim will visualize the altered organization of crypts in colon cancers and in preneoplastic lesions of the colon to assess the changes in the crypt organization during "sequential" strikes/hits" that propel conversion from normal to cancers. Finally, the third aim will assess the risk of neoplasia in pre-neoplastic lesions of the colon by comparing and contrasting cancer progression in adenomatous and serrated polyps. Insights gained will not only unlock hitherto unforeseen gene expression profiles in normal colon crypts and help generate detailed maps of gene expression changes due to mutations that drive normal-to-adenoma-to-cancer progression, but also unravel new biomarkers for prognostication in cancers and pre-neoplastic lesions (polyps).

LAY ABSTRACT
Colorectal cancer is the 3rd most common cancer in the US. Despite a large amount of work that has concentrated on understanding colon tumor formation, we still do not know the full complement of molecular lesions that are individually necessary (and together sufficient) to cause colorectal cancer. Neither do we understand why some specific mutations that are relatively rare in other tumors are extremely common in colorectal cancer. As a result, the race toward personalized and precision-based management plan in colorectal cancer has seen a lost decade with little to nothing that has translated to the cancer clinics after the discovery of the impact of tumor microsatellite instability status in the mid-90's. Big data (gene expression datasets), meanwhile, has accumulated from multiple studies, but has had a '*streetlight effect*': where and how we look affects what we see through our own biases. Thus, a lot of what has been accomplished using such big data is justification/ rationalization of research through the same bias-filled traditional reductionist approach [one protein, pathway at a time based on where one wanted to look]. Which can explain why most pre-clinically identified therapeutic target or

biomarkers identified at the bench stumble during translation into clinics. Currently, there is no effective way to extract meaningful information from such noisy collection of Big data in an unbiased way.

For this “DECODING COLON CANCER” project, we propose to use the tools of BOOLEAN PRINCIPLES AND ALGORITHMS to develop a quantitative and comprehensive model of colorectal tumor formation. It will use the precision of mathematical logic to sort through noisy RNASeq ‘big data’ from patients to first build a virtual colon crypt, and then validate it using human normal colon tissues. Once such a virtual crypt is built, we use that model to visualize changes in the expression of key genes at the top and bottom of the crypt in preneoplastic lesions [adenomas] and carcinomas to find out what changes, when, and what impact can such change have on cellular phenotype and on the organizational architecture of the colon crypt. The advantage of such an approach is that it removes the ‘street light’ effect of a reductionist approach and instead focuses on the strongest signals/clues that are usually representatives of invariants/constants in biology. The model will be analyzed for functional relationships between genes and relationships between cellular pathways to gain insight into how each component may contribute to tumor formation, and explain in detail how the genetic disposition of an individual can activate expression of genes that drive uncontrolled cell growth and lead to cancer. Eventually, this model will be used to find novel therapeutic targets, to guide genetic screening to identify individuals with elevated risk for developing colorectal cancer and to classify patients into sub-groups to select the treatment combination that is optimal for each patient (personalized medicine).

Overall, our study, led jointly by three PIs and a multidisciplinary team, each with non-overlapping expertise holds great promise in the synergy of math, biology, and translational medicine to improve our understanding of the pathogenesis of human colon cancers, and have limitless potential for delivering diagnostic and prognostic biomarkers and therapeutic targets to combat this disease.