12/5/2023 Update:

**1.a** is complete, **b.** complete, **c.** working on getting CIRCUST to work with our dataset and then implementing it into our docker container.

Working with Deepak biweekly to make sure we are hitting his targets. We have reached out to Bose a couple of times but no response, we should reach out again in Capstone 2. **Our Notes for the project:  
1. Timestamping TCGA Data:**

The initial phase is dedicated to the meticulous process of timestamping genomic data from The Cancer Genome Atlas (TCGA).

a. **Data Retrieval**:

* Automatic scripts will be developed to expedite the download of a substantial number of datasets from TCGA, specifically targeting between 5,000 to 10,000 datasets related to PDAC.
* The aim is to amass a rich dataset amalgam to enable a comprehensive analysis.

1. **Data Pre-processing**:

* Transcripts Per Million (TPM) values will be computed for normalization purposes, which is a precursor to the timestamping and subsequent analysis.

c. **Timestamping Methodology**:

* The absence of timestamping data necessitates the application of the CIRCUST methodology, as detailed in a recent publication, to generate a temporal context for each sample.
* The process involves shortlisting a set of 15 genes from a pool of 20,000 genes, known for their rhythmic behavior across various cancers.
* By clustering these genes, a two-dimensional plot is created, which is then converted into radian values representing a time dimension.
* Essentially, this step orders the samples from 0-24 hours, marking a critical step towards understanding the temporal dynamics of gene expression in PDAC.

d. **Identification of rhythmic genes:**

* By ordering the time stamped TCGA samples, rhythmic genes will be identified for pancreatic cancer. We will use multiple methods to deduce gene rhythmicity such as meta2D, RAIN, Limorrhyde. In addition, we will cluster the TCGA samples based on molecular features and subtypes to assess which genes are differentially rhythmic in each sub-group. This would aid in generating the first ever repo of rhythmic gene signature in pancreatic cancer.
* *[Time permitting, we plan to use the similar strategy for colorectal cancer as well.]*

d. **Infrastructure**:

* RUSH, equipped with robust computational resources including AWS, Azure, and a supercomputer, will facilitate handling and processing terabytes of data.

**2. Validation and Laboratory Experiments:**

The consequent phase hinges on validating the timestamped data and diving into laboratory experiments to explore the rhythmic genes and their interaction with drugs.

a. **Cell Line Dataset Validation**:

* Well-defined PDAC cell models will be utilized to validate the timestamped TCGA data.
* This step is crucial to confirm the rhythmic nature of identified genes in a controlled environment.

b. **Drug Interaction Analysis**:

* An extensive list of drugs interacting with the rhythmic genes will be compiled using resources like DrugBank.
* A computational analysis will be conducted to shortlist drugs based on their binding affinity to the proteins encoded by the rhythmic genes.
* The focus will be on FDA approved drugs to expedite potential clinical applications.

c. **In vitro Drug Testing**:

* Shortlisted drugs will be tested in the lab using organoid models grown from patient-derived cancer cell samples.
* This step aims to ascertain the therapeutic potential of these drugs against PDAC cells.

**3. Xenograft Experiments and Future Directions:**

The final phase extends to in vivo testing and delineates the roadmap for future explorations.

a. **Xenograft Mammalian Experiments**:

* Post in vitro validation, xenograft experiments in mice will be initiated to further evaluate the efficacy and safety of the shortlisted drugs.
* This step edges closer to potential clinical trials, marking a significant milestone towards real-world applications.

b. **Future Explorations**:

* The pipeline developed through this project could potentially be extended to other cancer types.
* Additionally, the project holds promise for spawning publications, research grants, and novel software tools, collectively propelling the domain of chronotherapeutic strategies forward.