Fr. Conceicao Rodrigues College of Engineering

Father Agnel Ashram, Bandstand, Bandra –west, Mumbai-50

Department of Computer Engineering

PROJECT PROPOSAL TEMPLATE

Group Information:

Roll No	Student Name	Div A/ B
9913 (Group Leader)	Mark Lopes	A
9914	Vivian Ludrick	A
9909	Pratyay Koley	A
9882	Aryan Daripkar	A

TITLE OF THE PROJECT:
Cross-Species Gene Mapping for Cardiac Regeneration Using Zebrafish and Human Orthologs

BROAD AREA (E.g. IoT, Machine Learning, Computer Vision, System Security etc.):
Machine Learning

Category of the Project (Product based, Application based, or Research based):
Application Based

ABSTRACT (maximum 300 words):

Heart disease remains a leading cause of mortality globally, with limited options for natural tissue regeneration in humans. In contrast, zebrafish exhibit remarkable cardiac regeneration capabilities due to the activation of specific genes post-injury. This project aims to develop an AI-powered bioinformatics platform that maps zebrafish heart regeneration genes to their human orthologs, identifies potential therapeutic targets, and ranks them based on relevance using machine learning techniques.

Starting with a curated list of regenerative zebrafish genes, our system uses public biological databases like Ensembl and ZFIN to find their human counterparts. Then, machine learning models are applied to biomedical literature and gene annotations to score and classify human genes based on their potential regenerative functions. The system also integrates keyword analysis and disease association frequency from sources like PubMed and DisGeNET.

An interactive web-based dashboard will allow users—especially researchers and students—to explore mapped genes, their functions, expression levels, and disease relevance. This combination of AI, biology, and data visualization provides a novel, interdisciplinary approach to assist in regenerative medicine research.

By democratizing access to complex genomic data and guiding researchers towards promising therapeutic genes, this project contributes to the advancement of bioinformatics tools and translational healthcare research.

Fr. Conceicao Rodrigues College of Engineering

Father Agnel Ashram, Bandstand, Bandra –west, Mumbai-50

Department of Computer Engineering

MOTIVATION (maximum 100 word):

Despite advances in medicine, humans lack the ability to naturally regenerate heart tissue after injury. Zebrafish, however, possess this capability. By understanding and translating their genetic mechanisms into human context, we can assist in identifying new therapeutic targets. Our motivation is to use machine learning and bioinformatics to bridge this cross-species biological gap and build a tool that supports researchers working in regenerative medicine.

PROBLEM STATEMENT:

How can we leverage zebrafish heart regeneration genes to identify and rank their human equivalents using machine learning, to aid biomedical researchers in discovering new therapeutic targets for cardiac repair?

SDGs Mapped:

SDG 3: Good Health and Well-being
SDG 9: Industry, Innovation, and Infrastructure
SDG 4: Quality Education

OBJECTIVES:

- Identify key zebrafish genes involved in cardiac regeneration.
- Map these genes to human orthologs using bioinformatics tools.
- Use machine learning to rank and prioritize human genes for regenerative potential.
- Design an interactive web interface to explore mapped genes, associated functions, and disease relevance.

METHODOLOGY (TENTATIVE IF ANY):

Data Collection: Gather known regenerative zebrafish genes from research papers and databases.

Gene Mapping: Use Ensembl BioMart and APIs to map zebrafish genes to human orthologs.

Literature Analysis: Apply NLP on PubMed abstracts to evaluate relevance of human genes in regeneration context.

ML Model: Train a classifier or scorer (e.g., logistic regression or decision tree) to prioritize human genes.

Visualization: Build a dashboard (e.g., using Streamlit/React) for gene exploration.

Fr. Conceicao Rodrigues College of Engineering

Father Agnel Ashram, Bandstand, Bandra –west, Mumbai-50

Department of Computer Engineering

HARDWARE / SOFTWARE REQUIREMENTS:

Python (Pandas, Scikit-learn, NLTK, BioPython)
Ensembl REST API, ZFIN database
PubMed (via Entrez API)
React/ Next (for UI)

INNOVATIVENESS:

Cross-species gene mapping integrated with machine learning for cardiac regeneration.
Combines AI, bioinformatics, and biomedical text mining.

SOCIETAL RELEVANCE? (e.g. Health, Agriculture, Environment, Smart Solution Etc…)

Health: Targets heart disease, aiding in discovery of future gene-based therapies.
Smart Solution**:** AI-powered research assistant for regenerative biology

****