A Platform for Integrating and Sharing Cancer Stem Cell Data

Irena Parvanova PhD, Kirill Borziak PhD, Jennifer Guarino, Joseph Finkelstein MD, PhD

October 8, 2021
ReMeDy Framework and Aims

- ReMeDy platform is based on Signature Commons framework
- Signature Commons was designed as a part of the BD2K-LINCS DCIC effort
- Common Data Elements (CDE) are uploaded using JSON format and validated against existing ontologies
- ReMeDy is designed to contain regenerative medicine projects data, including stem cell characteristics, study subject baseline and outcome information, and links to omics data files
- ReMeDy aims to facilitate cancer care innovation by providing an integrated source of information for cancer stem cell studies
Visual Representation of the ReMeDy Architecture: Interaction between the Docker Containers

ReMeDy architecture

Introduction

- Cancer stem cells (CSCs) are associated with resistance of tumors to treatments and formation of metastasis in multiple cancers

- There is a lack of CSC data homogeneity, organization, deposition, and visualization

- ReMeDy repository allows for the systematical collection and sharing of data

- The functionality and usability of the platform was tested by uploading 52 multi-modal CDE templates, based on 52 published CSC clinical, pre-clinical, and *in vitro* studies
Methods

- Database architecture and web interface
  - Signature Commons platform
  - BD2K-LINCS DCIC effort
  - API functionality
  - JSON format, ingested using a Python script

- Literature search, abstraction process and data collection
  - 52 published CSC projects
Results: Study Type

- Preclinical: 40%
- In Vitro: 17%
- Clinical: 42%
Results: Location of the Research Project
Results: Cancer Type

- Breast cancer
- Colorectal cancer
- Colon cancer
- Pancreatic cancer
- Blastoma
- Liver cancer
- Head and neck cancer
- Prostate cancer
- Bone cancer
- Melanoma
- Ovarian cancer
Discussion

- Cancer medicine requires the creation of a flexible and agile repository for cancer stem cell data aggregation, storage, visualization, and sharing

- ReMeDy platform and the multi-modular CDE framework for CSCs for 52 publicly available and PubMed indexed projects

- ReMeDy is an organized repository, which captures CSCs research project information in a standardized format, provides effortless visualization and search functions

- ReMeDy promotes accessibility to CSCs projects to facilitate data sharing and collaboration, allowing for standardized cross-discipline and cross-studies comparison
Future Plans and Conclusions

- Future aims for the project include increasing the database size to include all published stem cell trials and develop additional data visualization tools to improve usability.

- Plans for increasing the ReMeDy database size include promoting a crowdsourcing functionality of the platform and developing natural language processing (NLP) functionalities.

- This effort to include establishing and expanding an automated pipeline for uploading the CDE templates.

- Advantages of crowdsourcing are ensuring that the knowledge base for CSCs stays current.


Thank You for Listening!

Questions?

Irena.Parvanova@mountsinai.org