2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

A Platform for Integrating and Sharing Cancer Stem Cell Data*

Irena Parvanova, Kirill Borziak, Jennifer Guarino, Joseph Finkelstein

Abstract— **Advancements in cancer research and treatment have highlighted the need for standardization and sharing of cancer stem cell (CSC) data to facilitate research transparency and to promote collaboration within the scientific community. Although previous applications have attempted to gather and disseminate these data, currently no platform organizes the heterogeneous CSC information into a harmonized project-based framework. The aim of our platform, ReMeDy, is to provide an intelligent informatics solution integrating diverse CSC characteristics, outcomes information, and omics data across clinical, preclinical and in vitro studies. These heterogeneous data streams are organized within a multi-modular framework, subjected to a stringent validation by using standardized ontologies, and stored in a searchable format. To test usefulness of our approach for capturing diverse data related to CSCs, we integrated data from 52 publicly-available CSC projects. We validated the robustness of the platform, by efficiently organizing diverse data elements, and demonstrated its potential for promoting future knowledge discovery driven by aggregation of published data. Next steps include expanding number of uploaded CSC projects and developing additional data visualization tools. The platform is accessible through https://remedy.mssm.edu/.**

Clinical Relevance— **Our platform, ReMeDy, aims to facilitate cancer care innovation by providing an integrated source of information for cancer stem cell studies, including cell characteristics and clinical and pre-clinical outcomes.**

I. Introduction

Cancer stem cells (CSCs) have been associated with resistance of tumors to treatments and formation of metastasis in multiple cancers [1]. Currently, there is a lack of CSCs data standards and frameworks for their characterization. Existing CSC data is not consolidated, stored, or available for access by researchers in a centralized and unified manner. In our previous work, we have discussed the significance of stem cell research, voluminous amount of available stem cell data, and existence of many publicly-available stem cell data bases [2]. Based on the necessity for CSC data harmonization, organization, deposition, and visualization, we created a Regenerative Medicine Data Repository (ReMeDy) platform [3-4], which can be publicly accessed at https://remedy.mssm.edu/. Here, we present the expansion of this resource to include publicly available CSC data. The unique feature of our repository allows for the systematic collection and sharing of data using a multi-modular common data elements (CDE) framework, designed to incorporate clinical and pre-clinical trial outcomes and in vitro findings, CSC characteristics, and omics data for detailed comparisons across studies. This project is aimed at testing the ability of ReMeDy platform of capturing diverse CSC information, by uploading 52 multi-modal CDE templates, based on published CSC clinical, pre-clinical, and in vitro studies, indexed in the PubMed database [6-54].

II. Methods

A. Database architecture and web interface

ReMeDy platform is based on Signature Commons framework [4-5]. Signature Commons was designed as part of the BD2K-LINCS DCIC effort to catalogue biological data. Implemented through Docker, it aims to store and search diverse metadata in a harmonized and flexible manner. ReMeDy was installed on a Linux server using the default installation instructions [4]. The Signature Commons platform is available through GitHub at https://github.com/MaayanLab/signature-commons. It is composed of six repositories: controller, data-api, metadata-api, proxy, schema, and ui.

Validation, visualization, and user interface schema were ingested using the API functionality. Specifically, counting schemas were developed based on the CDE framework to provide counting and filtering functionality to the search results page. The schemas, formatted in JSON, were generated and ingested using a custom Python script. To improve usability of the Application programming interface (API), the upload process was improved by creating an upload interface. The upload interface was developed using ReactJS for the front-end and Spring Boot on the back-end. The interface allows for uploading and ingestion of CDE templates with minimal command line interface, while maintaining all of the validation features of the default ingestion pipeline.

B. Literature search, abstraction process and data collection

Data for ReMeDy pilot-testing, in the format of the multi-modular Common Data Elements (CDE) framework [3-4], described below, were obtained from 52 published CSC projects. The following articles were used as the dataset for this publication, identified by PubMed ID: 17283135, 16322242, 17548814, 18242515, 19665978, 18682804, 23747337, 26400441, 33537082, 33537099, 25121761, 33531658, 33527014, 33525605, 33478031, 33509274, 15994920, 33513805, 33500430, 33536785, 33500726, 33509660, 33482915, 33487592, 33471244, 30042390, 17020963, 29432735, 30091683, 30315141, 27588469, 29730663, 31083655, 27489355, 25193236, 25955492, 22585577, 30941956, 25384215, 25979230, 31151460, 32601337, 32369446, 30905410, 33453469, 33466690, 33430034, 33419140, 33456581, 33468992, 33469161, 33482272 [6-54]. Publications were selected using a

*This work was not supported by any organization.

I. Parvanova is with the Icahn School of Medicine at Mount Sinai, New York, NY 10029 USA (phone: 212-659-9594; e-mail: Irena.Parvanova@mountsinai.org).

K. Borziak, J. Guarino and J. Finkelstein are with the Icahn School of Medicine at Mount Sinai, New York, NY 10029 USA (Kirill.Borziak@mountsinai.org; jennifer.guarino@mssm.edu; Joseph.Finkelstein@mssm.edu)

randomized process from Google Scholar and PubMed search results for "cancer stem cells". Given our focus on original research publications, the resulting papers were filtered to exclude reviews, secondary sources and meta-analyses. The randomized selection process was designed to ensure the inclusion of the full range of CSC research from the initial reports of identification of CSCs to the most current research on CSCs. The resulting publication list ranges from 2005 to 2021. Secondly, the publication selection process was designed to include the full breadth of CSC, by ensuring the inclusion of in vitro, pre-clinical/animal model, and clinical trials on CSC research. The selected CSC publications were not restricted by journal, citations, or other criteria.

Following the selection of our publication set, the data from the publications was abstracted into our multi-modular CDE framework. The abstraction process was conducted manually by trained abstractors with experience in cancer and stem cell research. The focus of the abstraction was on extracting information related to performing CSC assays, such as carcinogenesis, malignancy and self-renewal, and their results. The majority of abstracted terms were extracted as values defined either by permissible values or ontologies. Some CDEs, such as outcomes/findings descriptions and cancer-specific assays, are not amenable to being extracted as specific values and these were recorded as short statements in free-text value fields. The abstracted data in the format of the multi-modular CDE framework was used to create a template for each CSC line examined in each of the published projects included in the test set of publications. Further, a template was created for each individual or grouped study subject (patient, animal model or cell line) reported in the set of publications, linking them to the stem cell product templates. The templates were then converted to JSON format and submitted for ingestion into the database using the upload interface utility.

III. Results and Discussion

ReMeDy is a user-friendly database, which provides detailed, comprehensive information on cancer stem cells. Since CSC data contained in ReMeDy are retrieved from published, publicly available projects, they are freely accessible on ReMeDy without password or registration requirements. The data can be quickly accessed through the search functionality available from the landing page. The landing page also provides direct access to the API interface. The projects contained within ReMeDy are all listed and accessible through the Projects page. While ReMeDy currently contains 52 CSC projects, we will continuously update the database by adding CSC publications, in order to provide an up-to-date resource to facilitate cancer research and help drive forward the generation of knowledge. ReMeDy is designed to provide detailed information not only on the cellular characteristics of CSCs, but also on the research systems (patients, animal models, or cell lines) and research findings. The advantage of this new data resource is the potential to generate novel insights into cancer and CSC research and gain a better understanding of the landscape of CSCs, by bringing together diverse data related to CSCs under one framework.

A. The ReMeDy platform

ReMeDy is an implementation of the Signature Commons platform [3-4], which is designed to store and search diverse metadata in an agile and flexible manner. The platform is based on a relational database for data storage. The advantage of relational databases such as PostgreSQL, utilized in ReMeDy, is their fast performance when storing and searching structured data that is organized within a well-structured schema. This requires the data structure to be defined ahead of time, but the advantages are easy updating, indexing, and fast searches. The large diversity of data and metadata related to CSCs is stored in ReMeDy in the form of common data elements (CDEs). This data technology along with the implemented schema for data storage aims to provide conformity to the FAIR guidelines (Findable, Accessible, Interoperable, and Reusable). The ReMeDy database contains tables for each different class of data object and defined relationships between them, which provides protections against erroneous updates. Finally, indexing of the data allows for fast searching of the data even as the size of the database expands.

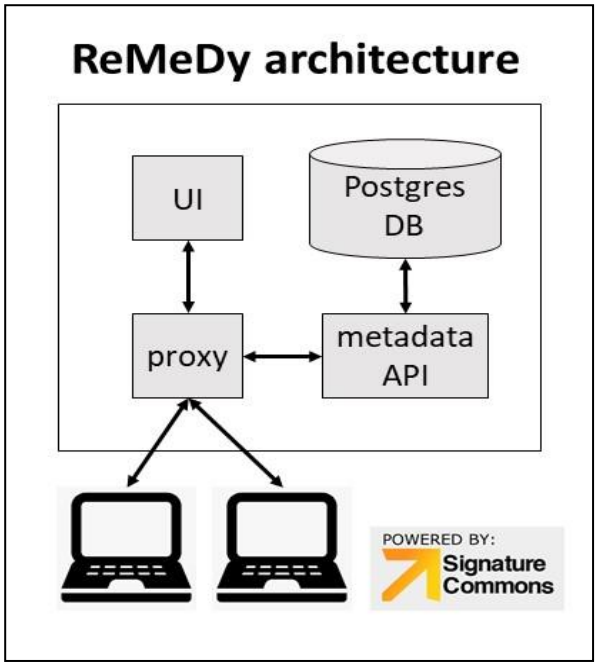

Figure 1. Visual representation of the ReMeDy architecture, showing the interaction between the Docker containers.

ReMeDy is composed of six repositories, which are installed together through the docker platform (Figure 1). These are the controller, data-api, metadata-api, proxy, schema, and UI, with the main functions being done by the proxy, metadata-api, and UI containers. The proxy is the primary access point to the ReMeDy and coordinates between the user interface (UI) and Application programming interface (API) services. The UI repository is the front-end of ReMeDy that is responsible for the visualization of stored data, search functionalities, and graphical representation tools. The API is powered by LoopBack and is responsible for communication with the PostgresDB for both retrieval and uploading of data. It is designed with a Swagger 2.0 JSON implementation, with all the RESTful endpoints returning structured JSON format data. The controller acts as an intermediary to aid the data ingestion process, as both data processing and uploading the processed data through API. The data-api repository functions to set up and install the ReMeDy instance. It is built using Gradle and controls the tasks of compilation, packaging,

testing, deployment, and publishing. The schema package details the JSON-Schema validators used for ReMeDy entities, which designed to flexibly validate the metadata ingested into the ReMeDy database.

To improve the usability of ReMeDy, we developed a user interface that allow quick and efficient data upload by consolidating the command line steps of interacting with the API. We provide public access to the upload interface in order to encourage researcher engagement with the platform and to promote crowd sourced uploading of published CSC data into ReMeDy. The main function of the upload interface is to generate JSON formatted data for upload into the API from user friendly CDE data formats such as Excel or tab delimited files. Additionally, the upload interface is capable of adding and incorporating all the features required for upload, including generating universally unique identifiers (UUIDs) and proper inter-data type connections. After the successful upload of the data file, the back-end reads the data using 'tsvReader' API and generates a JSON object with the defined type and content, which is then submitted to the ReMeDy API.

To ensure strict quality control of the ingested data. All CDE values are submitted for validation prior to the final API upload step. This is accomplished using validator schemas. The validators define which key value pair elements the ingested data can contain, the format of the values, and the required elements for ingestion. The formatting of the values is achieved either by hard coding CDE values, such as those for sex or acceptable ranges for age, or submitted for validation against existing ontologies, such as UniProt to validate protein and gene names.

B. *Multi-modular CDE Framework*

In order to efficiently facilitate data collection and promote a standardized data organization within ReMeDy, we have created a multi-modular CDE framework that reflects the full range of information generated by CSC studies. Previously create frameworks, such as the Minimum Information About a Cellular Assay for Regenerative Medicine (MIACARM) [55] for stem cells, have focused primarily on the cell characteristics and assays to characterize them. Although this format is in the process of being adopted by major stem cell banks such as hPSCreg, it lacks the full range of criteria required to completely characterize CSCs. It does not provide any CDEs to describe other relevant realms of information created by CSC project, such as research system and project outcome information. Our multi-modular CDE framework addresses these deficiencies using a scoping review approach to delineate relevant CDE categories [2]. The framework is organized into 5 modules: Project, Research System, CSC Properties, In-depth Characterization and Outcomes/ Findings (Figure 2). The CDEs within the modules are further hierarchically organized into sections and subsections, to allow for user-friendly navigation.

The Project module contains CDEs, which capture information about the Principal Investigator and general project summary information, Aside from the general project information, which includes CDEs for project title, description, key words, and hypothesis, the module contains four additional sections: PI Information, Publication/ Reference Information, Project Design, and Regulatory Compliance. The Project Design section contains CDEs that characterize the design of the project, such as trial phase, inclusion/exclusion criteria, target sample size, study design, and administrative practice. The Regulatory compliance section CDEs describe the study's compliance with regulatory bodies such as Institutional Review Board and Data Monitoring Committee, along with status of obtaining consent.

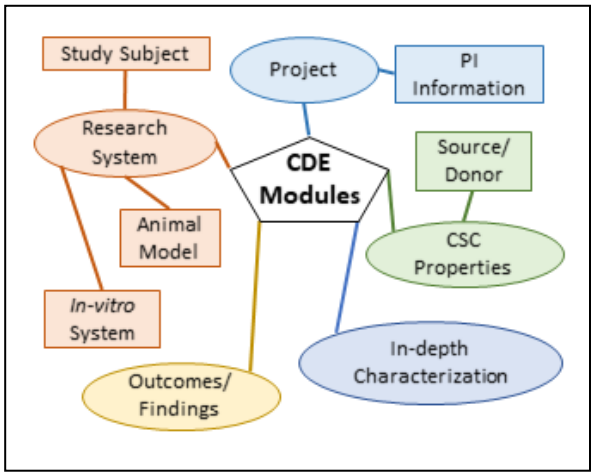

Figure 2. Multi-modular CDE Framework diagram, highlighting the modules and important, representative sections in the classification of their CDEs.

The Research System module contains CDEs that characterize the patients, animal models, and/or in vitro conditions that were used to study CSCs. This module also contains CDEs to characterize any experimental assays that were used to study or monitor the research system as described in a publication. It contains CDEs for Subject ID and Subject data type, which describes if subject data is from individuals or from grouped averages in the publication. Research System CDEs are then organized into four section: Study Subjects, Animal Model, In-vitro System, and Experimental Assays. The Study Subject CDEs are specific for clinical patients and contains CDEs on Informed Consent, Demographics, and Physical/Medical Status. The modular nature of the framework combined with our flexible database structure allows us to easily add disease, medical condition, and baseline/treatment measurement CDEs as required by newly added projects. The Animal Model section is designed to accommodate any animal model used for CSC research. The section is subdivided into Animal Characteristics, Animal source, Species and Strains, Sex, Genetic modification, Housing, and Diet subsections. The In-vitro System contains four subsections: Identification, Origin, Cell line properties, and Propagation. The Identification subsection contains CDEs that are used to identify the cell line or in vitro condition that was used to study CSCs. The Cell line properties CDEs describe growth rate, cell morphology and cellular products that define the cell line. The Propagation subsection details the cell line conditions, such as medium, seed density, and split ration. Finally, the Experimental Assays section features the experiments performed on the CSCs within the research system, such as imaging or transcriptome profiling. Specifically for CSC research, this section was expanded with Cancer-specific assays CDEs that describe assays used to characterize CSCs, such as Carcinogenesis, Chemoresistance, Eradication, Malignancy, and Self-renewal. The CSC Properties module is designed to capture information about the CSCs themselves. Although the initial aim of the framework

was to focus on stem cells, we have successfully adapted the framework to CSCs with minimal modifications. This module contains CDEs describing the source and origin of the CSCs, critical criteria used to define and discriminate CSCs, and information on the maintenance, growth, and/or derivation of the CSCs. The aim is to enable future collaborative sharing of scientific data by making it easier for researchers to differentiate and distinguish between different CSC lines. This module contains three sections: Source/Donor, Critical Quality Attributes, and Critical Process Parameters. The Source/Donor section contains CDEs used to identify and characterize the origin of the CSCs. It is further subdivided into Donor information, Anatomical origin, and Source cell information. The Critical Quality Attributes section details the CSC attributes that are used to define and delineate the CSCs. The CDEs in this section include cell potency, cell viability, morphology, population purity, differentiation propensity, viral copy number, in addition to epigenetic, protein, cell surface and metabolite markers that are used to define the stem cell products under investigation. The Critical Process Parameters section contains CDEs that are used to characterize the process that was used to derive, maintain, differentiate and store CSCs.

The In-depth Characterization module contains CDEs related to the assays used for the characterization of CSCs. This section is dedicated primarily to methodological CDEs that are used to describe the assays performed on CSCs to determine their cellular characteristics. This includes information on tumorigenicity, sterility, viability, pluripotency assays, transcriptomic and genomic profiling, imaging, MRI, and PCR, among others. As previously noted, this section will be expanded to ensure the complete characterization of assays performed on CSCs.

Finally, the Outcomes/ Findings module contains CDEs characterizing the published scientific finding from the CSC research. These CDEs are organized into Primary Clinical Outcomes, Secondary Clinical Outcomes, and Preclinical and In vitro Findings. The design of these sections ensures that the key findings from of each project are in a standardized format. The Primary Clinical Outcomes and Secondary Clinical Outcomes CDEs are modeled on the formatting of ClinicalTrials.gov outcomes. The Preclinical and In-vitro Findings subsection CDEs describe morphological, functional, and pathway results.

Our modular scheme allows for both flexible and comprehensive organization of CSC data, since not all sections or CDEs of the framework need to be shared across all studies. The modular framework creates a comprehensive organization that can be used to describe not only CSCs, but a wide variety of stem cell projects. This is primarily accomplished by allowing for the composition of CDEs to be adjusted, while maintaining a CDE core that allows for detailed comparisons across studies.

C. Data Visualization and Sharing

The landing page of ReMeDy is designed to provide easy access to the available functionalities of the platform. These include search, visualization tools, project links, and the API. The search functionality is designed to provide advanced search functions with intuitive controls, such as Boolean operator functionalities. The search functionality is capable of searching both by CDE name and by CDE value. Additionally, the CDE search is not case sensitive allowing for increased usability. The landing page provides visualization tools, such a pie chart showing the distribution of clinical (patients), pre-clinical (animal models), and in vitro (cell lines) study subjects across the uploaded projects. The ReMeDy Projects page lists all the uploaded published projects, including the 52 CSC studies. The projects links provide direct access to the data stored in the five modules of the CDE framework. The project records additionally serve as links to the individual or grouped patient, animal model, and in vitro research system.

In addition, the search functionality is used to access the individual and grouped research subject records. It works for both project and research system records types automatically. While searching for a specific CDE name/value returns specific record results, a blank search returns all the data currently in ReMeDy. This includes the 52-research system and 52 project records for CSC data. Selecting records displays all the CDEs contained within. The search results highlight our utilization of schemas, which aim to improve the user experience by providing the most relevant information regarding each project. To further improve utility of ReMeDy, we implement filtering schemas. They allow the user to incrementally refine their search query. Additionally, they provide statistical information on the distribution of data within the ReMeDy platform.

The API functionality allows researchers to download the data stored within ReMeDy. Due to the compartmentalization of the data, the data retrieval requests can be directed to either project records or research system data records. Since ReMeDy API uses the Swagger 2.0 JSON implementation, all RESTful endpoints return structured JSON format data. This allows for the data download to be restricted by specific CDE values or for retrieval of specific CDEs. All records can be retrieved using their UUIDs. Since the UUIDs are visible in the search results, this allows for on-demand data retrieval capabilities. The API is fully documented as part of the underlying Signature Commons. The API functionality is designed to increase the usability of the platform and data access to the end-user in order to further foster community data sharing and drive advancements in cancer research.

D. ReMeDy cancer stem cell project dataset

To test the functionality of our multi-modular framework and improve the usability of the ReMeDy platform, we included a diverse set of published CSC projects. Here, we tested a dataset of 52 CSC studies. In the current format, ReMeDy contains an average 70 CDE values per study out of the total 841 CDEs in the multi-modular framework. The CSC dataset also has a core set of over 20 CDEs common to all projects. Our randomized selection of CSC projects included clinical, pre-clinical and in vitro projects (Figure 3a). The majority of the publications are focused on in vitro (42%) and pre-clinical (40%) studies of CSCs. All preclinical trials were performed on mice. Clinical trials comprised the remaining 17% of publications. The majority of CSC research came from China, USA, and Italy (Figure 3b). China accounted for almost a third of our studies with 27%, with USA and Italy accounting for 23% and 13% each.

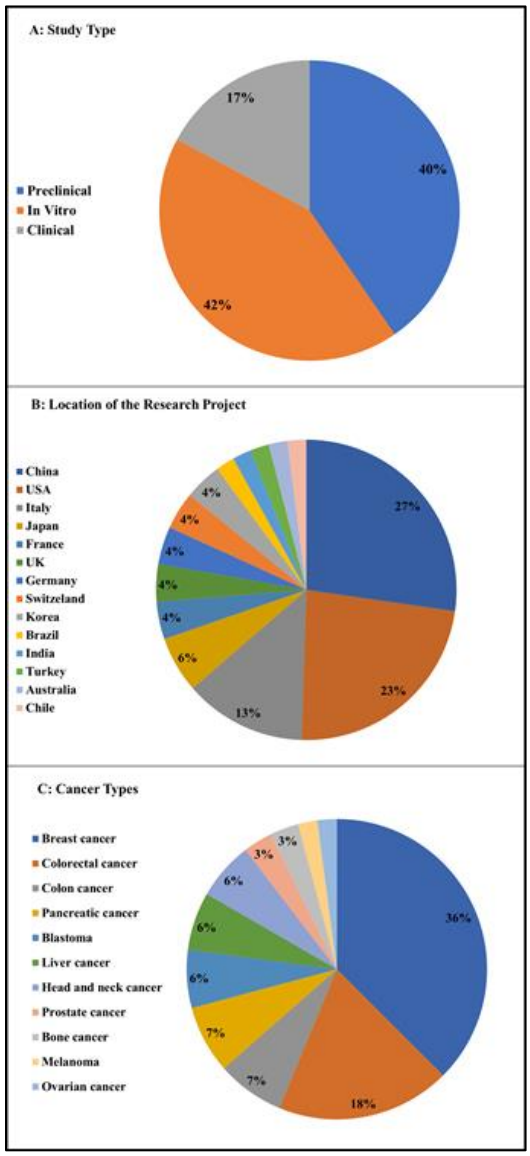

Figure 3. Distribution of CSCs projects in the ReMeDy platform across A) project type, B) country conducting the research, and C) cancer type.

Other countries with published CSC research in our set were France, Switzerland, Japan, Brazil, and UK. Our publication set covers 10 types of cancer (Figure 3c). The most common type of CSCs under investigation (36%) were those from breast cancer. Additionally, colorectal CSCs account for 18% of publications and colon and pancreatic CSCs for 7% each. Other cancer types covered by our publications are head and neck, colon, pancreatic, prostate cancer, melanoma, and leukemia. Similarly, the source organ for CSCs mirrors closely the cancer types under investigation, with breast accounting for the majority of source tissue for CSCs.

IV. Discussion

The expanding field of cancer medicine requires the creation of a flexible and agile repository for cancer stem cell data aggregation, storage, visualization, and sharing. To promote this, we have adapted the Regenerative Medicine Data Repository (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 CSC research project information in a standardized format and provides effortless visualization and search functions. The platform was tested by uploading clinical, preclinical, and in vitro studies, in an organized and systemic manner, using a flexible multi-modular framework, which confirmed ReMeDy functionality as a storage and visualization platform of CSC data. While the current number of projects in ReMeDy is a limitation to its utility, we aim to address this in our future aims.

The platform promotes easy accessibility to CSC projects to facilitate data sharing and collaboration within the field. The utilization of a relational JSON formatting database allows us to import CSC data sets in a schema-less nature, while simultaneously employs validators to ensure stringent quality control of the ingested data. Our data upload process promotes stringent quality control by necessitating validation against ontologies and identification of required elements for ingestion, which allows to select the most informative CSC CDEs. The essential CDEs not only include primary parameters to characterize CSCs, but also allows for standardized cross-discipline and cross-studies comparison.

V. Future Plans and Conclusion

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. In addition, plans for the ReMeDy database include promoting a crowdsourcing functionality of the platform [56-57]. In part, this effort includes establishing and expanding an automated pipeline for uploading the CDE templates, modeled on our upload interface, by providing researchers with a convenient tool to import their own CSC research data, allowing them to bypass more complicated computational stages of JSON formatting. We plan to transform the upload interface into a publicly facing webpage, accessible to registered users. To ensure the quality of data submission, we consider implementation of a login interface to allow researchers to create submission accounts. This will enable them to submit CDE templates in the format of a multi-modular CDE framework, drive data sharing and usability, encourage the development of collaborations, and promote advances in cancer research and treatment. Advantages of crowdsourcing are ensuring that the knowledge base for CSCs stays current.

References

[1] S. Dawood, L. Austin, and M. Cristofanilli, "Cancer stem cells: implications for cancer therapy," Oncology (Williston Park), vol. 28, no. 12, pp. 1101-7, 1110, 2014.

[2] J. Finkelstein, I. Parvanova, and F. Zhang, "Informatics Approaches for Harmonized Intelligent Integration of Stem Cell Research," Stem Cells Cloning, vol. 13, pp. 1-20, 2020.

[3] K. Borziak, I. Parvanova, J. Finkelstein, "ReMeDy: a platform for integrating and sharing published stem cell research data with a focus on iPSC trials," Database (Oxford), vol. 2021, baab038, 2021.

[4] K. Borziak, T. Qi, J. E. Evangelista, D. J. B. Clarke, A. Ma'ayan, and J. Finkelstein, "Towards Intelligent Integration and Sharing of Stem Cell Research Data," Stud Health Technol Inform, vol. 272, pp. 334, 2020.

[5] V. Stathias et al. "LINCS Data Portal 2.0: next generation access point for perturbation-response signatures," Nucleic Acids Res, vol. 48, pp. D431-D439, 2020.

[6] C. Li et al, "Identification of pancreatic cancer stem cells," Cancer Res, vol. 67, no. 3, pp. 1030-1037, 2007.

[7] A.T. Collins et al, "Prospective identification of tumorigenic prostate cancer stem cells," Cancer Res, vol. 65, no. 23, pp. 10946-10951, 2005.

[8] P. Dalerba et al, "Phenotypic characterization of human colorectal cancer stem cells," Proc Natl Acad Sci U S A, vol. 104, no. 24, 2007.

[9] Z.F. Yang et al, "Significance of CD90+ cancer stem cells in human liver cancer," Cancer Cell, vol. 13, no. 2, pp. 153-166, 2008.

[10] Y. Shimono et al, "Downregulation of miRNA-200c links breast cancer stem cells with normal stem cells," Cell, vol. 138, no. 3, pp. 592, 2009.

[11] A.P. Morel et al, "Generation of breast cancer stem cells through epithelial-mesenchymal transition," PLoS One, vol. 3, no. 8, 2008.

[12] W. Gao et al, "Isolation and phenotypic characterization of colorectal cancer stem cells with organ-specific metastatic potential," Gastroenterology, vol. 145, no. 3, pp. 636-46, 2013.

[13] Z.M. Feng et al, "Essential role of miR-200c in regulating self-renewal of breast cancer stem cells and their counterparts of mammary epithelium," BMC Cancer, vol. 23, no. 15, pp. 645, 2015.

[14] K. Li et al, "Metallothionein-1G suppresses pancreatic cancer cell stemness by limiting activin A secretion via NF-κB inhibition," Theranostics, vol. 11, no. 7, pp. 3196-3212, 2021.

[15] S.M. Afify et al, "A novel model of liver cancer stem cells developed from induced pluripotent stem cells," Br J Cancer, vol. 122, no. 9, 2020.

[16] F. Iacopino et al, "Isolation of cancer stem cells from three human glioblastoma cell lines: characterization of two selected clones," PLoS One, vol. 9, no. 8, 2014.

[17] G. Manic et al, "Control of replication stress and mitosis in colorectal cancer stem cells through the interplay of PARP1, MRE11 and RAD51," Cell Death Differ, 2021.

[18] G.M.M Fernandes et al, "Anti-EGFR treatment effects on laryngeal cancer stem cells," Am J Transl Res, vol. 13, no. 1, pp. 143-155, 2021.

[19] V. Castelli et al, "PPARα-Selective Antagonist GW6471 Inhibits Cell Growth in Breast Cancer Stem Cells Inducing Energy Imbalance and Metabolic Stress," Biomedicines, vol. 9, no. 2, pp. 127, 2021.

[20] E.L.M. Gelardi et al, "A Selective Competitive Inhibitor of Aldehyde Dehydrogenase 1A3 Hinders Cancer Cell Growth, Invasiveness and Stemness In Vitro," Cancers (Basel), vol. 13, no. 2, pp. 356, 2021.

[21] M. Chu, H. Wan, X. Zhang, "Requirement of splicing factor hnRNP A2B1 for tumorigenesis of melanoma stem cells," Stem Cell Res Ther, vol. 12, no. 1, pp. 90, 2021.

[22] S.Siljee S et al, "Cancer Stem Cells in Metastatic Head and Neck Cutaneous Squamous Cell Carcinoma Express Components of the Renin-Angiotensin System, "Cells, vol. 10, no. 2, pp. 243, 2021.

[23] K. Husain, D. Coppola, C.S. Yang, M.P. Malafa, "Farnesyl dimethyl chromanol targets colon cancer stem cells and prevents colorectal cancer metastasis," Sci Rep, vol. 11, no. 1, pp. 2185, 2021.

[24] J. Guo, M. Guo, J. Zheng, "Inhibition of Bone Morphogenetic Protein 2 Suppresses the Stemness Maintenance of Cancer Stem Cells in Hepatocellular Carcinoma via the MAPK/ERK Pathway," Cancer Manag Res, vol. 13, pp. 773-785, 2021.

[25] W. Shen et al, "CCL16 maintains stem cell-like properties in breast cancer by activating CCR2/GSK3β/β-catenin/OCT4 axis," Theranostics, vol. 11, no. 5, pp. 2297-2317, 2021.

[26] M. Tanori et al, "Microsecond Pulsed Electric Fields: An Effective Way to Selectively Target and Radiosensitize Medulloblastoma Cancer Stem Cells," Int J Radiat Oncol Biol Phys, vol. 109, no. 5, pp. 1495, 2021.

[27] J. Liu et al, "SOX4 maintains the stemness of cancer cells via transcriptionally enhancing HDAC1 revealed by comparative proteomics study," Cell Biosci, vol. 11, no. 1, pp. 23, 2021.

[28] S. Song et al, "Targeting cancer stem cells with a pan-BCL-2 inhibitor in preclinical and clinical settings in patients with gastroesophageal carcinoma," Gut, 2021.

[29] M. Balic et al, "Most early disseminated cancer cells detected in bone marrow of breast cancer patients have a putative breast cancer stem cell phenotype," Clin Cancer Res, vol. 12, no. 19, pp. 5615-5621, 2006.

[30] T.. Kim et al, "Cellular toxicity driven by high-dose vitamin C on normal and cancer stem cells," Biochem Biophys Res Commun, vol. 497, no. 1, pp. 347-353, 2018.

[31] F. Troschel et al, "miR-142-3p attenuates breast cancer stem cell characteristics and decreases radioresistance in vitro," Tumour Biol, vol. 40, no. 8, 2018.

[32] C. Yan et al, "Enhanced autophagy in colorectal cancer stem cells does not contribute to radio-resistance," Oncotarget, vol. 7, no. 29, 2016.

[33] M. Lin et al, "Prospective study of the safety and efficacy of a pancreatic cancer stem cell vaccine," J Cancer Res Clin Oncol, vol. 14, no. 10, pp. 1827-1833, 2015.

[34] A. Bartley et al, "Colorectal adenoma stem-like cell populations: associations with adenoma characteristics and metachronous colorectal neoplasia," Cancer Prev Res, vol. 6, no. 11, 2013.

[35] D. Kong et al, "Loss of let-7 up-regulates EZH2 in prostate cancer consistent with the acquisition of cancer stem cell signatures that are attenuated by BR-DIM," PLoS One, vol. 7, no. 3, 2012.

[36] C. Wang, Z. Wang, W. Liu, Z. Ai, "CD133 promotes the self-renewal capacity of thyroid cancer stem cells through activation of glutamate aspartate transporter SLC1A3 expression," Biochem Biophys Res Commun, vol. 511, no. 1, pp. 87-91, 2019.

[37] J. Cui et al, "New use of an old drug: inhibition of breast cancer stem cells by benztropine mesylate," Oncotarget, vol. 8, no. 1, 2017.

[38] M. D'Arcangelo et. al, "Cancer Stem Cells Sensitivity Assay (STELLA) in Patients with Advanced Lung and Colorectal Cancer: A Feasibility Study," PLoS One; vol. 10, no. 5, 2015.

[39] V.L Battula, et. al, "Ganglioside GD2 identifies breast cancer stem cells and promotes tumorigenesis," J Clin Invest, vol. 122, no. 6, pp. 2066.

[40] H.Ji, et. al, "The effect of omentin-1 on the proliferation and apoptosis of colon cancer stem cells and the potential mechanism," J BUON, vol. 24, no.1, pp. 91-98, 2019.

[41] D Zhao et al, "NOTCH-induced aldehyde dehydrogenase 1A1 deacetylation promotes breast cancer stem cells," J Clin Invest, vol. 124, no. 12, pp. 5453-65, 2014.

[42] M.I. James et al, "Curcumin inhibits cancer stem cell phenotypes in ex vivo models of colorectal liver metastases, and is clinically safe and tolerable in combination with FOLFOX chemotherapy," Cancer Lett, vol. 364, no. 2, pp. 135-141, 2015.

[43] Y Zhao et al, "Single-cell RNA sequencing reveals the impact of chromosomal instability on glioblastoma cancer stem cells," BMC Med Genomics, vol. 12, no. 1, pp. 79, 2019.

[44] C. Riether et al, "Targeting CD70 with cusatuzumab eliminates acute myeloid leukemia stem cells in patients treated with hypomethylating agents," Nat Med, vol. 26, no. 9, pp. 1459-1467, 2020.

[45] J.RBrown et al, "Phase II clinical trial of metformin as a cancer stem cell-targeting agent in ovarian cancer," JCI Insight, vol. 5, 2020.

[46] K.S. Lee, "Reprogramming of cancer stem cells into non-tumorigenic cells using stem cell exosomes for cancer therapy. Biochem Biophys Res Commun," vol. 512, no. 3, pp. 511-516, 2019.

[47] D.S. Schott et al, "Influence of adjuvant radiotherapy on circulating epithelial tumor cells and circulating cancer stem cells in primary non-metastatic breast cancer," Transl Oncol., vol. 14, no. 3, 2021.

[48] K. Kuramoto et al, "Inhibition of the Lipid Droplet-Peroxisome Proliferator-Activated Receptor α Axis Suppresses Cancer Stem Cell Properties. Genes (Basel)," vol. 12, no. 1, pp. 99, 2021.

[49] Y. Yu et al, "Targeting a Lipid Desaturation Enzyme, SCD1, Selectively Eliminates Colon Cancer Stem Cells through the Suppression of Wnt and NOTCH Signaling," Cells, vol. 10, no. 1, 2021.

[50] D. Kamble et al, "Keap1-Nrf2 Pathway Regulates ALDH and Contributes to Radioresistance in Breast Cancer Stem Cells," Cells, vol. 10, no. 1, pp. 83, 2021.

[51] S.Y Hyun et al, "Evodiamine inhibits both stem cell and non-stem-cell populations in human cancer cells by targeting heat shock protein 70," Theranostics, vol. 11, no. 6, pp. 2932-2952, 2021.

[52] C. Yoon et al, "PI3K/Akt pathway and Nanog maintain cancer stem cells in sarcomas," Oncogenesis, vol. 10, no. 1, pp. 12, 2021.

[53] P. Zhu et al, "A novel hypoxic long noncoding RNA KB-1980E6.3 maintains breast cancer stem cell stemness via interacting with IGF2BP1 to facilitate c-Myc mRNA stability," Oncogene, vol. 40 no. 9, pp. 1609-1627, 2021.

[54] F. Andrade et al, "Polymeric micelles targeted against CD44v6 receptor increase niclosamide efficacy against colorectal cancer stem cells and reduce circulating tumor cells in vivo," J Control Release, vol. 331, pp. 198-212, 2021.

[55] K. Sakurai, A. Kurtz, G. Stacey, M. Sheldon, and W. Fujibuchi, "First Proposal of Minimum Information About a Cellular Assay for Regenerative Medicine," Stem Cells Transl Med, vol. 5, no. 10, 2016.

[56] A. Elghafari, J. Finkelstein, "Automated Identification of Common Disease-Specific Outcomes for Comparative Effectiveness Research Using ClinicalTrials.gov: Algorithm Development and Validation Study," JMIR Med Inform, vol. 9, no. 2, 2021.

[57] A. Elghafari, J. Finkelstein, "Introducing an Ontology-Driven Pipeline for the Identification of Common Data Elements," Stud Health Technol Inform, vol. 272, pp. 379-382, 2020.