D2R DNA TO RNA

D2R RESEARCH **DATA MANAGEMENT GUIDELINES** AND **BEST PRACTICES**

McGill

Version: v1 March 2025
Unanimously approved by the D2R's Research Steering Committee on 2025 April the 8th.

TABLE OF CONTENTS

02 **BACKGROUND & CONTEXT**

03 **GUIDELINES STATEMENT**

03 / Scope
03 / Exclusion
03 / Data management requirements
04 / Guideline coverage

05 **ROLES AND RESPONSIBILITIES**

05 / D2R responsibilities
05 / D2R-funded researchers
06 / Support role of the D2R team

07 **APPENDIX 1 DEFINITIONS**

08 **APPENDIX 2 SUGGESTED BEST PRACTICES**

08 / Data management plans
09 / Data sharing and data catalogue
09 / Open Science
10 / Other recommendations

11 **APPENDIX 3 EXTERNAL DOCUMENTS, STRATEGIES AND REGULATIONS**

12 **APPENDIX 4 NON-EXHAUSTIVE LIST OF METADATA**

D2R DNA TO RNA

BACKGROUND & CONTEXT

McGill's DNA to RNA (D2R) Initiative is a federally funded program that aims to expand the therapeutic arsenal for precision medicine by fulfilling the potential of genomic discovery and capitalizing on the flexibility of RNA therapies to benefit all, including marginalized and medically underserved groups.

The highest professional and disciplinary standards, both nationally and internationally, are expected to be met by D2R researchers. These standards are in place to uphold research excellence. They ensure that research is conducted ethically while ensuring efficient use of funds and the replicability of experiments and studies. Research data management (RDM) is considered a necessary part of research excellence. To this end, the guideline recommends the FAIR guiding principles for research data management and the CARE Principles for Indigenous Data Governance. The guidelines also recognize the need to protect the security of sensitive Canadian research to prevent it from being misappropriated to the detriment of national security.

Appendix 1 provides definitions of key terms to aid in understanding the data management concepts presented in this document.

GUIDELINES STATEMENT

SCOPE

The D2R Research Data Management (D2R-RDM) Guidelines apply to all research projects funded in whole or in part by McGill's D2R Initiative. These guidelines cover research data, including but not limited to genomic, transcriptomic, proteomic, metabolomic, microscopy (image and video), screening, and surveys. It also includes the software and source code used to produce and process the data and the communications materials.

Clinical research is covered as well by the current guidelines, however, due to their sensitive nature, they may require additional provisions to ensure compliance with specific ethical, legal, and regulatory requirements.

EXCLUSION

The D2R Research Data Management (D2R-RDM) Guidelines do not apply to preliminary analyses, plans for future research or physical objects (e.g., laboratory specimens).

DATA MANAGEMENT REQUIREMENTS

Data management plan

Methodologies that reflect RDM best practices, such as FAIR, should be incorporated into all D2R-funded research projects. Every D2R-funded project must submit a data management plan (DMP). D2R will require data management plans to be submitted either at the time of application or after a short period not exceeding 3 months post-award.

Data deposit

Funded projects must deposit all relevant research data, metadata, and code that directly support the research conclusions in journal publications and preprints resulting from D2R-supported research into a digital repository. The definition of what qualifies as relevant research data is highly contextual and should be guided by specific disciplinary standards. In support of these practices, the D2R Data Catalog will be established as a centralized registry to enhance data sharing, accessibility, and reuse across D2R-funded projects. Additionally, the McGill Research Data Services provides Guidelines to Deposit your Research Data.

Regulatory compliance

Data management should be performed in accordance with the requirements of the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans - 2nd edition (2022) and the Tri-agency Research Data Management Policy (2021). These statements offer guidance on data management issues related to human subjects research, including consent, confidentiality and privacy, rights of Indigenous people, secondary use of data, and data linkage. Additionally, data management must be carried out in compliance with the Tri-Agency Framework: Responsible Conduct of Research (2021) guidelines.

The D2R Initiative and funded projects should ensure compliance with Quebec privacy legislation (for example, The Québec Loi 25 (Bill 64) "Act to modernize legislative provisions as regards the protection of personal information"), which governs the protection of personal information. In the case where the project Principal Investigator is from another Canadian province and where the data is stored only within that province, the privacy legislation of that province may apply instead.

Intellectual property

The McGill Policy on Inventions and Software and the McGill Policy on Copyright govern the reporting and administration of Intellectual Property (IP) and must be adhered to for all research activities conducted under the D2R Initiative. Researchers who are seeking to commercialize any research software or inventions are required to report them to the McGill Office of Innovation and Partnerships as soon as possible. It is the responsibility of researchers to ensure that all IP related to their work is identified, documented, and disclosed as early as reasonably possible. This includes software, algorithms, and other protectable IP resulting from data handling or analysis. To maintain confidentiality during the review and patenting process, information related to IP must be securely stored and accessible only to authorized individuals.

Research ethics

All research projects involving Human data must be reviewed and approved by an appropriate ethics board (IRB/REB Project Approval). It is essential to ensure that the study design takes Ethical Considerations into account, including minimizing potential harm and maximizing benefits to participants, protecting their rights and welfare, and respecting their dignity.

All research projects involving animals must be conducted in compliance with the guidelines of the The Animal Compliance Office.

Data sovereignty

For research conducted by and with First Nations, Métis, and Inuit communities, collectives, and organizations, Data management Policies, plans and strategies must be developed in collaboration with these entities in accordance with RDM principles or DMP formats that they accept. In the context of research involving these communities, DMPs should recognize Indigenous data sovereignty and incorporate provisions consistent with OCAP and CARE principles, including mechanisms for renegotiating the DMP as needed.

Data security and privacy

Protecting research data's privacy and security is crucial. Privacy protects study participants' personal information; data security makes sure that sensitive information is kept private and not compromised. D2R-funded projects should put strong security measures in place to safeguard against breaches, illegal access, and other security risks. Some of these methods could include response plans, access control, encryption, and anonymization. These measures are necessary to promote a safe and ethical research environment. The McGill Research Data Services provides general Guidelines and Resources on Cybersecurity.

GUIDELINE COVERAGE

Guideline review

As the context for research data management evolves, D2R, in consultation with the partners and experts for research data management, will review and revise this statement as appropriate.

Additional recommendations

D2R empowers researchers to design and manage their projects in alignment with their unique scientific vision and objectives. While every D2R-funded initiative must adhere to the mandatory guidelines specified in this section, we encourage researchers to embrace our supplementary best practices outlined in ***Appendix 2***. These recommendations have been crafted to help researchers improve their data management strategies to meet international standards and emerging best practices. By voluntarily adopting these additional best practices, researchers can significantly promote data interoperability and facilitate ethical, open research practices.

Complementary aspect

It's important to note that these guidelines are complementary and do not replace any existing regulations at the institutional, provincial, national, or international levels that pertain to D2R researchers and their teams. A non-exhaustive list of exiting documents that may apply to D2R researchers have been compiled in ***Appendix 3*** for information.

ROLES AND RESPONSIBILITIES

D2R RESPONSIBILITIES

D2R is dedicated to promoting excellence in research through effective data management practices. As part of this commitment, D2R is responsible for developing and maintaining guidelines and best practices that align with McGill University and the Tri-Agency vision. This ensures the effective management of research data across all funded projects. These guidelines and best practices provide a clear framework for creating and implementing DMPs, ensuring compliance with FAIR principles as well as adherence to ethical, legal, and security standards. D2R also manages the implementation of these guidelines to promote collaboration, innovation, and transparency within its research community.

D2R is promoting the highest standards of IP management. IP is a crucial aspect of the research process within the D2R Initiative. Acknowledging and respecting IP rights ensures that researchers are duly recognized for their contributions and promotes innovation. By emphasizing the importance of intellectual property, the D2R initiative aims to foster a research culture that prioritizes innovation, equity, and integrity. D2R adheres to intellectual property regulations and guidelines provided by McGill University to ensure that research activities meet ethical and legal standards and provide a mechanism so that inventors have legal protection for their inventions. This approach not only safeguards the interests of researchers and institutions but also contributes to the advancement of precision medicine and the wider scientific community. The McGill Office of Innovation + Partnerships offers essential resources to advise and support Faculty on the protection of their IP.

D2R is committed to ensuring that data is "as open as possible and as closed as necessary." We take responsibility for making our data accessible to promote reusability and accelerate research while also prioritizing the privacy of subjects, adhering to legal obligations, and safeguarding any identified commercial or strategic assets. To fulfill this commitment, we will provide resources, such as a Data Catalogue, that will help data sharing and reuse within the D2R research community.

D2R-FUNDED RESEARCHERS

Researchers funded by D2R are responsible for following the guidelines established by D2R to ensure the safety, privacy, and security of research data throughout their projects. This involves creating and maintaining comprehensive DMPs that align with D2R's guidelines and FAIR principles. Researchers must ensure that all data collected, stored, and shared comply with ethical and legal standards, protecting participant privacy and maintaining data integrity. Additionally, they are expected to implement adequate security measures to prevent unauthorized access or data breaches.

Researchers funded by D2R are responsible for following the guidelines established by D2R to ensure that data supporting their research conclusions are shared in accordance with the FAIR principles and the standards of their respective disciplines whenever ethical, cultural, legal, and commercial considerations permit. Data supporting articles should be provided on the publication date and linked to the publications. Whenever possible, these data, along with their metadata and code, should be linked to the publication using a persistent digital identifier. In specific circumstances, an extended period of exclusive access to data (embargo) for primary researchers is acceptable to protect D2R-funded research projects.

Researchers funded by D2R are expected to conduct administrative research activities responsibly, maintaining the highest levels of integrity and compliance with the D2R Research Data Management Guidelines outlined in this document. They must ensure that all supervised staff, students, and anyone involved in the conduct or management of the sponsored project are fully informed of these guidelines and agree to adhere to them.

SUPPORT ROLE OF THE D2R TEAM

The D2R team is committed to supporting D2R researchers in managing their research data effectively and enhancing the quality of their data strategies and protocols. A primary focus of this support is to help researchers comply with D2R's RDM guidelines and best practices, which include developing and maintaining comprehensive DMPs. To assist with this, the D2R team will provide clear guidance, practical templates, and personalized consultations to ensure researchers can efficiently manage their data throughout the research lifecycle.

Another essential responsibility of the D2R team is to improve research practices. By promoting better data strategies and streamlined workflows, the team will help researchers adhere to the FAIR principles ensuring that data is Findable, Accessible, Interoperable, and Reusable. This approach not only facilitates compliance but also enhances the scientific impact of their work by making data more discoverable and usable for the broader research community.

Central to these efforts is the D2R Data Catalogue, which the D2R team will be responsible for maintaining and developing. The team will ensure that this resource remains accessible and beneficial for all investigators. This includes creating intuitive user interfaces, offering detailed guidance, and providing training through webinars and workshops to help researchers maximize the catalogue's capabilities. For more personalized needs, the team will offer additional support, including assisted analyses and guidance on integrating datasets into broader research initiatives.

The D2R team is dedicated to reviewing the D2R-RDM guidelines and best practices as needed to ensure compliance with the data management strategy of the Tri-Agency, McGill, and any other relevant institutions. The D2R team will also keep up-to-date with international developments in data management to continue promoting and supporting best practice applications aligned with these international standards.

Collaboration and responsiveness are central to the D2R team's mission. The team will actively seek regular feedback from researchers to adapt tools and resources to meet evolving needs, ensuring that the catalogue and other support remain aligned with the community's priorities. By fostering an open and supportive environment, the D2R team enables researchers to manage their data effectively while concentrating on their scientific objectives, thus promoting a culture of excellence and collaboration across the initiative.

APPENDIX 1
DEFINITIONS

- **Research data:** Any recorded, factual material that is created, collected, or reused during research activity and widely accepted in the scientific community to validate research findings. This includes raw data, processed data, and datasets generated through experimentation, analysis, or repurposing of existing data.
- **Metadata:** Contextual or supplementary information about the data that facilitates its finding, comprehension, and utilization. Generally, metadata contains details on the author, creation date, methodology, and environment in which the data was produced.
- **Data management plan (DMP):** An evolving document that describes the procedures for managing data both during and after a study. It contains information about procedures for gathering, storing, preserving, protecting and exchanging data.
- **FAIR Principles:** Guidelines to ensure data is *Findable, Accessible, Interoperable, and Reusable.* These principles aim to enhance the ability of machines and humans to find, access, and reuse data.
- **Open access:** The process of providing unrestricted access to publications and research data, making them freely available to the public and academic community. Controlled access may be required before publication.
- **Controlled access:** Restrictions on data access to protect privacy, comply with legal requirements or safeguard commercial or strategic assets. Controlled access ensures that sensitive information is shared responsibly.
- **Indigenous data sovereignty:** Indigenous Peoples' right to control how their own data is gathered, owned, and used. This includes adhering to principles such as OCAP (Ownership, Control, Access, and Possession) and CARE (Collective Benefit, Authority to Control, Responsibility, and Ethics).
- **Ethics board:** A committee responsible for ensuring that participants' rights and welfare are carefully protected. To achieve this objective, it reviews research involving human subjects to ensure that it complies with high ethical standards.

APPENDIX 2
SUGGESTED BEST PRACTICES

Effective data management is essential for maintaining the integrity, accessibility, and reproducibility of research. This appendix provides recommendation for creating DMP and for effectively supporting the development of a comprehensive Data Catalogue. This appendix also provides general best practices suggestions on data management related topics such as open science, data sharing and data standard. All these recommendations are non-mandatory and could be adapted, modified depending on the specificity of each research project.

DATA MANAGEMENT PLANS

The DMP offers a structured approach to overseeing research data throughout its lifecycle, from collection to preservation and sharing.

Content

DMPs are dynamic documents that can be updated as a research project progresses to meet evolving requirements. Depending on the research project, DMP length and content may vary, but we recommend to include the following sections:

- **Data collection**
  - Describe the process and the type of data that will be collected during the project.
- **Metadata and documentation**
  - Describe the documentation and metadata that will accompany the data.
- **Backup and storage**
  - Describe where and how the data will be stored. This should include a description of cybersecurity and privacy measures in place to protect the data
- **Preservation**
  - Describe where and how the data be deposited for long-term preservation.
- **Data sharing and reuse**
  - Describe the appropriate data-sharing permission to be adopted.
- **Responsibility and resources**
  - Describe the role of the research team members concerning data management.
- **Contractual, legal and ethical compliance (if applicable)**
  - Describe any commercial, legal, and ethical issues related to the data.

Efficient and cost-effective

Data management should be economical and effective. While all data must be handled, not all data must be shared or kept forever; the costs and advantages of doing so should be taken into account while planning the data management process.

DATA SHARING AND DATA CATALOGUE

The development of a comprehensive D2R data catalogue is an essential part of our leading international program for excellence and innovation in discovering and developing genome-guided, RNA-based precision medicines. This catalogue will act as a central registry for various digital datasets, including genomic, proteomic, and metabolomic data between others, allowing researchers to easily access, share, and analyze high-quality datasets. It will be important if funded projects can accurately report their metadata as early as possible to keep the catalogue up-to-date and maximally beneficial for the D2R community. By including diverse data types, we can promote interdisciplinary research and develop comprehensive precision medicine solutions. Making this data available to other researchers will help reduce data duplication and promote the FAIR (Findable, Accessible, Interoperable, Reusable) principles, leading to cost-effective use of generated data. Establishing this data catalogue will significantly enhance research productivity, promote collaboration, streamline the discovery process, and drive the just and ethical advancement of precision medicine by providing timely access to critical data resources.

We recommend that D2R researchers share their data as early as possible in the research process, considering the data is informative and of sufficient quality. Data release can be phased in as the investigation progresses, starting with the sharing of metadata. For sensitive data, additional precautions may be necessary, such as implementing controlled access or data visiting methods. These approaches ensure secure and compliant sharing while protecting privacy and ethical considerations. Such methods are widely accepted for datasets that have confidentiality or security requirements, allowing for meaningful data sharing without compromising data integrity or participant privacy.

Consulting with the D2R team to plan your data-sharing strategy can help streamline your data management efforts throughout your research project. It can also maximize the value of your data, enhance its visibility, and potentially create new collaboration opportunities, all while contributing to the improvement of the D2R Data Catalogue.

OPEN SCIENCE

Open access

Open-access publishing and pre-print archives facilitate widespread dissemination of research results. Open access enables researchers to make their publications freely available to the domestic and international research community and to the public at large, thereby enhancing the use, application and impact of research results. D2R encourages the use of early open-access resources, to favor research results and findings dissemination into the community. However, the peer-review publication model should remain the standard for long-term sharing. To facilitate this process, D2R recommends the following steps:

- **Choose reputable open-access journals:** Select journals that adhere to rigorous peer-review standards to ensure the quality and credibility of your work.
- **Deposit in reputable pre-print archives**: Share preliminary results in established pre-print repositories relevant to your field, which can help garner feedback and improve the research before formal publication.
- **Consider institutional repositories:** Submit publications to your institution's open-access repository to enhance accessibility and discoverability.
- **Engage in outreach:** Actively promote your open-access publications through social media and academic networks to reach a broader audience.

Open software and protocol

Sharing of source code and protocols offers an opportunity to record and disseminate these research results and to recognize their significance in ensuring accuracy and promoting transparency in the scientific method. Research is also more transparent, rigorous, and reproducible when methods and software are shared. D2R recommend the following process while developing research software and protocols:

- **Publish in community registries:** Make your source code and protocols available in established community repositories (like GitHub, Bitbucket, etc), where other researchers can access and reference them easily.
- **Adopt version control:** Utilize version control systems (e.g., Git) to track changes and manage updates effectively, which helps maintain clarity and facilitates collaboration.

- **Provide comprehensive documentation:** Ensure that all shared resources are accompanied by clear and detailed documentation, including user guides and installation instructions, to facilitate their use and adaptation by others.
- **Share as early as possible:** Release your software and protocols as soon as they are functional, rather than waiting for final results, to encourage early feedback and improvements.
- **Align with FAIR principles:** Make certain that shared code and protocols adhere to the FAIR data principles by providing clear metadata, persistent identifiers, and machine-readable formats where applicable.

By implementing these recommendations, researchers can significantly contribute to a more transparent, rigorous, and collaborative scientific environment.

Shared data acknowledgment

Research data, used as primary sources to support scholarly work, are essential for scientific and technological inquiry. They serve as evidence of the research process and are gathered through various methods, like experimentation and analysis. The advancement of science increasingly depends on the accessibility and reuse of research data. It is important to acknowledge data as valuable and valid research products. When using research data, all users should provide credit to the original producers of the data through citation and other appropriate discipline-specific standards (e.g., authorship for unpublished data). Thus, users who gain from the reuse of the data should acknowledge researchers who share their data ethically and productively. Please consult McGill Library guidelines for Citing and publishing research data for more specific guidance on this point.

OTHER RECOMMENDATIONS

Data standards for long-term usability

It's advisable to manage data in line with the most appropriate and relevant standards and best practices of your research field, keeping in mind that these standards are evolving rapidly. During your research project, consider collecting and storing data using programs and file formats that ensure secure storage and facilitate data preservation and access even after the project concludes. This approach will enhance the longevity and usability of your research data.

Metadata

Making research data accessible, comprehensible, and reusable for future users, relies on robust metadata usage that adheres to disciplinary and international best practices. Good metadata not only aids the technologies that utilize or mine research data but also enhances the data's visibility through effective indexing in search engines. While metadata standards may vary across disciplines. Additionally, consider incorporating any information that facilitates comprehension and reuse. Whenever possible, following common metadata standards, like employing controlled vocabularies will improve the consistency and usability of your data. This approach could significantly benefit future users and ensure your research can be effectively built upon. ***Appendix 4*** provide a non-exhaustive list of examples of essential metadata.

APPENDIX 3
EXTERNAL DOCUMENTS, STRATEGIES AND REGULATIONS

Non-exhaustive list of regulation from the Tri-Agency, McGill and other D2R's partner institutions:

- **McGill regulation and guidelines for questions regarding IP:**
  a. Policy on copyright
  b. Policy on inventions and software
  c. Guidelines on inventions and software
- **McGill regulation for the conduct of research:**
  a. Regulation on the conduct of research
- **McGill and Tri-Agency regulation for sensitive, health and personal data:**
  a. Policy on the governance of personal information
  b. Policy on research with Human participants
  c. REB Office process
  d. Faculty of Medicine and Health Sciences - IRB process
  e. Tri-Agency Policy on Ethical Conduct for Research Involving Humans
- **RDM strategy resources at partner institutions (for non-McGill D2R researchers)**
  a. McMaster University - RDM strategy
  b. Université de Sherbrooke - RDM strategy (in French)
  c. UBC - RDM strategy
  d. University of Ottawa - RDM strategy
- **Ethics resources at partner institutions (for non-McGill D2R researchers)**
  a. McMaster University – Support for Researchers - Ethics
  b. Université de Sherbrooke - La recherche avec les êtres humains (in French)
  c. UBC – Office of Research Ethics
  d. University of Ottawa - Office of Research Ethics and Integrity

Non-exhaustive list of regulation from the Quebec and provincial governments

- **Quebec regulation applying to Quebec based researchers:**
  a. Act respecting the protection of personal information in the private sector (Law 25 – previously Bill 64)
  b. Civil Code of Quebec, Chapter III - Articles 35 to 41, Protecting privacy and reputation
- **Federal regulation applying to non-Quebec based researchers or inter-provinces projects:**
  a. Personal Information Protection and Electronic Documents Act (PIPEDA)

APPENDIX 4
NON-EXHAUSTIVE LIST OF METADATA

The current list has been created to help D2R researchers understand the importance and extent of the information that could be included as metadata.

It is important to note that the more information is compiled in the metadata, the more likely it is that your data will be shared, reused, and cited. This list has been compiled based on information and recommendations gathered from various sources, including McGill Digital Research Services, FAIR principles, MIAME and MINSEQ, IUPAC, MIxS, and others.

Metadata Category	Metadata Field	Description	Metadata Type	Example
General	Project Title	Name of the study or dataset	Text	LNP Optimization for mRNA Delivery
General	Principal Investigator (PI)	Lead researcher(s) responsible	Text	Dr. XX
General	Institution	Affiliated research institution(s)	Text	McGill University
General	Date of creation	When the data have been generated or created	Date	01/01/2025
General	Project Description	Summary of research goals & methodology	Text	This study optimizes lipid nanoparticles for efficient mRNA delivery to lung cells.
General	License & Access	Data usage restrictions or sharing policy	Text	CC-BY 4.0, Controlled Access
Genomics	Organism	Species from which RNA is extracted	Text	Homo sapiens, Mus musculus
Genomics	Sample Type	Type of biological material used	Text	Whole blood, Tissue biopsy
Genomics	Tissue / Cell Type	Source of RNA within the organism	Text	Liver, T cells
Genomics	Condition / Treatment	Experimental variables (e.g., disease state, drug treatment)	Text	Healthy vs. Diseased, Treated vs. Control
Genomics	Library Preparation Method	Details of RNA extraction and preparation	Text	Poly(A) selection, Ribosomal RNA depletion
Genomics	Sequencing Platform	Technology used for RNA sequencing	Text	Illumina NovaSeq 6000
Genomics	Read Length & Type	Sequencing read specifications	Number	150
Genomics	Library Type	Sequencing library type	Text	Paired-end
LNP	Lipid Composition	Type and ratio of lipids used	Text	Ionizable lipid (50%), DSPC (10%)
LNP	mRNA Sequence Encapsulated	Target sequence of the RNA within LNP	Text	SARS-CoV-2 Spike Protein
LNP	Encapsulation Efficiency (%)	% of RNA successfully encapsulated in LNPs	Number	[0-100]
LNP	Particle Size Distribution (nm)	Size characterization of nanoparticles	Number/ Interval	80
LNP	Surface Charge (Zeta Potential, mV)	Measurement of surface charge affecting stability	Number	-10
LNP	Cryo-EM Imaging Data	Structural confirmation of LNP morphology	Text	Electron Microscopy Files (TIFF)
LNP	LNP Delivery Route	Method of administration	Text	Intravenous, Intramuscular
Chemistry/Mollecule	Molecule Name	Name of the designed molecule	Text	siRNA-Inhibitor-101
Chemistry/Mollecule	Chemical Structure	Structural formula	Text	SMILES or InChI notation
Chemistry/Mollecule	Molecular Weight (g/mol)	Weight of the designed compound	Number	650
Chemistry/Mollecule	Purity (%)	Chemical purity after synthesis	Number	[0-100]
Chemistry/Mollecule	Solubility (mg/mL)	Solubility in different solvents	Number	10
Chemistry/Mollecule	Spectral Data	NMR, Mass Spectrometry, UV-Vis characterization	Text	NMR Spectra (Bruker), LC-MS Files
Chemistry/Mollecule	Target Binding Assay	Data from molecule-target interaction studies	Text	SPR, ITC, ELISA
Synthetic Biology	Vector Backbone	Plasmid or construct used	Text	pET28a
Synthetic Biology	Promoter Type	Regulatory region controlling expression	Text	T7, CMV
Synthetic Biology	Codon Optimization Strategy	Was the sequence optimized for expression?	Text	Humanized Codon Usage
Synthetic Biology	Expression System	Host system for sequence expression	Text	E. coli, HEK293
Synthetic Biology	Sequence Verification Method	How was the construct validated?	Text	Sanger Sequencing, NGS
Synthetic Biology	GC Content (%)	GC content of the designed sequence	Number	[0-100]
Processing/ Computational	Raw Data Format	File format for raw experimental data	Text	FASTQ, CSV, TIFF
Processing/ Computational	Processed Data Format	Final data output format	Text	BAM, VCF, JSON
Processing/ Computational	Alignment Tool & Version	Software used for read mapping	Text	STAR v2.7.10a
Processing/ Computational	Expression Quantification Tool	Software for RNA expression analysis	Text	Salmon v1.5
Processing/ Computational	Normalization Method	Method used for RNA expression normalization	Text	TPM, RPKM, DESeq2
Processing/ Computational	Other Software & Algorithms Used	Computational tools for data analysis	Text	Goseq, Homer
Processing/ Computational	Metadata Standard Used	Standard metadata format	Text	MIAME, MINSEQE
Management	Data Repository	Where the dataset is stored	Text	D2R Data Catalog, GEO, ENA
Management	Data Availability	Public, Restricted, Controlled	Text	Controlled
Management	Data Security Measures	Encryption, access controls, etc.	Text	Encrypted SFTP
Management	Data Governance Compliance	Compliance with OCAP, CARE principles	Text	OCAP, CARE, ERB

D2R DNA TO RNA

McGill