

Data Management in Toxicological Science: Collating materials from life scientists to deliver high quality and readily accessible biological data in the SEURAT-1 cluster

Introduction

ToxBank establishes a dedicated web-based warehouse for toxicity data management and modelling, a "gold standards" compound database and repository of selected test compounds, and a reference resource for cells, cell lines and tissues of relevance for *in vitro* systemic toxicity research carried out across the SEURAT-1 cluster. The project develops infrastructure and service functions to create a sustainable predictive toxicology support resource going beyond the lifetime of the program.

Simply putting all protocols and data into a centralized database will not ensure efficient use of that data if it is not available to scientists in user friendly formats and organized to support the needs of the cluster. An important early step is establishing a good level of understanding between database constructors and life scientists. Here we report the development of a set of Biological Materials Requirements (BMR) for the ToxBank database and a subsequent consultation with SEURAT-1 scientists as the future end users.

Key Objective

To construct requirements for biological information content and format for ToxBank to serve the SEURAT-1 cluster.

Drafting the biological materials requirements

The main activity was to identify the key data types that would be required within each of the SEURAT-1 groups.

Described below are each group of biological materials requirements identified.

The additional elements included after consultation with SEURAT-1 scientists are shown **in grey**.

1. Cell Line Descriptors

- Name of cell line
- Supplier reference (critical identifier for traceability)
- iPSC Nomenclature
- Cell origins (tissue: type, disease mutation present, **cell line passage number/information-genotype/phenotypic stability**)
- Derivation method:
 - hESC isolation method
 - iPSC method and constructs
- Ethical issues:
 - Fully informed consent (statement from originator; EC and hESCCreg database)
 - Donor constraints

2. Supplier Specific Information

- Supplier identity and contacts
- Minimal data set generated (standards - ISCB 2009; CSC 2010) specific to each supplier
 - Quality Control
 - Safety testing
 - Characteristics
- Access arrangements
- General description of charges and application process
- Summary of terms and conditions, agreements and supplier constraints

3. Publications

- Scientific literature/**and other resources**
 - Derivation
 - Differentiation
- Standard Operating Procedures:
 - Culture, preservation and differentiation
 - SOPs for same cell type from multiple sources

Notes: Origin ref and version numbers will change frequently

HeMiBio and DETECTIVE projects are currently preparing SOPs for all assays that will be carried out on primary hepatocytes with the purpose of probing/monitoring their functionality

Scoping biological materials requirements

For the BMR analysis, ToxBank investigators were cell biologists engaged in stem cell laboratory work and had previously developed stem cell line scientific data bases (www.hescreg.eu).

Accordingly, they established a draft set of types of biological data that ToxBank would expect to collate:

1. Cell identifiers
2. Supplier specific data: cells, SOPs, QC data
3. Published information: known characteristics of specific cells
4. User feedback on cells/methods: enable continuous updates
5. Tissues and primary cells
6. Specialist cell culture requirements
7. European legislative issues
8. -omics data: (including raw data or meta data)

4. Types of Information Preferred

- iPSC growth and differentiation – HeMiBio and SCR&Tox input
- iPSC stability – HeMiBio and DETECTIVE (epigenetic data) input: Data would probably include DNA methylation patterns. Histone modifications appear to be less consistent patterns. And data may be less useful. MicroRNA profiles would be important here. Further information to come from HeMiBio who will also provide transcriptome data.
- Hepatic cell lines
- High throughput
- Materials requirements for cellular barrier assays - skin/blood-brain barrier etc
- **Further information will come from the HeMiBio project who will work on endothelial cell barrier model from M18**
- Materials requirements for 3D architectures, bioreactors, specialist cell culture applications and miniaturised sub-systems

Notes: It is important to get further input from NOTOX on the hepatic cell lines and SCR&Tox on the high throughput. It is also desirable to identify additional SEURAT-1 contacts concerning the cellular barrier assays. No bioreactor will be developed in DETECTIVE. This information will be sought from the HeMiBio project.

5. Tissues and Primary Cells

Human and animal tissues and primary cells will be used at some level within the project with the following types of information needed:

- Sourcing
- Harvest/isolation, preservation, culture and stability
- QC assessment criteria for release and use
- SOPs for cell preparation (stem cells, **feeder cells genetic modification, hepatocyte selection (HeMiBio) for tox assays**).
- Ethics and other issues including transport described within supplier info

Notes: SOPs for cell preparation are currently being prepared with the DETECTIVE and HeMiBio project (will be in place by 2012).

6. Further Specialised Data Types

A wide range of technologies involving biological materials will be used across SEURAT-1. Initial feedback on these requirements is given below.

HeMiBio feedback:

- **Culture conditions, species differences, genotypic stability**
- **Lists of reagents (antibodies, PCR primers from HeMiBio for differentiation to hepatocytes and miRNA expression by M6)**
- **Information on development of bioreactors (HeMiBio from M18)**

Authors

Glyn Stacey, Lyn Healy

UK Stem Cell Bank, National Institute for Biological Standards and Control (NIBSC-HPA)

Glenn Myatt

Leadscope

Barry Hardy

Douglas Connect

7. Legislation

Feedback from Partners requested ToxBank to capture legislation information on biological materials (e.g. safety, shipment, genetic manipulation, infectious hazards). There was also a request for ToxBank to record variations arising from National legislation and rules which would include regulation on the use of human embryonic stem cell lines.

8. 'omics and other data output

- Standard Operating Procedures (**HeMiBio Published information; will develop with time**)
- User feedback on cells/methods
- **DETECTIVE:**
 - **Genomics: Affymetrix Gene expression arrays**
 - **Proteomics: 2D Page and selective characterisation by MALDI-TOF, QTOF MS**
 - **Metabonomics: GC-MS, NMR spectroscopy**
 - **Functional readouts**
 - **Electrophysiology data for cardiomyocytes (LAT, ARI, Q-T interval, beat freq)**
 - **Impedance measurement (Xcelligence, Roche)**
 - **High Content Imaging**

Notes: Additional details are to be obtained from HeMiBio, DETECTIVE, SCR&Tox and NOTOX

Conclusions and Next Steps

Having completed the draft biological materials requirements, initial feedback from SEURAT-1 partners has been incorporated and information gathered on timing of new data arising from the different consortia (a second round of review and consultation is already under way)

- ToxBank will also explore the possibility of holding information on most relevant national legislation such as regulation of hESC lines.
- ToxBank is now considering extending its wiki resource used for compound information to the biological materials requirements
- ToxBank is currently testing a first version of the data warehouse for uploading, tracking, archiving, and searching SOPs

Acknowledgements: We acknowledge the generosity of SEURAT-1 scientists (particularly NOTOX, HeMiBio, SCR&Tox and DETECTIVE) in contributing their time for discussions with ToxBank investigators.