OECD Extended Advisory Group for Molecular Screening and Toxicogenomics (EAGMST)

Guidance Document for Consistent Reporting of ‘Omics Data From Various Sources

Transcriptomics Reporting Framework (TRF)

EAGMST Meeting, Paris FR
June 29th, 2018
Disclaimer

• The views expressed in this presentation are those of the author(s) and do not necessarily represent the views or policies of the U.S. Environmental Protection Agency.
Agenda

• Background
  • Regulatory Acceptance of transcriptomic data
  • ‘Omics Reporting Frameworks

• Transcriptomics Reporting Framework (TRF)
  • Objective & Scope
  • Topic Areas
  • Document Structure
  • Case Studies

Workgroups
  • Membership & Leadership
  • Charge

• Timelines
  • Overall Project Timeline
  • Progress-to-Date

• Q & A
1) Poor experimental design and data quality (early days) = bad reputation for omics technologies
2) Lack of accepted quality control standards and data quality assessment tools
3) Lack of availability of metadata necessary for interpretation and regulatory application
4) Lack of transparency, public availability and best practices/standards for data processing methods
5) Variances in methods and prior knowledge used to analyse and interpret genomics data
6) Lack of standardized reporting frameworks to ensure that all required and appropriate data, metadata, and analytical processes are available.

Buesen et al. Reg Tox Pharm. 2017
To develop frameworks for the standardisation of reporting of ‘omics data generation and analysis, to ensure that all of the information required to understand, interpret and reproduce an ‘omics experiment and its results are available.

**Purpose:** to ensure that sufficient information is available to enable an evaluation of the quality of the experimental data and interpretation, and support reproducibility.

**NOT** to stipulate the methods of data analysis or interpretation....**Rather**, provide guidance on reporting of information that fosters transparency and reproducibility.

<table>
<thead>
<tr>
<th>Project Name</th>
<th>Project Lead</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metabolomics Reporting Framework (MRF)</td>
<td>Mark Viant (U. Birmingham, UK)</td>
</tr>
<tr>
<td>Transcriptomics Reporting Framework (TRF)</td>
<td>Joshua Harrill (USEPA)</td>
</tr>
<tr>
<td></td>
<td>Carole Yauk (Health Canada)</td>
</tr>
<tr>
<td>Reference Baseline Analysis (RBA)</td>
<td>Tim Gant (PHE, UK)</td>
</tr>
</tbody>
</table>
OBJECTIVE: Development of a Transcriptomics Reporting Framework (TRF) for processing of ‘omics data that will facilitate acceptance of transcriptomics studies in a regulatory setting.

WORKING GROUP CHARGE: The TRF working group is tasked with determining what information should be captured by the TRF to support interpretation and computational reproducibility of ‘omics experiments by members of the regulatory community. Such information will also be of value to researchers in academia and industry.

SCOPE: The transcriptomics reporting framework (TRF) is a tool for documenting the details of laboratory-based toxicology studies that utilize a transcriptomics technology: i.e. an assay that measures the abundance of many transcripts simultaneously and that provides highly multiplexed outputs. The TRF is appropriate for use in documenting experiments involving the use of either in vivo or in vitro laboratory models. The information captured by the TRF should be of sufficient detail for other researchers to replicate all aspects of the transcriptomics experiment including administration of chemicals, sample processing, raw data collection and computational methods used to generate processed data. The TRF is designed to be coupled with downstream analysis reporting modules (DARMs) that detail the steps and resources necessary to reproduce a computational analysis of the processed data. Specific DARMs are coupled to the TRF based on the researcher’s specific use case.
EXPERIMENT:

• The experiment should be described in sufficient detail that would allow another researcher to replicate the experiment.
• Adapted from existing sources
• Information in this section is independent of ‘omics platform

PROCESSING AND ANALYSIS OF ‘OMICS DATA:

• The transcriptomics technology, sample processing procedures, methods used to collect raw data and methods used to generate processed data.
• Described in Gant et al. (2017).
• Information in this section is dependent on ‘omics platform

DOWNSTREAM ANALYSIS REPORTING MODULES [DARMs]

• Detail the steps and resources necessary to reproduce a computational analysis of the processed data.
Transcriptomics Reporting Framework (TRF)
- Several technology-specific documents
- Redundancy in sections i-iv across documents

Downstream Analysis Reporting Modules:
- Separate documents with reporting structure for more complex analyses.
- To be mixed and matched with TRF based on use case.

I've completed an ‘omics experiment!
And I think it may be useful to regulators...
How do I share these findings effectively?

Review what you did...
Pick your path...
Fill in the TRF!
**Draft TRF Outline**

**i. ABSTRACT**

**ii. INTRODUCTION**
  I. Purpose / Aims
  II. Background
  III. Scope
  IV. Related ‘Oomics Standard Projects

**iii. DEFINITIONS / ABBREVIATIONS**

**iv. EXPERIMENT**
  I. Study Rationale
  II. Study Design
  III. Subject / Test System Characteristics
  IV. Test Article
  V. Treatment Conditions
  VI. Study Exit
  VII. Sample Collection & Pre-processing
  VIII. Sample Identification Codes
  IX. Supporting Data Streams

**iv. PROCESSING OF ‘OMICS DATA**
  I. Technology
  II. Sample Processing
  III. Transcriptomics Study Design
  IV. Specification of Raw Data
  V. Data Normalization
  VI. Data filtering
  VII. Identification and Removal of Low Quality or Outlying Datasets

- **Stylistic alignment:**
  - Previous OECD guidance in the biological sciences (where applicable)
  - Metabolomics Reporting Framework (MRF) – In Progress

- **Reporting Format**
  - Narrative text followed by Reporting Fields

- **Consistent vocabulary across modules**

- **Database compatibility (?)**
i. ABSTRACT

ii. INTRODUCTION
   I. Purpose / Aims
   II. Background
   III. Scope
   IV. Related ‘Omics Standard Projects

iii. (TABLE OF) DEFINITIONS / ABBREVIATIONS

These sections will be drafted by the leadership team and sent to the entire project group for comment.
iv. EXPERIMENT

I. Study Rationale
II. Study Design
III. Subject / Test System Characteristics
IV. Test Article
V. Treatment Conditions
VI. Study Exit
VII. Sample Collection & Pre-processing
VIII. Sample Identification Codes
IX. Supporting Data Streams

Content is technology independent

Section to be drafted by a section workgroup under guidance of a section leader

Content leverages previously existing works
Experiment Module, Existing Resources

CEBS

MIAME

ToxRTool

SOAR
iv. PROCESSING AND ANALYSIS OF ‘OMICS DATA

I. Technology
II. Sample Processing **
III. Transcriptomics Study Design
IV. Specification of Raw Data **
V. Data Normalization **
VI. Data filtering
VII. Identification and Removal of Low Quality or Outlying Datsets **

Content is platform-specific
Sections to be drafted by a section workgroup under guidance of a section leader

** Emphasis on the use and description of Quality Control procedures / samples / performance metrics.
Downstream Analysis Reporting Modules (DARMs)

- Originally conceived by project leadership as a set of reporting templates complementary to the TRF.
- Originally thought to be beyond the scope of current TRF project (i.e. follow-up work).
- **DOES** prompt user to list all components of an analysis necessary to computationally reproduce results
- **DOES NOT** tell the user which method, or iteration of a method, they should be using.

- Identification of Differentially Expressed Genes (DEGs) will be piloted as the first DARM.
Round Robin Case Study

Objectives: Evaluate the utility of the TRF in fostering reproducibility of ‘omics data analysis by different research groups.

<table>
<thead>
<tr>
<th>Step 1.</th>
<th>Identify three (or more) analysis teams from various organizations.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 2.</td>
<td>Coordinate with the leadership team to identify an existing dataset from each team</td>
</tr>
<tr>
<td>Step 3.</td>
<td>Ask each team to: 1) Analyze their data &amp; determine DEGs (no other instructions or restrictions). 2) Report DEGs and 3) Fill out the TRF describing what they did</td>
</tr>
<tr>
<td>Step 4.</td>
<td>Provide raw data and completed TRFs (blinded, sans DEG list) to other analysis teams</td>
</tr>
<tr>
<td>Step 5.</td>
<td>Ask teams to: 1) Try and reproduce the analysis described in the TRF 2) Report DEGs to leadership team 3) Identify areas in the completed TRFs which were unclear</td>
</tr>
<tr>
<td>Step 6.</td>
<td>Leadership team assesses concordance of DEG call results and report results back to analyses teams.</td>
</tr>
<tr>
<td>Step 7.</td>
<td>Refine TRF (if necessary)</td>
</tr>
</tbody>
</table>
**Section Workgroups**

Each workgroup will consist of the following:

<table>
<thead>
<tr>
<th>Title</th>
<th>Identity</th>
<th>Roles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Section Leads</td>
<td>Raffaella Corvi [ JRC ]</td>
<td>Coordinate workgroup activities</td>
</tr>
<tr>
<td>Microarray</td>
<td>Vikrant Vijay [ NCTR ]</td>
<td>Maintain draft of section</td>
</tr>
<tr>
<td>RNA-Seq</td>
<td>Florian Caiment [ Maastricht ]</td>
<td>Manage timelines for deliverables</td>
</tr>
<tr>
<td>q-PCR array</td>
<td>Jason O'Brien [ ECCC ]</td>
<td></td>
</tr>
<tr>
<td>TempO-Seq</td>
<td>Scott Auerbach [ NTP ]</td>
<td></td>
</tr>
<tr>
<td>DARM.1 [DEG]</td>
<td>Lyle Burgoon [ ERDC ]</td>
<td></td>
</tr>
</tbody>
</table>

**Workgroup Members**

*See Next Slide*

Contribute text and content for sections

**“Floating” Facilitators**

Joshua Harrill [ USEPA ]

Carole Yauk [ Health Canada ]

Ensure consistency and cross-talk with other workgroups.

Monitor progress in accordance with project timeline

Foster discussion.

**OECD Secretariat**

Magda Sachana

Project administration / OECD liaison

All members of the TRF workgroup will have the opportunity to comment on each section.

Project group leads (Harrill & Yauk) will integrate sections into the final document.
The section workgroups are tasked with:

- Determining what information the user may list under each heading.

- Identify gaps (if any) that need to be added to the TRF structure.

- Determine the level of descriptive detail that is appropriate for each section

- “Beta test” the section using a couple of examples.

- Don’t forget to use existing resources if available!
## Project Timeline

<table>
<thead>
<tr>
<th>Date</th>
<th>Milestone</th>
</tr>
</thead>
<tbody>
<tr>
<td>April, 2018</td>
<td>Kickoff teleconference / recruiting for workgroups</td>
</tr>
<tr>
<td>May – June, 2018</td>
<td>Begin work on Introduction, Experiment, Microarray and DARM.1 modules</td>
</tr>
<tr>
<td>June, 2018</td>
<td>OECD WPHA &amp; EAGMST Meeting – Project update (presentation)</td>
</tr>
<tr>
<td>Dec, 2018</td>
<td>First drafts of Introduction, Experiment and Microarray sections due OECD Winter Meeting</td>
</tr>
<tr>
<td>June, 2019</td>
<td>Near Final Draft of Introduction, Experiment and Microarray sections Kickoff of Round Robin Case Study for Microarray First drafts of RNA-Seq, PCR array, TempO-Seq due OECD Spring Meeting</td>
</tr>
<tr>
<td>Dec, 2019</td>
<td>Final document(s) – project completion OECD Winter Meeting</td>
</tr>
</tbody>
</table>
# Progress To Date

<table>
<thead>
<tr>
<th>Date</th>
<th>Milestone</th>
</tr>
</thead>
</table>
| February-April, 2018| Project leadership planning calls  
Drafting and circulation of TRF outline document |
| April, 2018         | Kick-off teleconference with entire TRF working group  
• Solicitation of comment on TRF outline  
• Recruiting for section working groups  
• Addition of industry members |
| May, 2018           | Second teleconference with entire TRF working group  
• Presentation of scoping statement  
• Follow-up on discussion points on document content / structure  
• Alignment of TRF with OECD Harmonized Templates (Alberto Martin, EFSA) |
| June, 2018          | Kickoff TC for Experiment working group  
Kickoff TC for Microarray working group |
| July – Nov, 2018    | Kickoff TC for DARM.1 Working Group  
Drafting of Experiment, Microarray and DARM.1 sections  
Monthly TC with each active working group. |
Acknowledgements

• Leadership Team
  • Carole Yauk
  • Tim Gant
  • Magda Sachana

• OECD TRF Working Group Members

Questions?