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

Update on Development of Guidance Document(s) for Consistent Reporting of 'Omics Data From Various Sources

Transcriptomics Reporting Framework (TRF)

WPHA Meeting, Boulogne FR
June 17th, 2019

OECD
BETTER POLICIES FOR BETTER LIVES

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.*

Project Description

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.

Project Name	Project Lead
Metabolomics Reporting Framework (MRF)	Mark Viant (U. Birmingham, UK)
Transcriptomics Reporting Framework (TRF)	Joshua Harrill (USEPA) Carole Yauk (Health Canada)
Reference Baseline Analysis (RBA)	Tim Gant (PHE, UK)

TRF Objective, Working Group Charge & Scope

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.

TRF Document, Major Topic Areas

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
- Includes relevant fields from relevant OECD harmonized reporting templates (OHT 201)

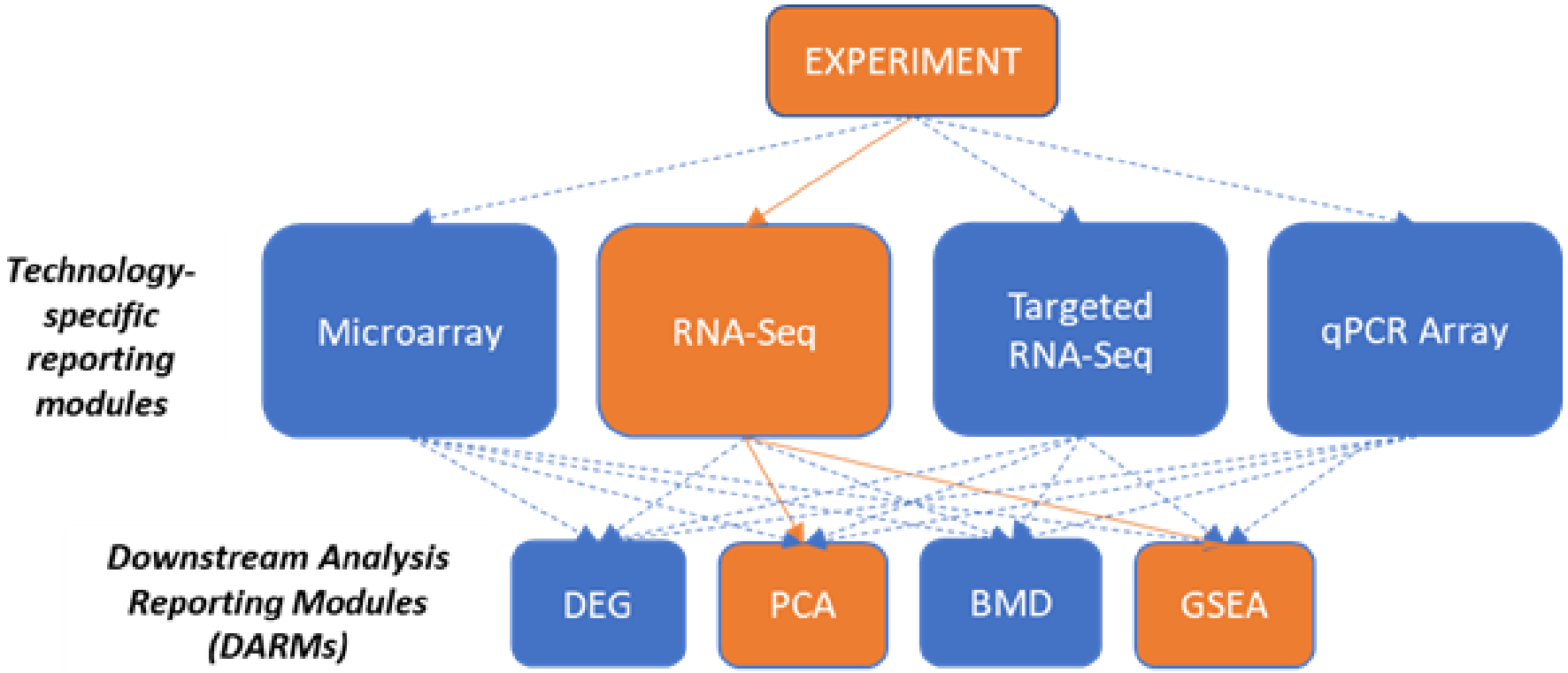
PROCESSING AND ANALYSIS OF ‘OMICS DATA:

- 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.

Modular Structure of Transcriptomics Reporting Framework



- To input information into the TRF, a researcher selects reporting modules relevant to the technology platform and computational analyses used to conduct a study.
- Report the information that would be required by an end-user to fully comprehend and replicate the analyses.

Format of the TRF

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

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

DARM.1 (DEGs)

I. Statistical Software
II. Contrasts for DEG Identification
III. Assay Experimental Design
IV. Statistical Analysis
V. Outputs

REPORT:

3.1. Type and version of the platform, manufacturer's name (e.g., Affymetrix U133 Plus 2.0 Array)
3.2. The unique identifier (e.g. serial number)
3.3. Feature type (e.g. spotted oligonucleotide)
3.4. Annotation (e.g. probe IDs)
3.5. Purpose (e.g. target gene expression, quality control, etc.)
3.6. Composition (i.e. oligo sequence, ligated product sequence)
3.7. Control console operating system
3.8. Any other relevant information

- The TRF provides narrative descriptions and basic background info for each reporting field (consistent with MERIT and MRF)
- To ease reporting efforts, a tabular companion document to the TRF (i.e. spreadsheet) is under development.
- Will indicate MANDATORY and OPTIONAL reporting fields as determined by the TRF working groups.

Section Workgroups

Each workgroup will consist of the following:

Title	Identity	Roles
Section Leads	**Experiment** **Raffaella Corvi [JRC]** **Microarray** **Vikrant Vijay [NCTR]** RNA-Seq Florian Caiment [Maastricht] q-PCR array Jason O'Brien [ECCC] Targeted RNA-Seq Scott Auerbach [NTP] **DARM.1 [DEG]** **Lyle Burgoon [ERDC]**	Coordinate workgroup activities Maintain draft of section Manage timelines for deliverables
Workgroup Members (n = 2 - ?)	Various gov't, academic and industry scientists	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.

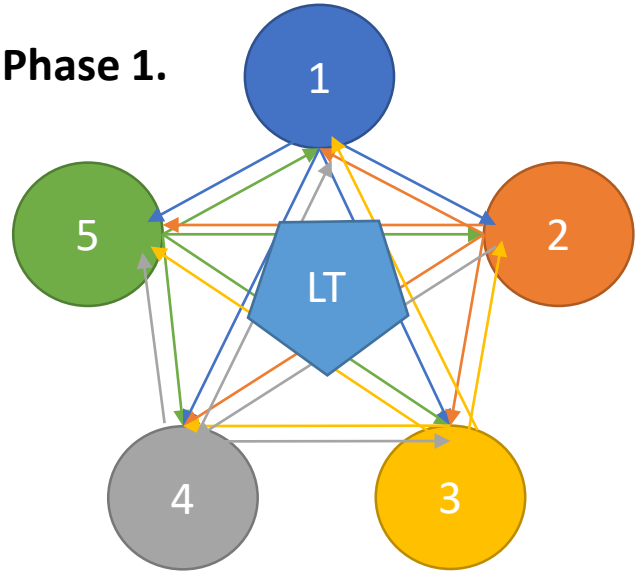
Project group leads (Harrill & Yauk) will integrate sections into the final document.

Round Robin Case Study

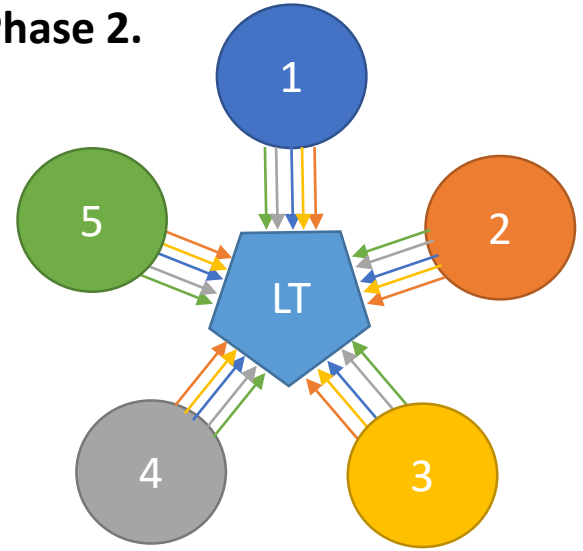
Objectives: Evaluate the utility of the TRF in fostering reproducibility of 'omics data analysis by different research groups.

Phase 1	**Step 1.**	Identify multiple analysis teams across various organizations.
	Step 2.	Coordinate with the leadership team to identify an existing microarray dataset from each team
	Step 3.	Ask each team to: 1) Analyze their data & determine DEGs (no other instructions or restrictions). 2) Report DEGs and 3) Fill out the TRF describing what they did
Phase 2	**Step 4.**	Provide raw data and completed TRFs (blinded, sans DEG list) to the other analysis teams
	Step 5.	Ask teams to: 1) Try and reproduce the analysis described in each TRF 2) Report DEGs to leadership team 3) Identify areas in the completed TRFs which were unclear and may have lead to inconsistencies.
	Step 6.	Leadership team assesses concordance of DEG call results and report results back to analyses teams.
	Step 7.	Refine TRF (if necessary)

Phase 1.



Phase 2.



Project Timeline (Revised)

Date	Milestone
April, 2018	Kickoff teleconference / recruiting for workgroups
May – June, 2018	Begin work on Introduction, Experiment, Microarray and DARM.1 modules
June, 2018	OECD WPHA & EAGMST Meeting – Project update (presentation)
Dec, 2018	First drafts of Introduction, Experiment and Microarray sections due OECD Winter Meeting
June, 2019	Near Final Draft of Introduction, Experiment and Microarray sections Submit TRF drafts to EAGMST TRF Working Group members for review OECD Summer Meeting
July-Aug, 2019	Initiate RNA-Seq, targeted RNA-Seq and PCR array working groups Recruit for working members for additional DARM modules (BMD, PCA, etc.) Kickoff of Round Robin Case Study for Microarray
Dec, 2019	Near final drafts of RNA-Seq, targeted RNA-Seq and PCR array documents due OECD Winter Meeting

Project will likely require extension to the summer of 2020.

Acknowledgements

OECD TRF Contributors

EXPERIMENT

Rafaella Corvi

Gina Hilton

Ian Cotgreave

Ed Perkins

Mirjam Luitjen

John Colbourne

Tao Chen

Monica Vaccari

Sofia Batista-Leite

Ivana Campia

MICROARRAY

Vikrant Vijay

Tim Gant

Laura Gribaldo

Andrew Williams

DARM.1

Lyle Burgoon

Andrew Williams

Natalia Reyero

Andy White

OHT Mapping

Alberta Martin-Aparicio

Raffaela Corvi

Sofia Batista-Leite

Ivana Campia

TRF Leadership Team

Carole Yauk

Tim Gant

Magda Sachana

Questions?