

## **Appendix C: AIRR-C Sub-committee (SC) and Working Group (WG) Reporting Template**

*Instructions.* This form is to be used for AIRR-C Meeting updates. The form should be completed by SC or WG (Co)-leaders, with input from other SC or WG members, and submitted to the Chair of the AIRR-C Executive SC at least one week prior to the AIRR-C Meeting.

### **Current**

**Date of this report: April 8, 2022**

**SC/WG Name: Common Repository Working Group**

**SC/WG Co-leaders: Brian Corrie, Artur Rocha**

### **SC/WG Active Members (list): Attended Meetings**

Brian Corrie, Artur Rocha, Scott Christley, Christian Busse, George Blanck, Lindsay Cowell, Eric Waltari, Felix Breden, Katharina Imkeller, Ulrik Stervbo, Veronique Giudicelli, Marco Amaro Oliveira, David Klatzmann, Adrien Six, Enkelejda Miho, Corey Watson

### **Purpose:**

To promote and facilitate deposit, access, and sharing and reuse of IG and TCR AIRR-seq datasets through the creation of common repositories that enable:

- Standardized queries of processed AIRR-seq data
- Re-analysis of raw and processed AIRR-seq data utilizing repository analysis tools
- Download of raw and processed AIRR-seq data for offline re-analysis

### **Goals:**

- To develop and promote the AIRR Data Commons (ADC) as an international resource for storing, finding, accessing, and reusing AIRR-seq and related data.
- To develop and promote the ADC API, a web-based query API for querying repositories in the ADC.

### **Long-term vision and how WG products integrate with the AIRR-C mission:**

- The ADC relies on and promotes adoption of the standards established in the Standards Working Group.
- The ADC is a central resource for the AIRR Community and beyond for sharing, accessing, and reusing AIRR-seq data.
- We anticipate the ADC to expand both in size (more data and more repositories) and in scope (broader types of data – e.g. single-cell gene expression and receptor binding characteristics) in the future.

## Products (if any):

### AIRR Data Commons

- Although not provided by the CRWG (repositories are provided by community members) the AIRR Data Commons is the main product produced by the CRWG.
- The ADC API is the main output developed by the CRWG, enabling users to query the ADC in a consistent and interoperable way.

## Resources (if any):

- AIRR documentation web site
- AIRR Data Commons – Network of repositories

## Progress report on current purpose, goals, products and resources:

- Continual growth in data – added 1B annotations
  - 5.1 billion annotations, 7128 repertoires, 7 repositories, 84 studies.
- Continual growth in use – iReceptor Gateway as a proxy for ADC use
  - 10 – 20 new users per month
  - 489 new users since COVID-19 data made available
- FAIR AIRR-seq data citations – ADC data reuse
  - 12 COVID-19 data related citations
    - C. Schultheiß et al., “Next-Generation Sequencing of T and B Cell Receptor Repertoires from COVID-19 Patients Showed Signatures Associated with Severity of Disease,” *Immunity*, vol. 53, no. 2, pp. 442-455.e4, Aug. 2020.
    - J. D. Galson et al., “Deep Sequencing of B Cell Receptor Repertoires From COVID-19 Patients Reveals Strong Convergent Immune Signatures,” *Front. Immunol.*, vol. 11, p. 3283, Dec. 2020.
    - D. Simnica et al., “Landscape of T-cell repertoires with public COVID-19-associated T-cell receptors in pre-pandemic risk cohorts,” *Clin. Transl. Immunol.*, vol. 10, no. 9, p. e1340, Jan. 2021.
    - R. R. Goel et al., “mRNA vaccines induce durable immune memory to SARS-CoV-2 and variants of concern,” *Science* (80-. ), Dec. 2021.
    - A. Mohamad-Gabriel Alameh et al., “Lipid nanoparticles enhance the efficacy of mRNA and protein subunit vaccines by inducing robust T follicular helper cell and humoral responses,” *Immunity*, vol. 0, no. 0, pp. 1–16, Nov. 2021.
    - P. Meysman, A. Postovskaya, N. De Neuter, B. Ogunjimi, and K. Laukens, “Tracking SARS-CoV-2 T cells with epitope-T-cell receptor recognition models,” *bioRxiv*, p. 2020.09.09.289355, Sep. 2020.
    - M. Kuchroo et al., “Multiscale PHATE Exploration of SARS-CoV-2 Data Reveals Multimodal Signatures of Disease,” *bioRxiv*, p. 2020.11.15.383661, Nov. 2020.
    - M. Heming et al., “Neurological Manifestations of COVID-19

Feature T Cell Exhaustion and Dedifferentiated Monocytes in Cerebrospinal Fluid,” Immunity, vol. 54, no. 1, pp. 164-175.e6, Jan. 2021.

- R. A. Porritt et al., “HLA class I–associated expansion of TRBV11-2 T cells in multisystem inflammatory syndrome in children,” J. Clin. Invest., vol. 131, no. 10, May 2021.
- A. J. Schmitz et al., “A vaccine-induced public antibody protects against SARS-CoV-2 and emerging variants,” Immunity, vol. 54, no. 9, pp. 2159-2166.e6, Sep. 2021.
- J. Y. Humrich, J. P. Bernardes, R. J. Ludwig, D. Klatzmann, and A. Scheffold, “Phenotyping of Adaptive Immune Responses in Inflammatory Diseases,” Front. Immunol., vol. 11, p. 3010, Nov. 2020.
- F. P. Caruso, G. Scala, L. Cerulo, and M. Ceccarelli, “A review of COVID-19 biomarkers and drug targets: resources and tools,” Brief. Bioinform., vol. 22, no. 2, pp. 701–713, Mar. 2021.
- Repository growth
  - 7 repositories in total currently, four “recent” repositories
    - iReceptor Public Archive – iReceptor/SFU
    - iReceptor COVID-19 – iReceptor/SFU
    - VDJServer – VDJServer/UTSW
    - VDJBase (Israel) – Gur Yaari/Bar Ilan
    - sciReptor (Germany) – Christian Busse/DKFZ
    - NICD (South Africa) – Cathrine Scheepers/NICD
    - UKM (Germany) – Nicholas Schwab/Muenster
- ADC API extensions
  - Existing: /repertoire, /rearrangement
  - Version 1.4: /clone, /cell, /expression
  - Longer term: /receptor, /stats, /germline

### **Proposed plans for the coming interval:**

#### **Purpose:**

To promote and facilitate deposit, access, and sharing and reuse of IG and TCR AIRR-seq datasets through the creation of common repositories that enable:

- Standardized queries of processed AIRR-seq data
- Re-analysis of raw and processed AIRR-seq data utilizing repository analysis tools
- Download of raw and processed AIRR-seq data for offline re-analysis

#### **Goals:**

- Recommendations: Continue to adapt as required
- Outreach: to community to expand the AIRR Data Commons
- Standards: AIRR Standard and ADC API evolution
- Registry: Programmatic registry of ADC repositories that is searchable

**Products (if any):**

- AIRR Data Commons
- ADC API
- ADC Registry

**Resources (if any):**

- AIRR documentation web site
- AIRR Data Commons – Network of repositories

**Long-term vision and how WG or SC products integrate with the AIRR-C mission:**

Same as above...

**Proposed SC/WG Co-leaders: Brian Corrie, Artur Rocha**