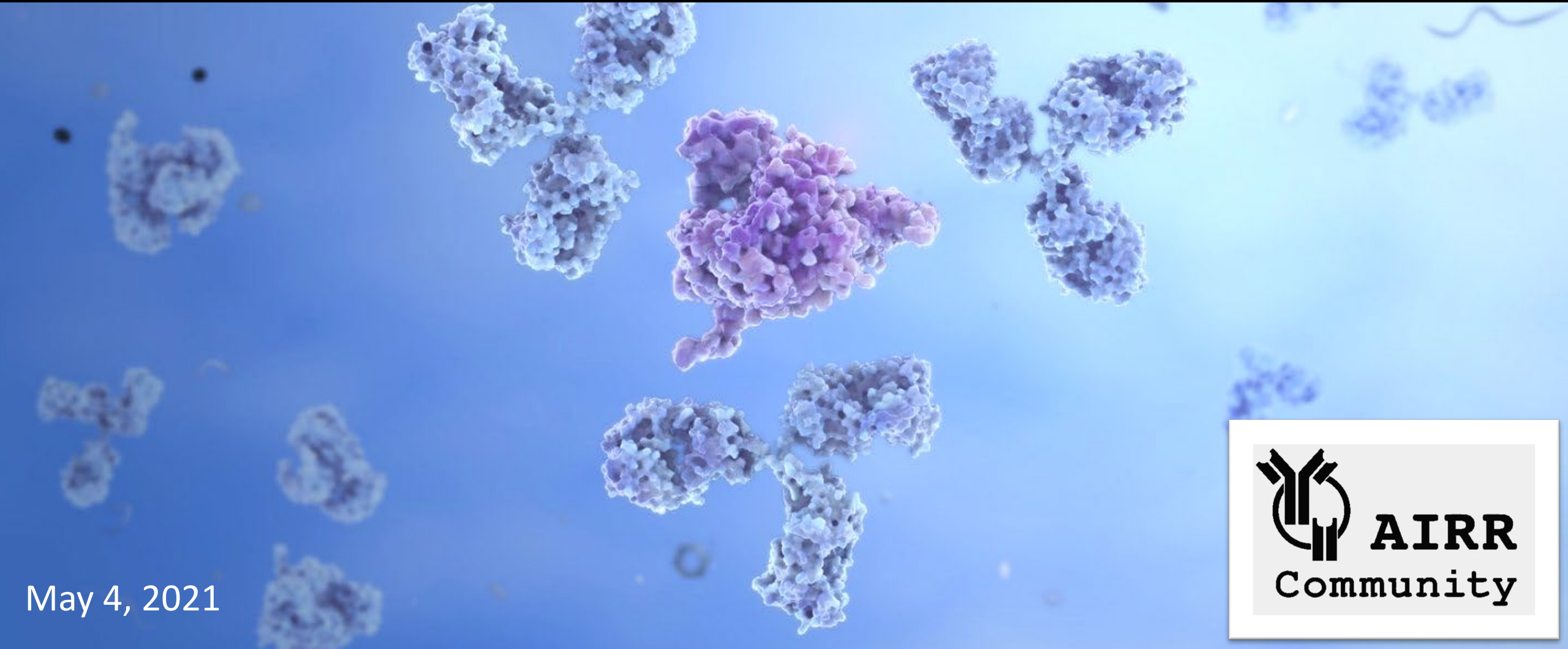


The AIRR Data Commons: 4 billion reasons to store, analyze and share antibody/B-cell and T-cell receptor repertoire data



May 4, 2021



Today's speakers

- Dr. Felix Breden
 - Executive Director, iReceptor/iReceptor Plus, Simon Fraser University
 - Member and former chair, AIRR Community Executive
- Dr. Brian Corrie
 - Technical Director, iReceptor/iReceptor Plus, Simon Fraser University
 - Co-chair, AIRR Common Repository Working Group
- Dr. Kira Neller
 - Data curator/bioinformatician, iReceptor/iReceptor Plus, Simon Fraser University
 - Member, AIRR Standards Working Group
- Dr. Scott Christley
 - Computational Biologist, UT Southwestern Medical School
 - Project Leader, VDJSERVER
 - Co-chair, AIRR Standards Working Group

Today's Webinar

Part I: Adaptive Immune System and the AIRR Community

Part II: The AIRR Data Commons

Part III: AIRR Data Commons Tool Demos

- *iReceptor Gateway*
- *AIRR-seq Data Curation*
- *VDJServer Analysis Portal*

Part I: Adaptive Immune System and the AIRR Community

Adaptive Immune System

- Comprises antibodies/B-cell receptors and T-cell receptors – AIRR-seq data (Adaptive Immune Receptor Repertoire)
- Incredibly variable to recognize and *remove* bacteria and viruses (including new ones, *e.g.* novel coronavirus)
- Also, must recognize and *not remove* >25,000 expressed “self” proteins
- Basis of vaccines, drugs suppressing autoimmune diseases, new cancer immunotherapies, *etc.*

Why a platform specific to AIRR-seq data?

AIRR-seq data are difficult to share and compare:

Size of data sets

AIRR-seq repertoires are highly diverse:

$\sim 10^{13}$ potential human B-cell receptors

Individual can have $\sim 10^9$ B cells and T cells:

10^6 or 10^7 receptors sequenced per sample

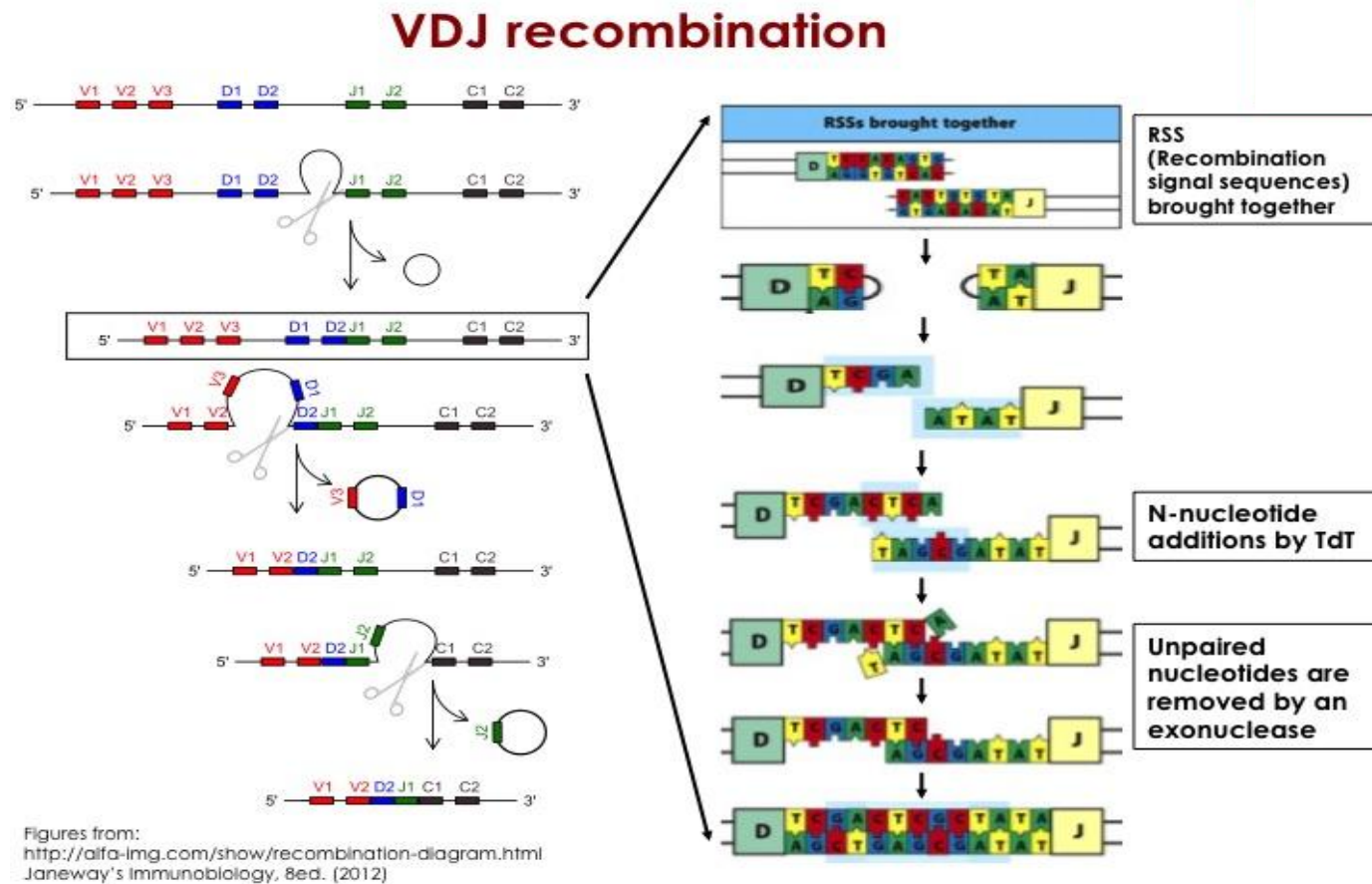
Studies produce huge data sets:

Burkitt lymphoma (Lombardo *et al.*) – 1.2B B-cell receptors

Systemic sclerosis (de Bourcy *et al.*) - 700M B-Cell receptors

Several recent cancer studies – >500M T-cell receptors

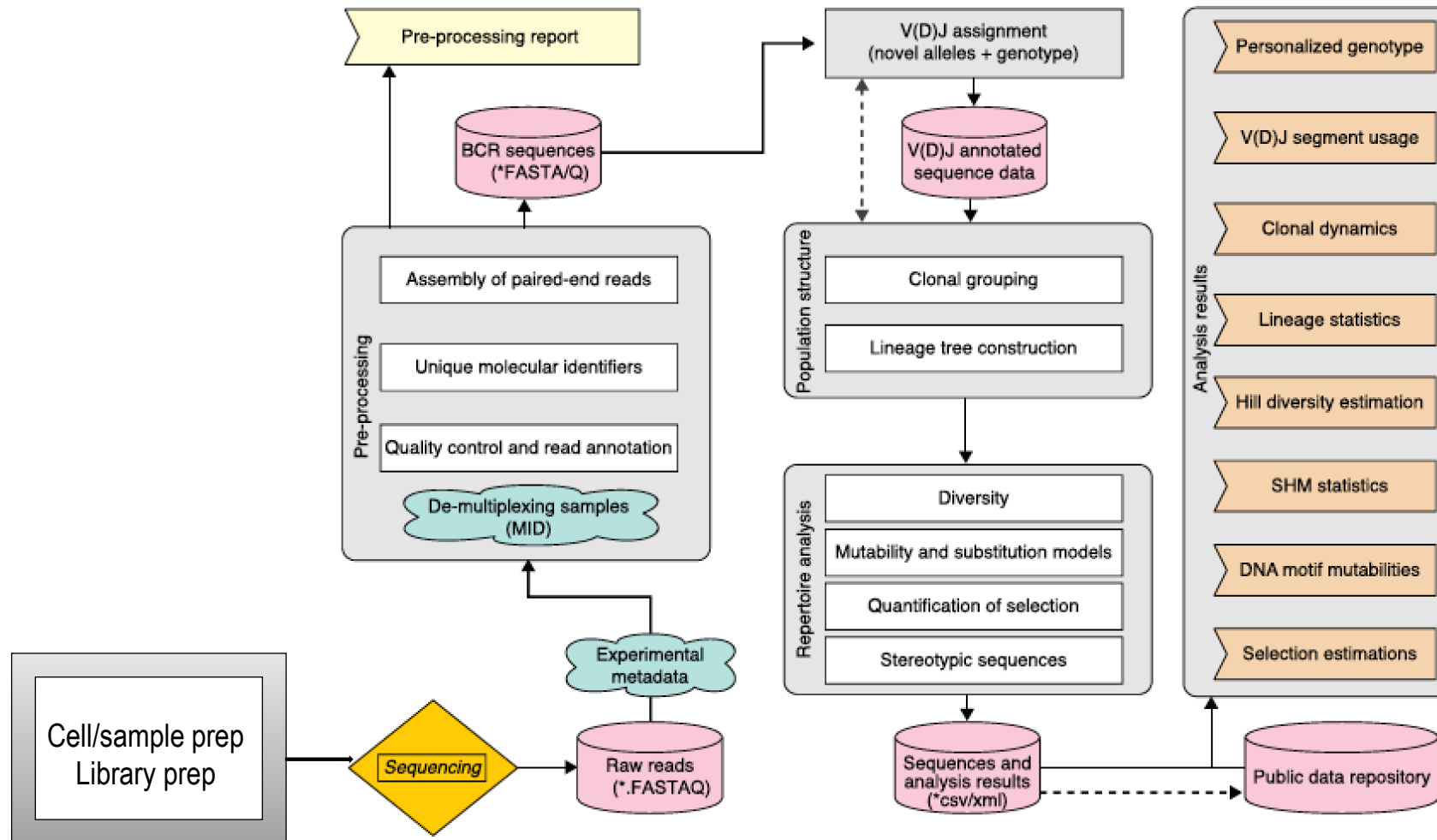
AIRR-seq data are difficult to share and compare:
VDJ recombination demands unique data base model for
annotated receptor sequences



Clones are sets of
B-cells or
T-cells descended
from ancestral cell
produced by
V(D)J
recombination

Immunoglobulin and T-cell Receptor genes are only genes in eukaryote
genome that undergo this somatic recombination

AIRR-seq data are difficult to share and compare:
Many ways for experiments to differ!



AIRR-seq data are difficult to share and compare: Complexity and confidentiality of data sets

Require data on patient demography, treatment, clinical outcome, *etc.*

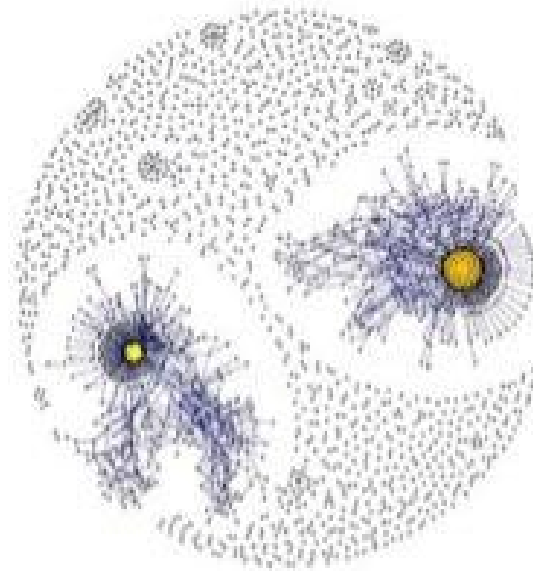
Complex analyses demand unique bioinformatic tools

AIRR-seq data could lead to personal identification?

B-cell Clonal Lineage Expansion in Health and Disease

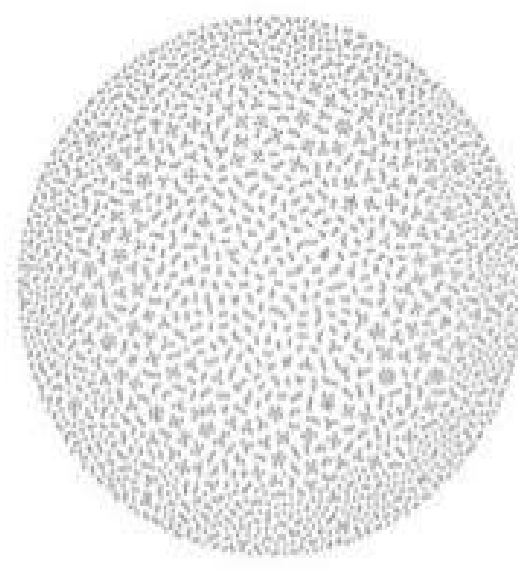
CLL sample 5

26,086 sequences total
48.0 % maximum cluster



Healthy donor 1

12,316 sequences total
0..49 % maximum cluster



Chronic Lymphocytic Leukemia (CLL) is characterized by expansion of a few dominant clones in B-cell repertoire (Bashford-Rogers *et al.* 2019)

FDA approved Adaptive Biotechnologies clonoSEQ® test for Minimal Residual Disease (MRD) based on searching for these CLL-associated, expanded clones

Patterns Observed in COVID-19 patient AIRR-seq data repository: 20 studies, 3531 repertoires, 1B sequences

- Stereotypical immune response mediated by convergent clonotypes
 - Galson et al., Montague et al., Nielsen et al., Robbiani et al., Sokal et al., Xiang et al. Zhang et al.
 - Adaptive Biotechnologies FDA approved T-Detect COVID test for prior infections
- Pre-existing immunity in unexposed individuals
 - Precursor nABs: Kreer et al., Mor et al.
 - SARS-CoV-2 T-cell epitopes cross-reactive to common viruses: Mahajan et al., Grifoni et al., Zhao et al.
- Neuro-COVID
 - 7 studies from ADC show T-cell expansion in "long haulers" with neurological defects: Hemming et al.
- Biased V-gene usage
 - IGHV3-53 and IGHV3-66 greatly increased in anti-SARS-CoV-2 antibodies: Kim et al., Mor et al., Barnes et al.
 - T-cell repertoires biased toward specific TRVB genes in severe hyper-inflammatory COVID-19 patients: Cheng et al.

Adaptive Immune Receptor Repertoire (AIRR) Community

The AIRR Community is a global, grass-roots group of immunologists, bioinformaticists, computer scientists, experts in legal, ethical and IP issues, who are developing guidelines and standards for the generation, annotation and storage of high-throughput AIRR-seq data to facilitate its use by the larger research community.

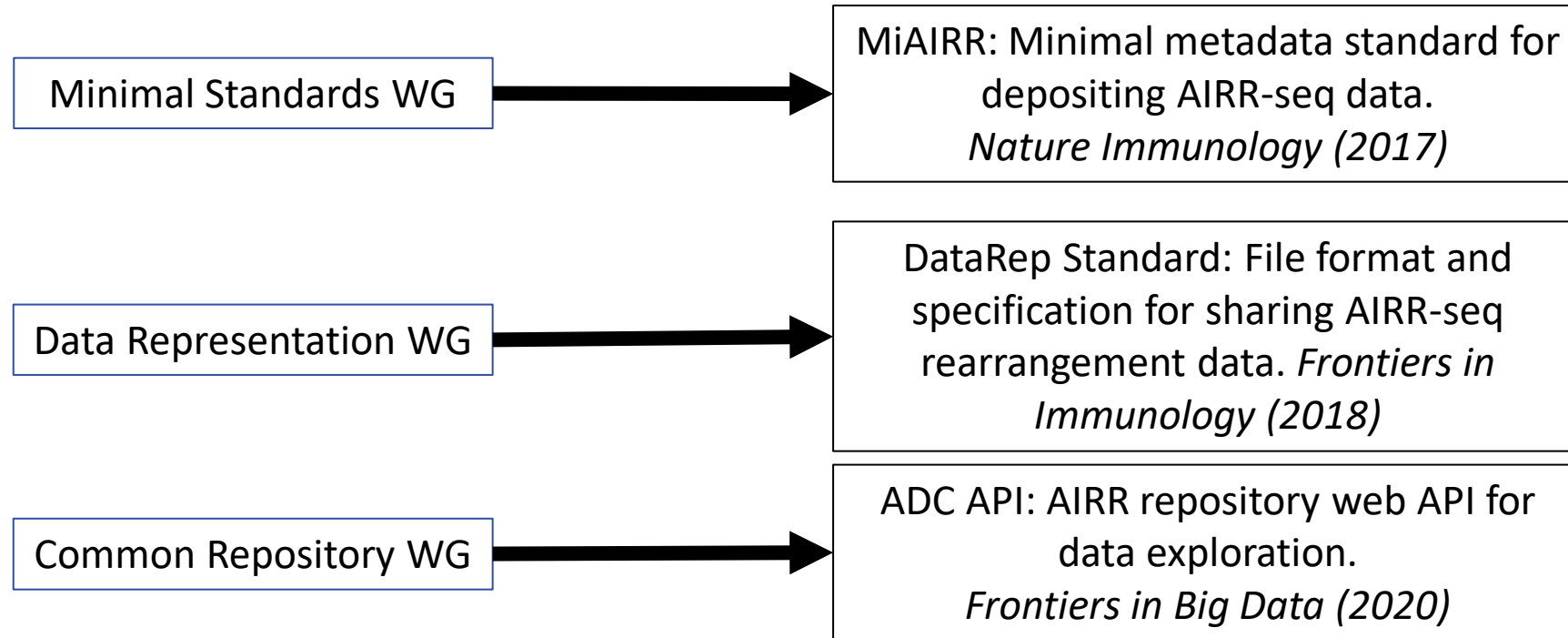
Ability to share AIRR-seq data greatly increases the value of any one data set:

- each researcher, small N, large amount of data per sample
- increase sample sizes, statistical power
- AI approaches demand huge sample sizes
- facilitate comparisons between affected/controls/multiple disease states

AIRR Community Working Groups

1. Minimal Standards for publishing or depositing AIRR-seq data (MiAIRR)
2. Data Representation – common formats for annotated data
3. Software – interoperability of analysis software
4. Common Repository – Data Commons for AIRR-seq data, following FAIR principles
5. Diagnostics – develop diagnostics and markers for disease; one goal is sustainable business model
6. Legal and Ethics – standards for human subjects
7. Biological Resources – testing molecular protocols
8. Germline Database – inference from AIRR-seq data

AIRR Community Working Groups Develop Standards



Standards (Publications) must be ratified by
full AIRR Community

Join: www.airr-community.org

Part II: The AIRR Data Commons

Making AIRR-seq data FAIR and TRUSTed

- Goal of the AIRR Community:
 - Promote the Reproducibility and Reuse of AIRR-seq data (Breden et al., 2017)

Making AIRR-seq data FAIR and TRUSTed

- Goal of the AIRR Community:
 - Promote the Reproducibility and Reuse of AIRR-seq data (Breden et al., 2017)
- Researcher needs: AIRR-seq data that is **FAIR**
 - **Find** data of interest from multiple labs/institutions
 - Combine the data of interest (**Accessible** and **Interoperable**)
 - **Reuse** combined data in new analyses to **derive new insights** (**FAI** leads to **Reuse**)

Making AIRR-seq data FAIR and TRUSTed

- Goal of the AIRR Community:
 - Promote the Reproducibility and Reuse of AIRR-seq data (Breden et al., 2017)
- Researcher needs: AIRR-seq data that is **FAIR**
 - **Find** data of interest from multiple labs/institutions
 - Combine the data of interest (**Accessible** and **Interoperable**)
 - **Reuse** combined data in new analyses to **derive new insights** (**FAI** leads to **Reuse**)
- Reproducible science requires **TRUST**ed AIRR-seq data repositories
 - Transparent, Responsible, User focused, Sustainable, and Technologically reliable
 - Community involvement is critical to gaining TRUST – The AIRR Community

The AIRR Data Commons (ADC)

- Network of distributed AIRR-seq repositories

AIRR Data Commons

International network of
distributed repositories



The AIRR Data Commons (ADC)

- Network of distributed AIRR-seq repositories

AIRR Data Commons

International network of
distributed repositories



Based on standards developed
by the AIRR Community!

The AIRR Data Commons (ADC)

- Network of distributed AIRR-seq repositories
 - 4 repositories, 60 studies, 4 billion sequences

AIRR Data Commons

International network of
distributed repositories



Based on standards developed
by the AIRR Community!

The AIRR Data Commons (ADC)

- Network of distributed AIRR-seq repositories
 - 4 repositories, 60 studies, 4 billion sequences

AIRR Data Commons

International network of
distributed repositories

Data Query



Based on standards developed
by the AIRR Community!

The AIRR Data Commons (ADC)

- Network of distributed AIRR-seq repositories
 - 4 repositories, 60 studies, 4 billion sequences

AIRR Data Commons

International network of
distributed repositories



Based on standards developed
by the AIRR Community!

The AIRR Data Commons (ADC)

- Network of distributed AIRR-seq repositories
 - 4 repositories, 60 studies, 4 billion sequences
 - MiAIRR Standard for study reporting
 - Rubelt et al. (Nat. Immunol., 2017)
 - File formats for data sharing
 - Vander Heiden et al. (Front. Immunol., 2018)
 - Queryable through ADC web API
 - Christley et al. (Front. Big Data, 2020)

AIRR Data Commons

International network of
distributed repositories



Based on standards developed
by the AIRR Community!

Making AIRR-seq data FAIR and Easy!

AIRR Data Commons

International network of
distributed repositories



Based on standards developed
by the AIRR Community!

Making AIRR-seq data FAIR and Easy!

iReceptor Scientific Gateway

Interactive web-based data
discovery, exploration, and analytics



AIRR Data Commons

International network of
distributed repositories



Based on standards developed
by the AIRR Community!

Making AIRR-seq data FAIR and Easy!

iReceptor Scientific Gateway

Interactive web-based data
discovery, exploration, and analytics



Web based portal that hides the
complexity of finding, searching
and federating data across many
repositories!

AIRR Data Commons

International network of
distributed repositories



Based on standards developed
by the AIRR Community!

Making AIRR-seq data FAIR and Easy!

iReceptor Scientific Gateway

Interactive web-based data
discovery, exploration, and analytics



Web based portal that hides the
complexity of finding, searching
and federating data across many
repositories!

AIRR Data Commons

International network of
distributed repositories



Based on standards developed
by the AIRR Community!

Data Query

Making AIRR-seq data FAIR and Easy!

iReceptor Scientific Gateway

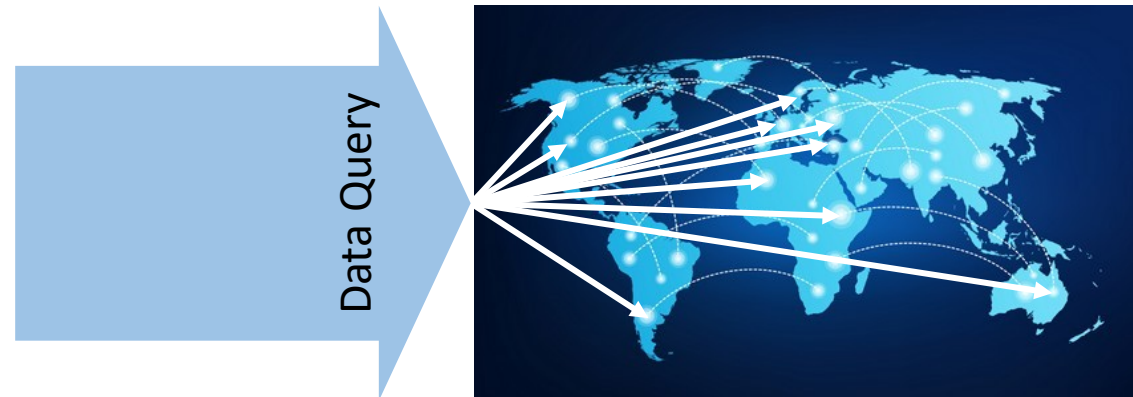
Interactive web-based data
discovery, exploration, and analytics



Web based portal that hides the
complexity of finding, searching
and federating data across many
repositories!

AIRR Data Commons

International network of
distributed repositories



Based on standards developed
by the AIRR Community!

Making AIRR-seq data FAIR and Easy!

iReceptor Scientific Gateway

Interactive web-based data
discovery, exploration, and analytics



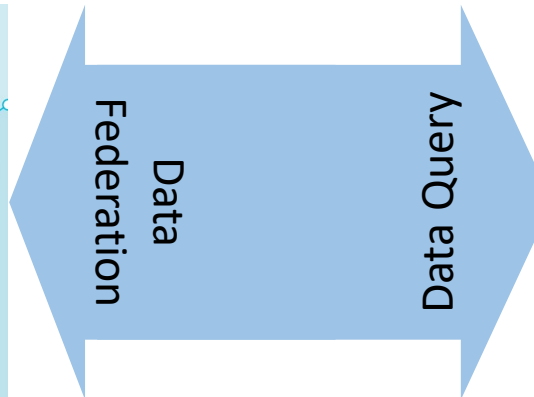
Web based portal that hides the
complexity of finding, searching
and federating data across many
repositories!

AIRR Data Commons

International network of
distributed repositories

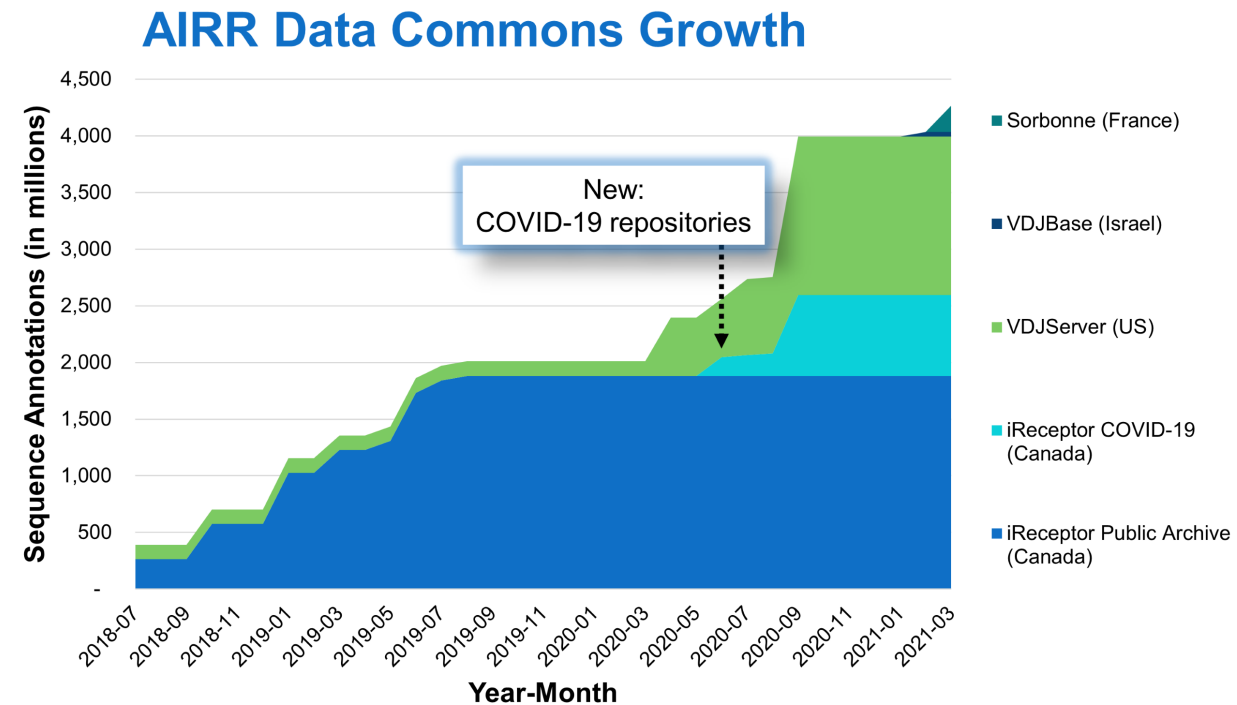


Based on standards developed
by the AIRR Community!



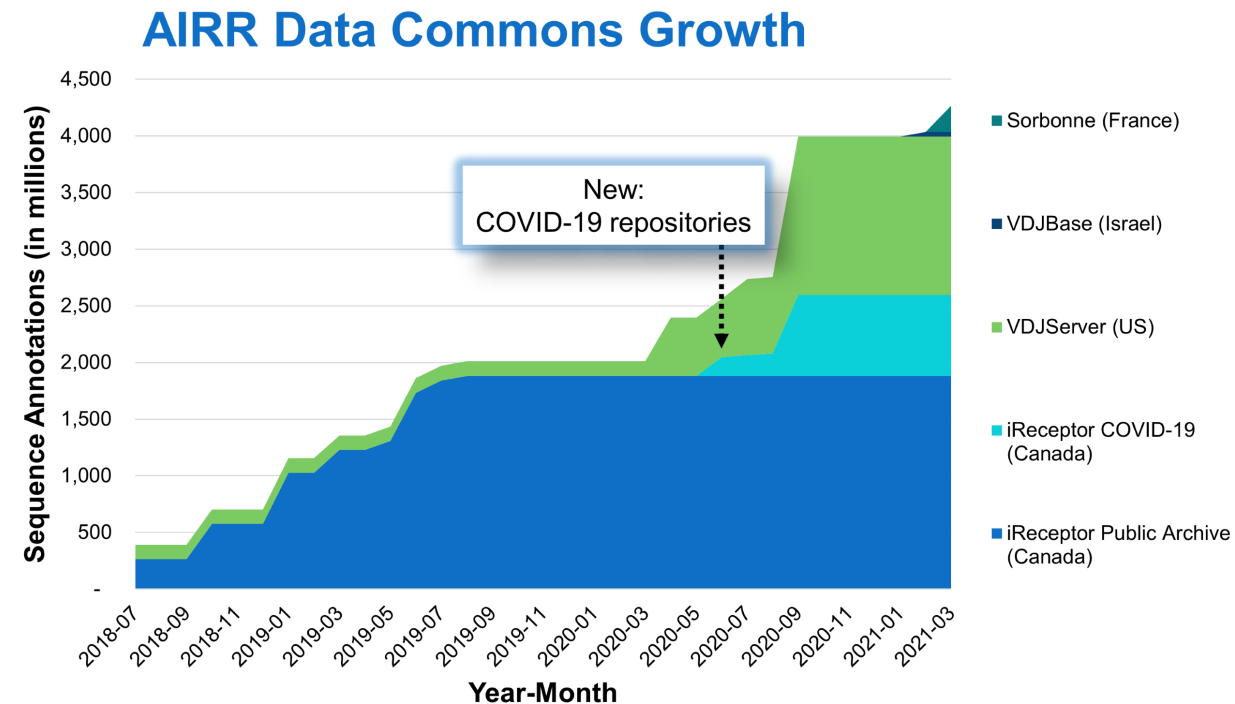
Why should I use the AIRR Data Commons?

- The ADC continues to grow...
 - We encourage everyone to share!
- 4 billion reasons to use the ADC!
 - 60 studies, 4 billion sequences



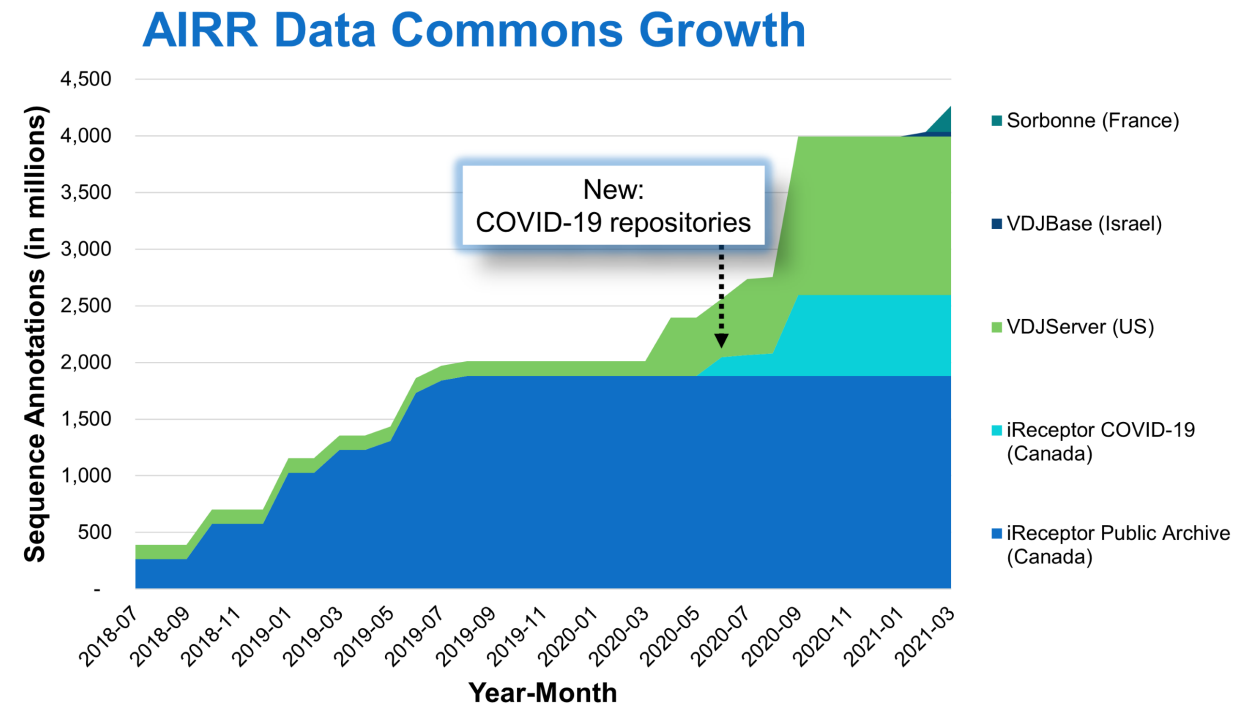
Why should I use the AIRR Data Commons?

- The ADC continues to grow...
 - We encourage everyone to share!
- 4 billion reasons to use the ADC!
 - 60 studies, 4 billion sequences
 - 15 COVID-19 studies, 1 billion sequences
 - Nielsen et al. data available June 2020
 - Preprint, May 2020, Cell Host & Microbe, Oct 2020



Why should I use the AIRR Data Commons?

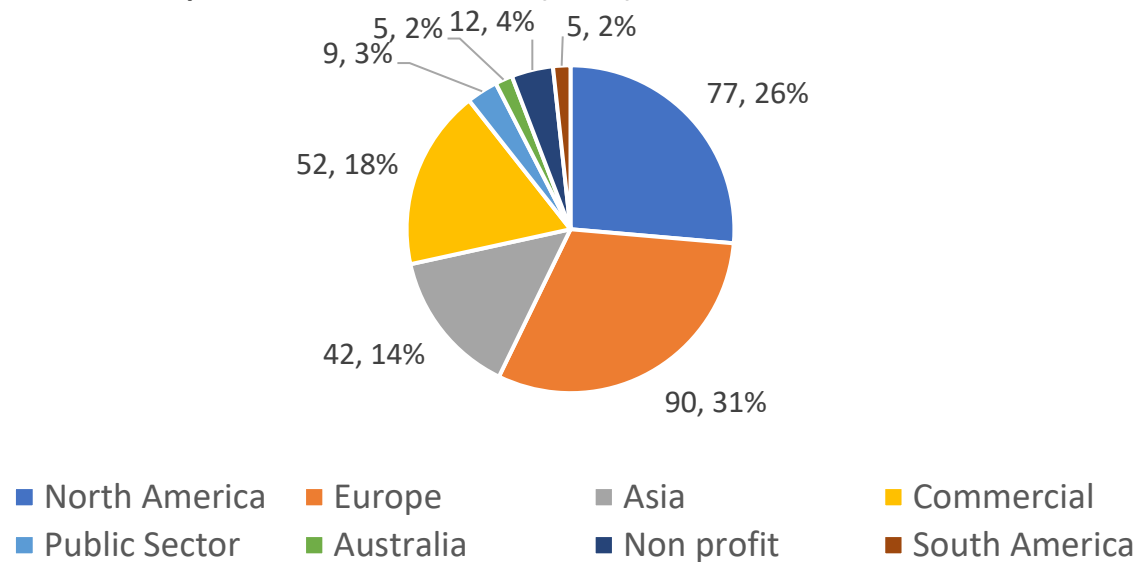
- The ADC continues to grow...
 - We encourage everyone to share!
- 4 billion reasons to use the ADC!
 - 60 studies, 4 billion sequences
 - 15 COVID-19 studies, 1 billion sequences
 - Nielsen et al. data available June 2020
 - Preprint, May 2020, Cell Host & Microbe, Oct 2020
- Data reuse is becoming common
 - Partially driven by COVID-19 data
 - Schultheiß et al. (Immunity, Aug 2020) - Data
 - Meysman et al. (bioRxiv, Sep, 2020) - Reuse
 - Heming et al. (Immunity, Jan 2021) - Reuse
 - Chang et al. (Front. Immunol., Apr 2021) - Reuse
 - Data reuse = paper citations!



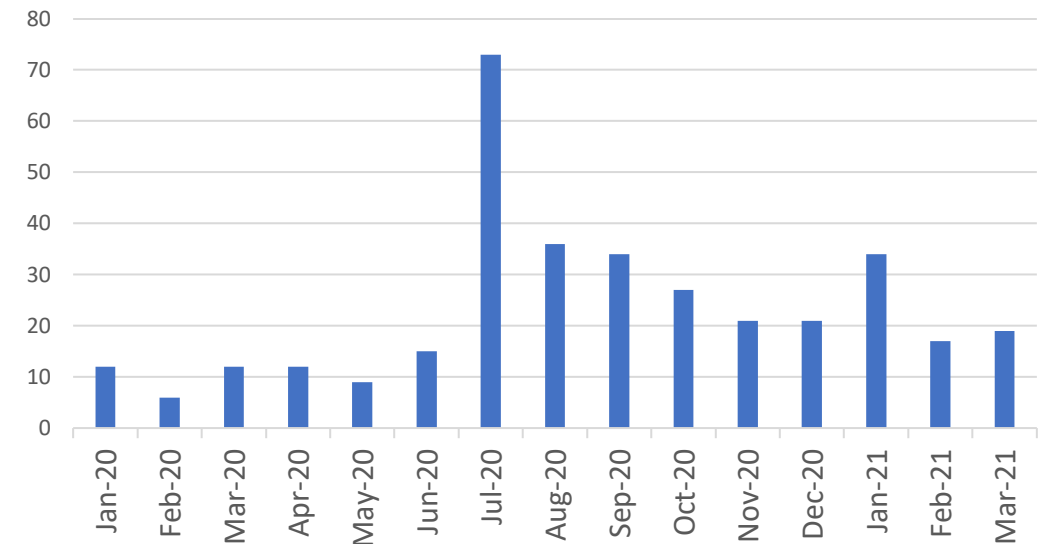
Why should I use the AIRR Data Commons?

- Because everyone is doing it...
 - Significant user increase in last 12 months
 - Critical mass of data? COVID-19?

iReceptor Users Added (297): Jun 2020 - Mar 2021




iReceptor Users Added per Month



How do I use the AIRR Data Commons?

- iReceptor Gateway - Search and ye shall find!
 - Search 60 studies for over 80 MiAIRR study/subject/sample metadata fields
 - E.g. Finding COVID-19 data...



Search ▾

1. Repertoire Metadata

2. Sequences

1. Repertoire Metadata Search

Filter by study/subject/sample and choose repertoires to analyze relevant sequence data

Filters

Full-text search ⓘ

Filter by study

+

Filter by subject

+

Filter by sample

+

More filters

+

Active filters

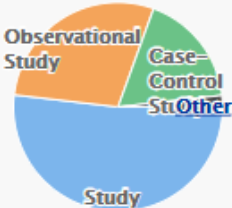
✕ Full-text search

Remove all filters

Search results statistics

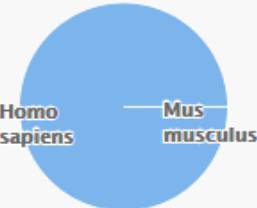
1,074,424,168 sequences (3624 repertoires) returned from [3 remote repositories](#), [15 research labs](#) and [21 studies](#).

5 STUDY TYPES



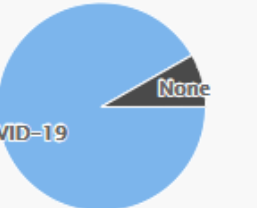
Study Type	Count
Study	~1,000,000
Case-Control Study	~100,000
Observational Study	~100,000
Other	~100,000
Other	~100,000

2 ORGANISMS



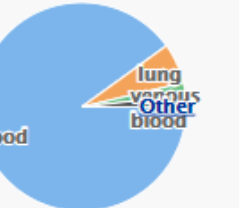
Organism	Count
Homo sapiens	~1,000,000
Mus musculus	~100,000

2 DIAGNOSES



Diagnosis	Count
COVID-19	~1,000,000
None	~100,000

6 TISSUES




Tissue	Count
blood	~1,000,000
lung	~100,000
various	~100,000
blood	~100,000
Other	~100,000
Other	~100,000

Individual Repertoires 1-120 of 3624 [Customize displayed columns](#)

How do I use the AIRR Data Commons?

- iReceptor Gateway - Search and ye shall find!
 - Search 60 studies for over 80 MiAIRR study/subject/sample metadata fields
 - E.g. Finding COVID-19 data...



Search ▾

1. Repertoire Metadata

2. Sequences

1. Repertoire Metadata Search

Filter by study/subject/sample and choose repertoires to analyze relevant sequence data

Filters

Full-text search ⓘ
covid-19

Filter by study +

Filter by subject +

Filter by sample +

More filters +

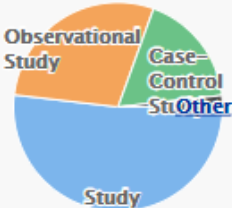
Active filters

✕ Full-text search Remove all filters

Search results statistics

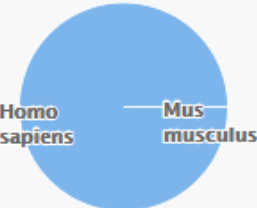
1,074,424,168 sequences (3624 repertoires) returned from [3 remote repositories](#), [15 research labs](#) and [21 studies](#).

5 STUDY TYPES



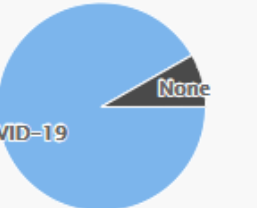
Study Type	Count
Study	1
Case-Control Study	1
Observational Study	1
Other	1
Other	1

2 ORGANISMS



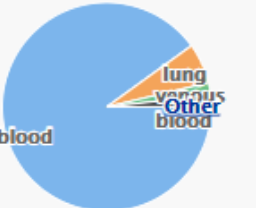
Organism	Count
Homo sapiens	1
Mus musculus	1

2 DIAGNOSES



Diagnosis	Count
COVID-19	1
None	1

6 TISSUES

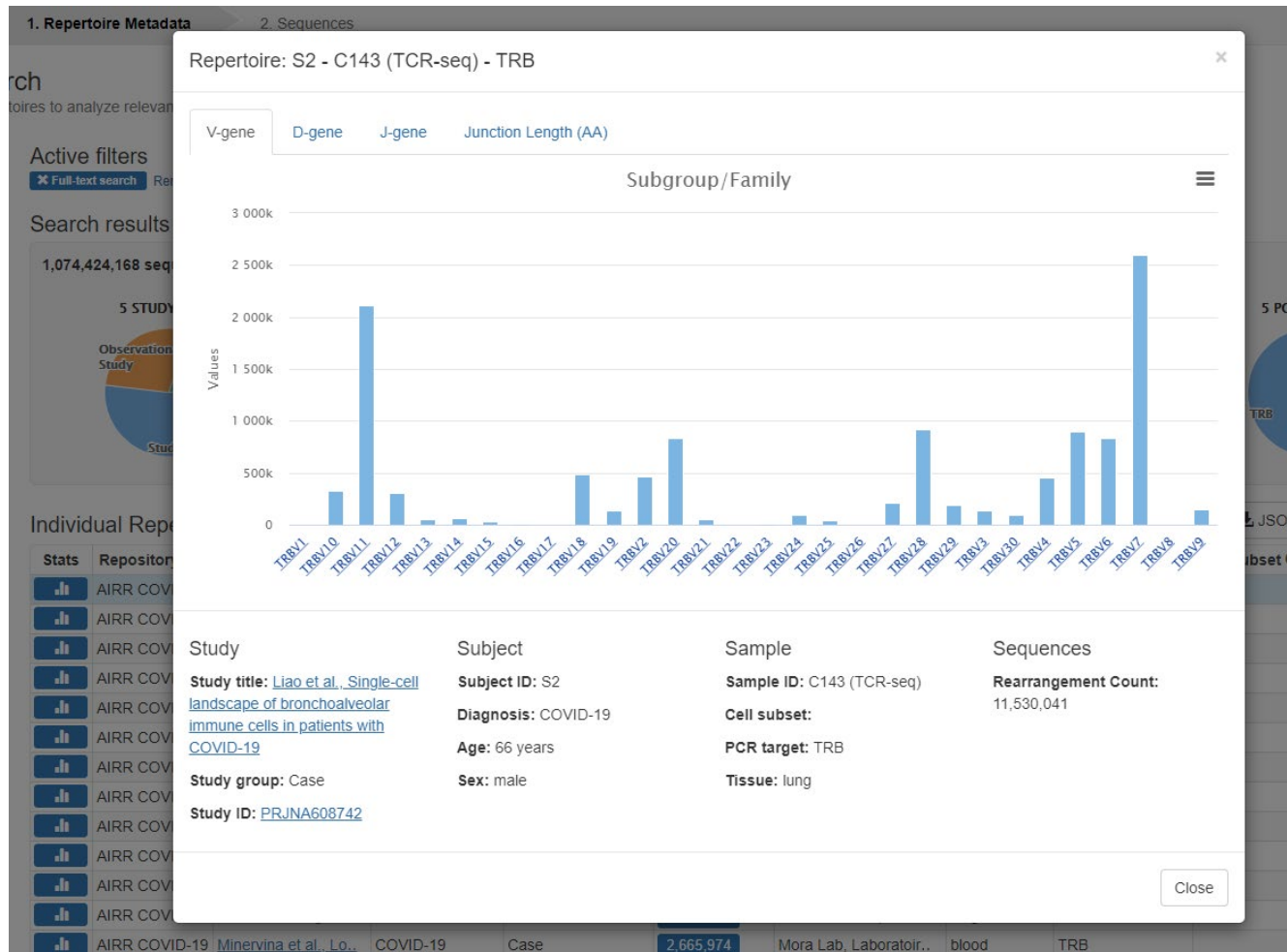


Tissue	Count
blood	1
lung	1
various	1
blood	1
Other	1
Other	1

Individual Repertoires 1-120 of 3624 [Customize displayed columns](#)

How do I use the AIRR Data Commons?

- Simple analyses are trivial - Gene usage, CDR3 length distributions



How do I use the AIRR Data Commons?

- Want to do something more complicated/computational?
 - Use the ADC API directly – through the command line or code - Full search and download power

How do I use the AIRR Data Commons?

- Want to do something more complicated/computational?
 - Use the ADC API directly – through the command line or code - Full search and download power
- Use case – V gene usage for antibodies that are antigen specific for SARS-CoV2

How do I use the AIRR Data Commons?

- Want to do something more complicated/computational?
 - Use the ADC API directly – through the command line or code - Full search and download power
- Use case – V gene usage for antibodies that are antigen specific for SARS-CoV2
 - Go to Oxford (OPIG) CoV-AbDab (<http://opig.stats.ox.ac.uk/webapps/covabdab/>)
 - Find a SARS-CoV2 specific antibody and get its CDR3 (e.g. 0304-2F8 - **ARDLYYGMDV**)

How do I use the AIRR Data Commons?

- Want to do something more complicated/computational?
 - Use the ADC API directly – through the command line or code - Full search and download power
- Use case – V gene usage for antibodies that are antigen specific for SARS-CoV2
 - Go to Oxford (OPIG) CoV-AbDab (<http://opig.stats.ox.ac.uk/webapps/covabdab/>)
 - Find a SARS-CoV2 specific antibody and get its CDR3 (e.g. 0304-2F8 - **ARDLYYYGMDV**)
 - Use the ADC API and “curl” to find out which v-call is associated with this CDR3 in the ADC

Query

```
curl -d '{
  "filters": {
    "op": "contains", "content": {
      "field": "junction_aa", "value": "ARDLYYYGMDV"
    }
  },
  "fields": ["v_call", "junction_aa"]
}' http://covid19-1.ireceptor.org/airr/v1/rearrangement
```

How do I use the AIRR Data Commons?

- Want to do something more complicated/computational?
 - Use the ADC API directly – through the command line or code - Full search and download power
- Use case – V gene usage for antibodies that are antigen specific for SARS-CoV2
 - Go to Oxford (OPIG) CoV-AbDab (<http://opig.stats.ox.ac.uk/webapps/covabdab/>)
 - Find a SARS-CoV2 specific antibody and get its CDR3 (e.g. 0304-2F8 - **ARDLYYYGMDV**)
 - Use the ADC API and “curl” to find out which v-call is associated with this CDR3 in the ADC
 - Found in data from 5 COVID-19 studies across 21 subjects

Query

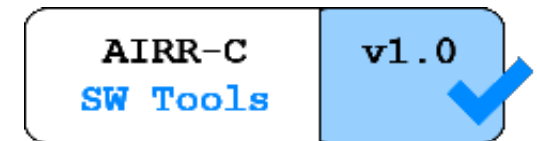
```
curl -d '{
  "filters": {
    "op": "contains", "content": {
      "field": "junction_aa", "value": "ARDLYYYGMDV"
    }
  },
  "fields": ["v_call", "junction_aa"]
}' http://covid19-1.ireceptor.org/airr/v1/rearrangement
```

Response (partial)

```
[ ...
  { "junction_aa": "CARDLYYYGMDVW", "v_call": "IGHV3-53*04"},
  { "junction_aa": "CARDLYYYGMDVW", "v_call": "IGHV3-53*04"},
  { "junction_aa": "CARDLYYYGMDVW", "v_call": "IGHV3-53*04"},
  { "junction_aa": "CARDLYYYGMDVW", "v_call": "IGHV3-53*04"},
  { "junction_aa": "CARDLYYYGMDVW", "v_call": "IGHV3-53*01"},
  { "junction_aa": "CARDLYYYGMDVW", "v_call": "IGHV3-66*02"},
  { "junction_aa": "CARDLYYYGMDVW", "v_call": "IGHV3-53*01"},
  { "junction_aa": "CARDLYYYGMDVW", "v_call": "IGHV3-53*01"}
  ... ]
```

How do I use the AIRR Data Commons?

- Analyze the data you find in the ADC
- The AIRR Software compliance badge
 - Requirements
 - Open source, support AIRR Standards, example data, containerization, user support
 - https://docs.airr-community.org/en/stable/swtools/airr_swtools_standard.html
- Current AIRR certified analysis tools:
 - Sonar (Schramm et al. Front Immunol, 2016)
 - ImmuneDB (Rosenfeld et al. Front Immunol, 2018)
 - Scirpy (Sturm et al. Bioinformatics, 2020)
 - Immcantation (Vander Heiden et al. Bioinformatics, 2014)
 - Your tool goes here...



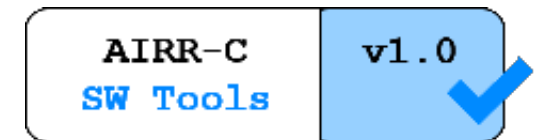
How do I share data in the AIRR Data Commons?

- Reach out to your colleagues in the AIRR Community
 - We can help!
- Run your own repository (iReceptor Turnkey)
 - Docker containers, simple to install and run (Github download)
 - www.ireceptor.org/repositories#turnkey
- Upload your data to an existing repository (VDJServer)
 - Upload data, run it through the VDJServer pipeline, publish!
 - www.vdjserver.org



What is next for the AIRR Data Commons?

- AIRR extensions to the standard – coming soon to a repository near you!
 - Clones and clonal lineages
 - Single cell capabilities
- More advanced analyses - iReceptor Plus (www.ireceptorplus.org)
 - Extensions to existing Statistics capabilities
 - Don't just download - Integrated AIRR analysis tools
 - E.g. Run Immcantation analysis on data in the ADC
 - Never leave the web based UIs (iReceptor Gateway, VDJServer)



Part III:
AIRR Data Commons Tool Demos

The iReceptor Gateway

Using the iReceptor Gateway: home page



Username

kira_neller

Password

.....

Log In →

Apply for an account by emailing
support@ireceptor.org.

What's New

New Data and New Statistics!

New Data: New study from Sokal et al. [Maturation and persistence of the anti-SARS-CoV-2 memory B cell response](#) added.

Statistics: Remember to check out our new [Stats analysis capability](#).

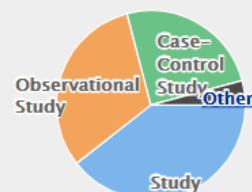
[Read all news →](#)

A **science gateway** that enables the discovery, analysis and download of [AIRR-seq data](#) (antibody/B-cell and T-cell receptor repertoires) from multiple independent repositories (the [AIRR Data Commons](#)), including:

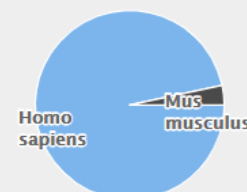
- [iReceptor Public Archive](#)
- [AIRR COVID-19](#)
- [VDJServer](#)

Search study metadata and sequence features from **4 billion sequences**, **6110 repertoires**, and **60 studies** across **4 remote repositories**.

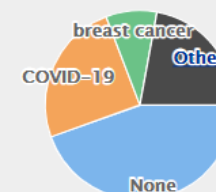
5 STUDY TYPES



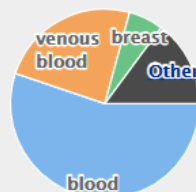
2 ORGANISMS



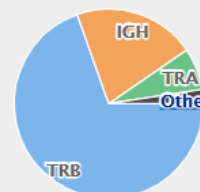
44 DIAGNOSES



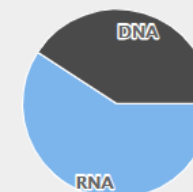
24 TISSUES



5 PCR TARGETS



2 TARGET SUBSTRATES



Using the iReceptor Gateway: home page



Username

kira_neller

Password

.....

Log In →

Apply for an account by emailing
support@ireceptor.org.

What's New

New Data and New Statistics!

New Data: New study from Sokal et al. [Maturation and persistence of the anti-SARS-CoV-2 memory B cell response](#) added.

Statistics: Remember to check out our new [Stats analysis capability](#).

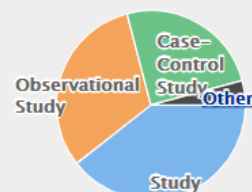
[Read all news →](#)

A **science gateway** that enables the discovery, analysis and download of AIRR-seq data (antibody/B-cell and T-cell receptor repertoires) from multiple independent repositories (the [AIRR Data Commons](#)), including:

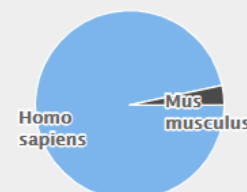
- [iReceptor Public Archive](#)
- [AIRR COVID-19](#)
- [VDJServer](#)

Search study metadata and sequence features from **4 billion sequences**, **6110 repertoires**, and **60 studies** across **4 remote repositories**.

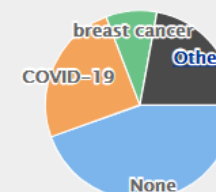
5 STUDY TYPES



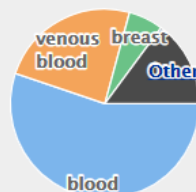
2 ORGANISMS



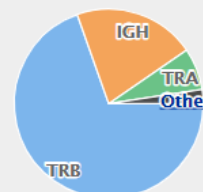
44 DIAGNOSES



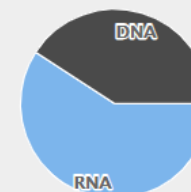
24 TISSUES



5 PCR TARGETS



2 TARGET SUBSTRATES



Using the iReceptor Gateway: home page



A **science gateway** that enables the discovery, analysis and download of [AIRR-seq data](#) (antibody/B-cell and T-cell receptor repertoires) from multiple independent repositories (the [AIRR Data Commons](#)), including:

- [iReceptor Public Archive](#)
- [AIRR COVID-19](#)
- [VDJServer](#)

Username

kira_neller

Password

.....

Log In →

Apply for an account by emailing support@ireceptor.org.

What's New

New Data and New Statistics!

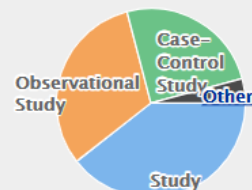
New Data: New study from Sokal et al. [Maturation and persistence of the anti-SARS-CoV-2 memory B cell response](#) added.

Statistics: Remember to check out our new [Stats analysis capability](#).

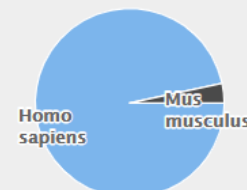
[Read all news →](#)

Search study metadata and sequence features from **4 billion sequences**, **6110 repertoires**, and **60 studies** across **4 remote repositories**.

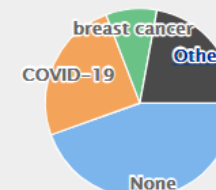
5 STUDY TYPES



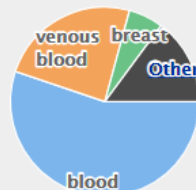
2 ORGANISMS



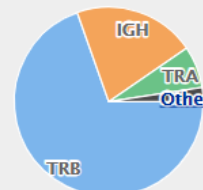
44 DIAGNOSES



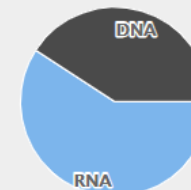
24 TISSUES



5 PCR TARGETS



2 TARGET SUBSTRATES



Using the iReceptor Gateway: home page



Username

kira_neller

Password

.....

Log In →

Apply for an account by emailing
support@ireceptor.org.

What's New

New Data and New Statistics!

New Data: New study from Sokal et al. [Maturation and persistence of the anti-SARS-CoV-2 memory B cell response](#) added.

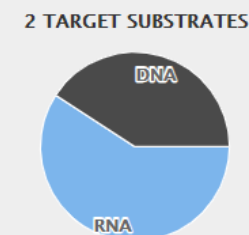
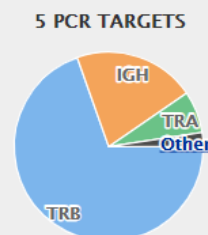
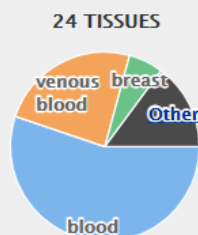
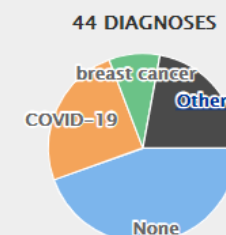
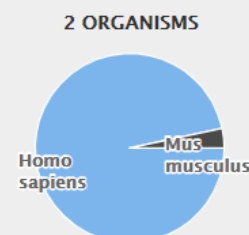
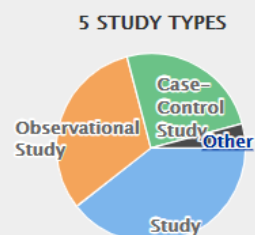
Statistics: Remember to check out our new [Stats analysis capability](#).

[Read all news →](#)

A **science gateway** that enables the discovery, analysis and download of [AIRR-seq data](#) (antibody/B-cell and T-cell receptor repertoires) from multiple independent repositories (the [AIRR Data Commons](#)), including:

- [iReceptor Public Archive](#)
- [AIRR COVID-19](#)
- [VDJServer](#)

Search study metadata and sequence features from **4 billion sequences**, **6110 repertoires**, and **60 studies** across **4 remote repositories**.



Using the iReceptor Gateway: home page



A **science gateway** that enables the discovery, analysis and download of [AIRR-seq data](#) (antibody/B-cell and T-cell receptor repertoires) from multiple independent repositories (the [AIRR Data Commons](#)), including:

4 remote repositories, 49 research labs, 60 studies

🏠 AIRR COVID-19 *747 million sequences*

📁 Lab: Adaptive Biotechnologies *543.4 million sequences*

- 📄 Study: [Nolan et al., COVID-19-Adaptive: A large-scale database of T-cel..](#) *111.7 million sequences*
- 📄 Study: [Nolan et al., COVID-19-BWNW: A large-scale database of T-cell re..](#) *33.2 million sequences*
- 📄 Study: [Nolan et al., COVID-19-DLS: A large-scale database of T-cell rec..](#) *155.4 million sequences*
- 📄 Study: [Nolan et al., COVID-19-HUniv12Oct: A large-scale database of T-c..](#) *84.2 million sequences*
- 📄 Study: [Nolan et al., COVID-19-IRST/AUSL: A large-scale database of T-ce..](#) *29 million sequences*
- 📄 Study: [Nolan et al., COVID-19-ISB: A large-scale database of T-cell rec..](#) *74.3 million sequences*
- 📄 Study: [Nolan et al., COVID-19-NIH/NIAID: A large-scale database of T-ce..](#) *55.4 million sequences*

Using the iReceptor Gateway: home page



Username

kira_neller

Password

.....

Log In →

Apply for an account by emailing
support@ireceptor.org.

What's New

New Data and New Statistics!

New Data: New study from Sokal et al. [Maturation and persistence of the anti-SARS-CoV-2 memory B cell response](#) added.

Statistics: Remember to check out our new [Stats analysis capability](#).

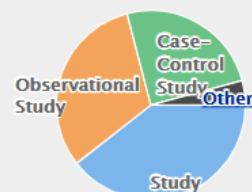
[Read all news →](#)

A **science gateway** that enables the discovery, analysis and download of [AIRR-seq data](#) (antibody/B-cell and T-cell receptor repertoires) from multiple independent repositories (the [AIRR Data Commons](#)), including:

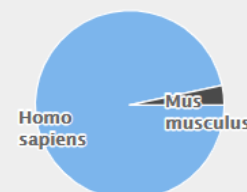
- [iReceptor Public Archive](#)
- [AIRR COVID-19](#)
- [VDJServer](#)

Search study metadata and sequence features from **4 billion sequences**, **6110 repertoires**, and **60 studies** across **4 remote repositories**.

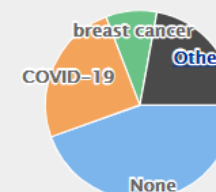
5 STUDY TYPES



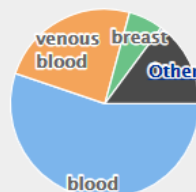
2 ORGANISMS



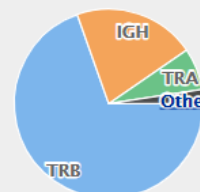
44 DIAGNOSES



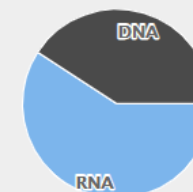
24 TISSUES



5 PCR TARGETS



2 TARGET SUBSTRATES



Using the iReceptor Gateway: home page



Username

kira_neller

Password

.....

Log In →

Apply for an account by emailing
support@ireceptor.org.

What's New

New Data and New Statistics!

New Data: New study from Sokal et al. [Maturation and persistence of the anti-SARS-CoV-2 memory B cell response](#) added.

Statistics: Remember to check out our new [Stats analysis capability](#).

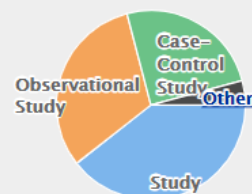
[Read all news →](#)

A **science gateway** that enables the discovery, analysis and download of [AIRR-seq data](#) (antibody/B-cell and T-cell receptor repertoires) from multiple independent repositories (the [AIRR Data Commons](#)), including:

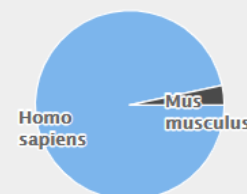
- [iReceptor Public Archive](#)
- [AIRR COVID-19](#)
- [VDJServer](#)

Search study metadata and sequence features from **4 billion sequences**, **6110 repertoires**, and **60 studies** across **4 remote repositories**.

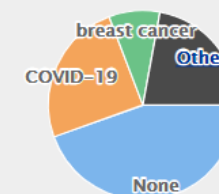
5 STUDY TYPES



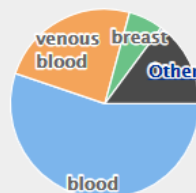
2 ORGANISMS



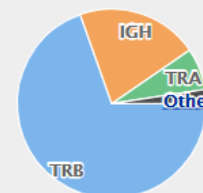
44 DIAGNOSES



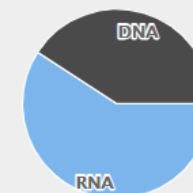
24 TISSUES



5 PCR TARGETS



2 TARGET SUBSTRATES



Using the iReceptor Gateway: home page



A **science gateway** that enables the discovery, analysis and download of [AIRR-seq data](#) (antibody/B-cell and T-cell receptor repertoires) from multiple independent repositories (the [AIRR Data Commons](#)), including:

- [iReceptor Public Archive](#)
- [AIRR COVID-19](#)
- [VDJServer](#)

Username

kira_neller

Password

.....

Log In →

Apply for an account by emailing support@ireceptor.org.

What's New

New Data and New Statistics!

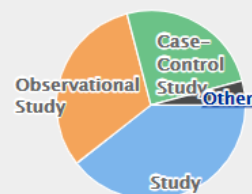
New Data: New study from Sokal et al. [Maturation and persistence of the anti-SARS-CoV-2 memory B cell response](#) added.

Statistics: Remember to check out our new [Stats analysis capability](#).

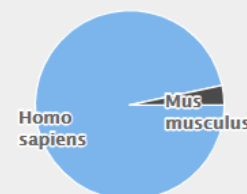
[Read all news →](#)

Search study metadata and sequence features from **4 billion sequences**, **6110 repertoires**, and **60 studies** across **4 remote repositories**.

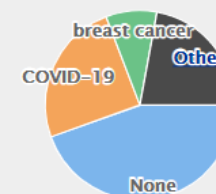
5 STUDY TYPES



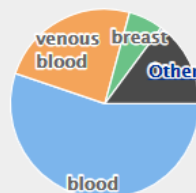
2 ORGANISMS



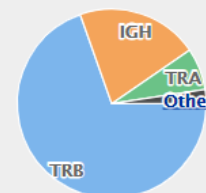
44 DIAGNOSES



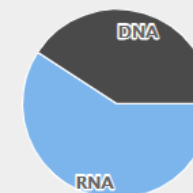
24 TISSUES



5 PCR TARGETS



2 TARGET SUBSTRATES



Using the iReceptor Gateway: Metadata Search



Search ▾

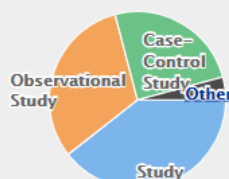
Help ▾

admin ▾

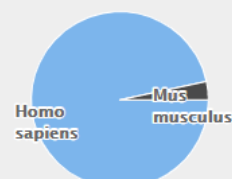
kira_neller ▾

4 billion sequences and **6110 repertoires** are currently available, from [4 remote repositories](#), [49 research labs](#) and [60 studies](#).

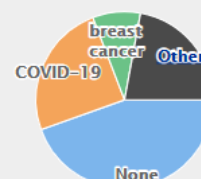
5 STUDY TYPES



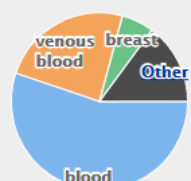
2 ORGANISMS



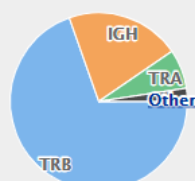
44 DIAGNOSES



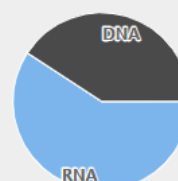
24 TISSUES



5 PCR TARGETS



2 TARGET SUBSTRATES



Repertoire Metadata Search

Find interesting sequences and sequence annotations by exploring study, subject, and sample metadata

[Browse Repertoire Metadata →](#)

Sequence Quick Search

Find sequences through all repositories with a specific Junction/CDR3 AA substring.

Junction/CDR3 AA ?

Cell subset ?

Any ▾

Organism ?

Any ▾

[Search](#)

About iReceptor

Our Mission

iReceptor facilitates the curation, analysis and sharing of antibody/B-cell and T-cell receptor repertoires (Adaptive Immune Receptor Repertoire or AIRR-seq data) from multiple labs and institutions. We are committed to providing a platform for researchers to increase the value of their data through sharing with the community. This will greatly increase the amount of data available to answer complex questions about the adaptive immune response, accelerating the development of vaccines, therapeutic antibodies against autoimmune diseases, and cancer immunotherapies.

Getting Started

- [Displaying repertoire statistics](#)
- [Finding COVID-19 data](#)
- [Finding B-cell or T-cell repertoires](#)
- [Finding sequences using a specific V-Gene](#)
- [Downloading sequences](#)
- [Frequently Asked Questions](#)

See the [iReceptor Gateway documentation](#) for more.

Using the iReceptor Gateway: Metadata Search



Search ▾

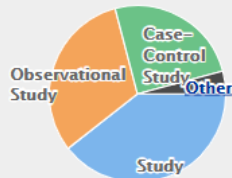
Help ▾

admin ▾

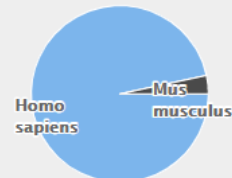
kira_neller ▾

4 billion sequences and **6110 repertoires** are currently available, from [4 remote repositories](#), [49 research labs](#) and [60 studies](#).

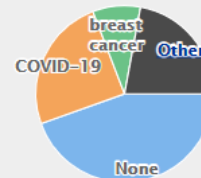
5 STUDY TYPES



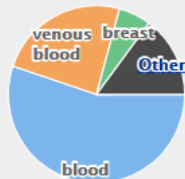
2 ORGANISMS



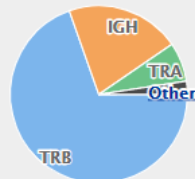
44 DIAGNOSES



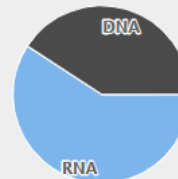
24 TISSUES



5 PCR TARGETS



2 TARGET SUBSTRATES



Repertoire Metadata Search

Find interesting sequences and sequence annotations by exploring study, subject, and sample metadata

[Browse Repertoire Metadata →](#)

Sequence Quick Search

Find sequences through all repositories with a specific Junction/CDR3 AA substring.

Junction/CDR3 AA ?

Cell subset ?

Any ▾

Organism ?

Any ▾

[Search](#)

About iReceptor

Our Mission

iReceptor facilitates the curation, analysis and sharing of antibody/B-cell and T-cell receptor repertoires (Adaptive Immune Receptor Repertoire or AIRR-seq data) from multiple labs and institutions. We are committed to providing a platform for researchers to increase the value of their data through sharing with the community. This will greatly increase the amount of data available to answer complex questions about the adaptive immune response, accelerating the development of vaccines, therapeutic antibodies against autoimmune diseases, and cancer immunotherapies.

Getting Started

- [Displaying repertoire statistics](#)
- [Finding COVID-19 data](#)
- [Finding B-cell or T-cell repertoires](#)
- [Finding sequences using a specific V-Gene](#)
- [Downloading sequences](#)
- [Frequently Asked Questions](#)

See the [iReceptor Gateway documentation](#) for more.

Using the iReceptor Gateway: Metadata Search

1. Repertoire Metadata Search

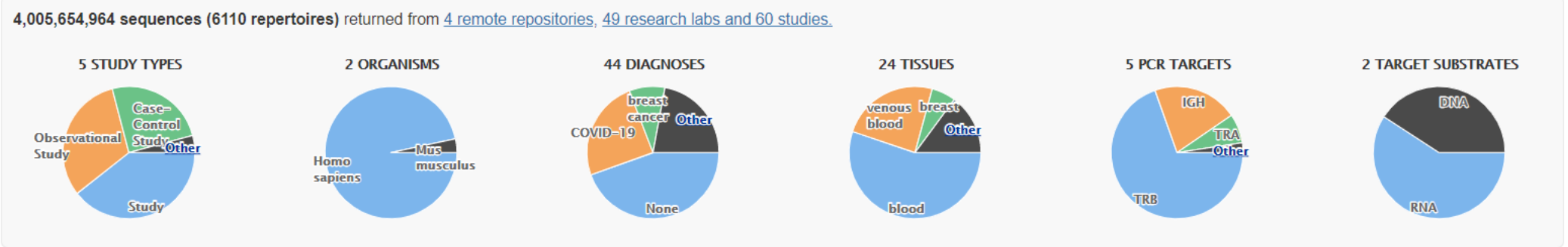
Filter by study/subject/sample and choose repertoires to analyze relevant sequence data

Filters

Full-text search ?

- Filter by study +
- Filter by subject +
- Filter by sample +
- More filters +

Search results statistics



Individual Repertoires 1-120 of 6110 [Customize displayed columns](#)

[JSON](#)

[TSV](#)

[Browse sequences from 6110 repertoires →](#)

Stats	Repository	Study title ?	Diagnosis ?	Study group ?	Sequences	Lab name ?	Tissue ?	PCR target ?	Cell subset ?	Cell subset phenotype ?
	AIRR COVID-19	Minervina et al., Lo..	COVID-19	Case	229,933	Mora Lab, Laboratoir..	blood	TRA	CD8-positive, alpha-..	CD8+
	AIRR COVID-19	Nolan et al., COVID-..			98,230	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	83,967	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	194,683	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	71,321	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	27,389	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	91,860	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	134,902	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	157,489	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	150,840	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	236,319	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	105,620	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	63,558	Adaptive Biotechnolo..	blood	TRB		

Using the iReceptor Gateway: Metadata Search

1. Repertoire Metadata Search

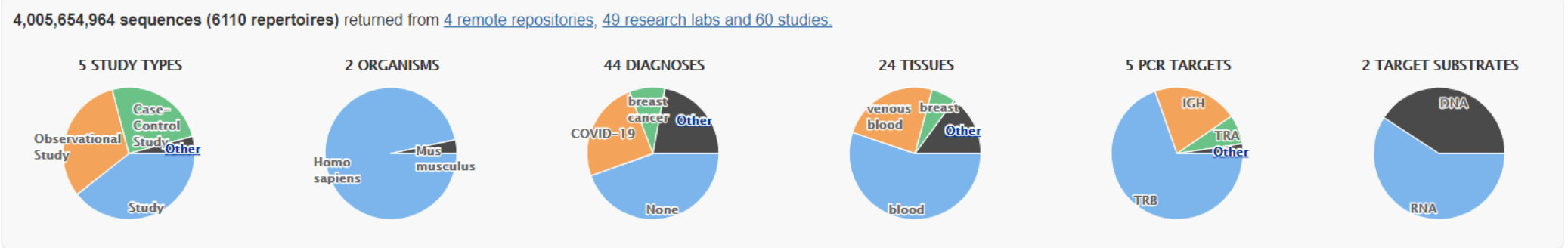
Filter by study/subject/sample and choose repertoires to analyze relevant sequence data

Filters

Full-text search

- Filter by study +
- Filter by subject +
- Filter by sample +
- More filters +

Search results statistics



Individual Repertoires 1-120 of 6110 [Customize displayed columns](#)

[JSON](#)

[TSV](#)

[Browse sequences from 6110 repertoires →](#)

Stats	Repository	Study title	Diagnosis	Study group	Sequences	Lab name	Tissue	PCR target	Cell subset	Cell subset phenotype
	AIRR COVID-19	Minervina et al., Lo..	COVID-19	Case	229,933	Mora Lab, Laboratoir..	blood	TRA	CD8-positive, alpha-..	CD8+
	AIRR COVID-19	Nolan et al., COVID-..			98,230	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	83,967	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	194,683	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	71,321	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	27,389	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	91,860	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	134,902	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	157,489	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	150,840	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	236,319	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	105,620	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	63,558	Adaptive Biotechnolo..	blood	TRB		

Using the iReceptor Gateway: Metadata Search

Customize displayed columns

✕ Close

Repertoire

- ☐ ? Repertoire Description
- ☐ ? Repertoire ID
- ☐ ? Repertoire Name

Study

- ☐ ? Contact (collection)
- ☐ ? Contact (deposition)
- ☐ ? Funding
- ☐ ? Inclusion criteria
- ☐ ? Lab address
- ☒ ? Lab name
- ☒ ? Publications
- ☐ ? Study Description
- ☒ ? Study ID
- ☐ ? Study keywords
- ☒ ? Study title
- ☐ ? Study type
- ☐ ? Study type (Ontology ID)

Subject

- ☐ Age (deperrecated)
- ☐ ? Age event
- ☐ ? Age maximum
- ☐ ? Age minimum
- ☐ ? Age unit

- ☐ ? Age unit (Ontology ID)

- ☐ ? Ancestry
- ☐ ? Ethnicity
- ☐ ? Organism
- ☐ ? Organism (deprecated)
- ☐ ? Organism (deprecated)
- ☐ ? Organism (Ontology ID)
- ☐ ? Race
- ☐ ? Relation (subjects)
- ☐ ? Relation type
- ☐ ? Sex
- ☐ ? Strain name
- ☒ ? Subject ID
- ☐ ? Synthetic library

Sample

- ☐ ? Anatomic site
- ☐ ? Biomaterial provider
- ☐ ? Collection event
- ☐ ? Collection time
- ☐ ? Disease state
- ☒ ? Sample ID
- ☐ ? Sample type
- ☒ ? Tissue
- ☐ ? Tissue (Ontology ID)

Diagnosis

- ☒ ? Diagnosis
- ☐ ? Diagnosis (Ontology ID)
- ☐ ? Disease stage
- ☐ ? Immunogen/agent
- ☐ ? Intervention
- ☐ ? Length of disease
- ☐ ? Medical history
- ☐ ? Prior therapies
- ☒ ? Study group

Cell Processing

- ☐ ? # cells/experiment
- ☐ ? # cells/sequencing reaction
- ☐ ? Cell isolation procedure
- ☐ ? Cell quality
- ☐ ? Cell species
- ☐ ? Cell species (Ontology ID)
- ☐ ? Cell storage
- ☒ ? Cell subset
- ☐ ? Cell subset (Ontology ID)
- ☒ ? Cell subset phenotype
- ☐ ? Processing protocol
- ☐ ? Single-cell sort
- ☐ ? Tissue processing

Nucleic Acid Processing

- ☐ ? Complete sequences
- ☐ ? Library generation method
- ☐ ? Library generation protocol
- ☐ ? Linkage of loci
- ☐ ? Protocol IDs
- ☒ ? Target substrate
- ☐ ? Target substrate quality
- ☐ ? Template amount

Sequencing Run

- ☐ ? Batch number
- ☐ ? Date of sequencing run
- ☐ ? Reads passing QC
- ☐ ? Sequencing facility
- ☐ ? Sequencing kit
- ☒ ? Sequencing platform

Data Processing

- ☐ ? Analysis ID
- ☐ ? Collapsing method
- ☐ ? Data processing ID
- ☐ ? Data processing protocols
- ☐ ? Paired read assembly
- ☐ ? Primary annotation
- ☐ ? Primer match cutoffs

- ☐ ? Processed files

- ☐ ? Quality thresholds
- ☐ ? Software tools/versions
- ☐ ? V(D)J germline database

Sample Processing

- ☐ ? Sample Processing ID

Raw Sequence Data

- ☐ ? Forward read length
- ☐ ? Paired read direction
- ☐ ? Paired read length
- ☐ ? Paired sequencing file name
- ☐ ? Read direction
- ☐ ? Sequencing file name
- ☐ ? Sequencing file type

PCR Target

- ☐ ? Forward PCR target
- ☒ ? PCR target
- ☐ ? Reverse PCR target

Other

- ☒ Repository
- ☒ Sequences

Using the iReceptor Gateway: Search Filters

1. Repertoire Metadata Search

Filter by study/subject/sample and choose repertoires to analyze relevant sequence data

Filters

Full-text search ⓘ

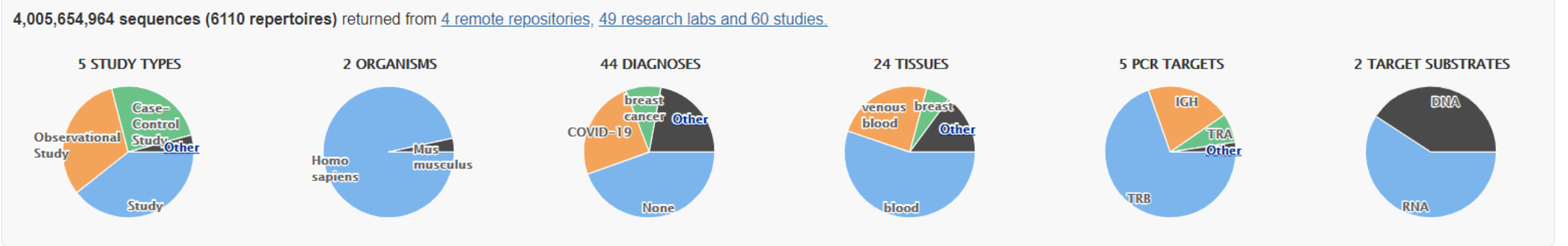
Filter by study +

Filter by subject +

Filter by sample +

More filters +

Search results statistics



Individual Repertoires 1-120 of 6110 [Customize displayed columns](#)

↓ JSON

↓ TSV

Browse sequences from 6110 repertoires →

Stats	Repository	Study title ⓘ	Diagnosis ⓘ	Study group ⓘ	Sequences	Lab name ⓘ	Tissue ⓘ	PCR target ⓘ	Cell subset ⓘ	Cell subset phenotype ⓘ
	AIRR COVID-19	Minervina et al., Lo..	COVID-19	Case	229,933	Mora Lab, Laboratoir..	blood	TRA	CD8-positive, alpha-..	CD8+
	AIRR COVID-19	Nolan et al., COVID-..			98,230	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	83,967	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	194,683	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	71,321	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	27,389	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	91,860	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	134,902	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	157,489	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	150,840	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	236,319	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	105,620	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	63,558	Adaptive Biotechnolo..	blood	TRB		

Using the iReceptor Gateway: Search Filters

1. Repertoire Metadata Search

Filter by study/subject/sample and choose repertoires to analyze relevant sequence data

Filters

Full-text search ⓘ

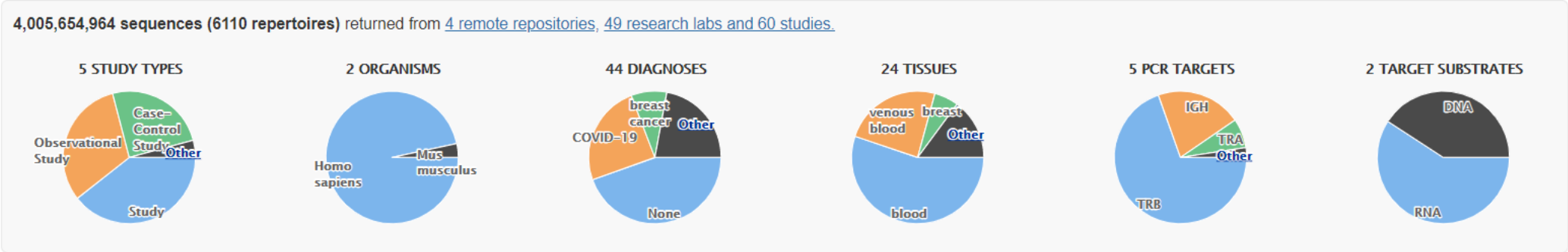
Filter by study +

Filter by subject +

Filter by sample +

More filters +

Search results statistics



Individual Repertoires 1-120 of 6110 [Customize displayed columns](#)

↓ JSON

↓ TSV

Browse sequences from 6110 repertoires →

Stats	Repository	Study title ⓘ	Diagnosis ⓘ	Study group ⓘ	Sequences	Lab name ⓘ	Tissue ⓘ	PCR target ⓘ	Cell subset ⓘ	Cell subset phenotype ⓘ
	AIRR COVID-19	Minervina et al., Lo..	COVID-19	Case	229,933	Mora Lab, Laboratoir..	blood	TRA	CD8-positive, alpha-..	CD8+
	AIRR COVID-19	Nolan et al., COVID-..			98,230	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	83,967	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	194,683	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	71,321	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	27,389	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	91,860	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	134,902	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	157,489	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	150,840	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	236,319	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	105,620	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	63,558	Adaptive Biotechnolo..	blood	TRB		

Using the iReceptor Gateway: Search Filters

Filters

Full-text search

covid-19

Filter by study



Filter by subject



Filter by sample



More filters



Using the iReceptor Gateway: Search Filters

Filters

Full-text search ?

covid-19

Filter by study +

Filter by subject +

Filter by sample +

More filters +

Filter by study -

Study ID ?

Study title ?

Study type ?

Any ▾

Study group ?

Lab name ?

Apply filters →

Using the iReceptor Gateway: Search Filters

Filters

Full-text search ?

covid-19

Filter by study +

Filter by subject +

Filter by sample +

More filters +

Filter by study -

Study ID ?

Study title ?

Study type ?

Any ▼

Study group ?

Lab name ?

Apply filters →

Filter by subject -

Subject ID ?

Organism ?

Any ▼

Sex ?

Any ▼

Ethnicity ?

Any ▼

Age minimum ?

Age maximum ?

Diagnosis ?

Any ▼

Apply filters →

Using the iReceptor Gateway: Search Filters

Filters

Full-text search ?

covid-19

Filter by study +

Filter by subject +

Filter by sample +

More filters +

Filter by study -

Study ID ?

Study title ?

Study type ?

Any

Study group ?

Lab name ?

Apply filters →

Filter by subject -

Subject ID ?

Organism ?

Any

Sex ?

Any

Ethnicity ?

Any

Age minimum ?

Age maximum ?

Diagnosis ?

Any

Apply filters →

Filter by sample -

Sample ID ?

PCR target ?

Any

Cell subset ?

Any

Tissue ?

Any

Target substrate ?

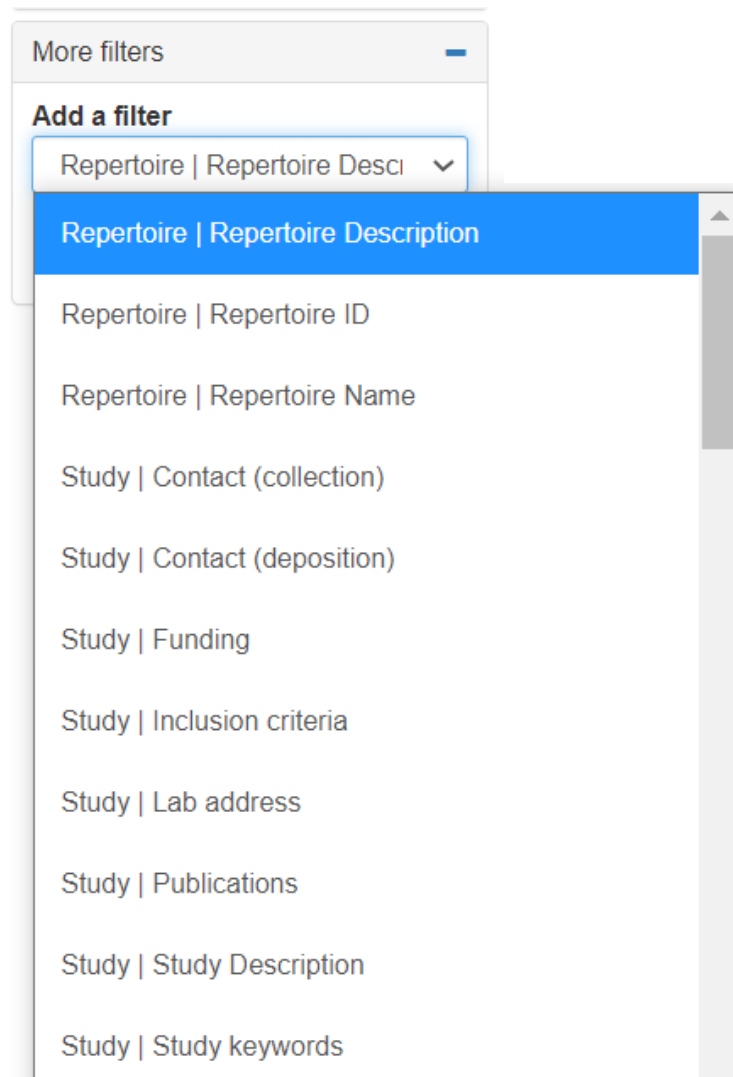
Any

Cell subset phenotype ?

Sequencing platform ?

Apply filters →

Using the iReceptor Gateway: Search Filters



Using the iReceptor Gateway: Search Filters

More filters —

Add a filter

Repertoire | Repertoire Desci ▼

- Repertoire | Repertoire Description
- Repertoire | Repertoire ID
- Repertoire | Repertoire Name
- Study | Contact (collection)
- Study | Contact (deposition)
- Study | Funding
- Study | Inclusion criteria
- Study | Lab address
- Study | Publications
- Study | Study Description
- Study | Study keywords

More filters —

Apply filters →

Add a filter

Repertoire | Repertoire Desci ▼

+ Add filter

Using the iReceptor Gateway: Search Filters

More filters

Add a filter

Repertoire | Repertoire Desci

Repertoire | Repertoire Description

Repertoire | Repertoire ID

Repertoire | Repertoire Name

Study | Contact (collection)

Study | Contact (deposition)

Study | Funding

Study | Inclusion criteria

Study | Lab address

Study | Publications

Study | Study Description

Study | Study keywords

More filters

Apply filters →

Add a filter

Repertoire | Repertoire Desci

+ Add filter

More filters

Repertoire Description ? x


Apply filters →

Add a filter

Repertoire | Repertoire ID

+ Add filter

Using the iReceptor Gateway: Filtered Results



Search ▾

1. Repertoire Metadata

2. Sequences

Help ▾

admin ▾

kira_neller ▾

1. Repertoire Metadata Search

Filter by study/subject/sample and choose repertoires to analyze relevant sequence data

Filters

Full-text search ?

covid-19

Filter by study +

Filter by subject +

Filter by sample +

More filters +

Active filters

✕ Full-text search

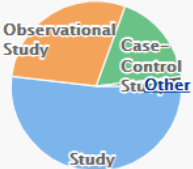
Remove all filters

☆ Bookmark this search


Search results statistics

1,074,424,168 sequences (3624 repertoires) returned from 3 remote repositories, 15 research labs and 21 studies.

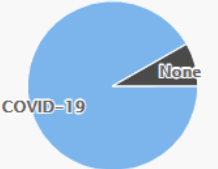
5 STUDY TYPES



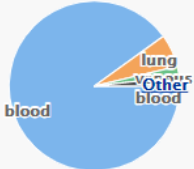
2 ORGANISMS



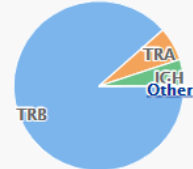
2 DIAGNOSES



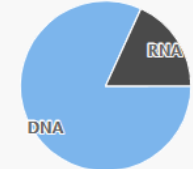
6 TISSUES



5 PCR TARGETS



2 TARGET SUBSTRATES



Individual Repertoires 1-120 of 3624

Customize displayed columns


Download JSON

Download TSV

Browse sequences from 3624 repertoires →

Stats	Repository	Study title	Diagnosis	Study group	Sequences	Lab name	Tissue	PCR target	Cell subset	Cell subset phenotype
	AIRR COVID-19	Schultheis et al., N...	COVID-19	Case	83,468	Binder Lab, Departme..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	171,810	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	496,996	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	160,409	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	294,339	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	333,999	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	296,487	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	340,151	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	360,309	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	413,604	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	128,286	Adaptive Biotechnolo..	blood	TRB		

Using the iReceptor Gateway: Filtered Results



Search ▾

1. Repertoire Metadata

2. Sequences

Help ▾

admin ▾

kira_neller ▾

1. Repertoire Metadata Search

Filter by study/subject/sample and choose repertoires to analyze relevant sequence data

Filters

Full-text search ?

covid-19

Filter by study +

Filter by subject +

Filter by sample +

More filters +

Active filters

✕ Full-text search

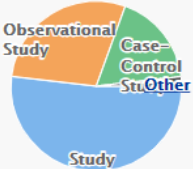
Remove all filters

☆ Bookmark this search


Search results statistics

1,074,424,168 sequences (3624 repertoires) returned from 3 remote repositories, 15 research labs and 21 studies.

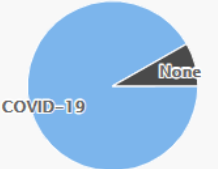
5 STUDY TYPES



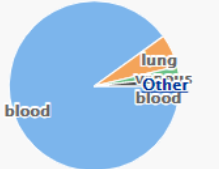
2 ORGANISMS



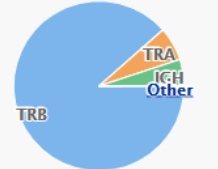
2 DIAGNOSES



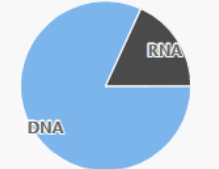
6 TISSUES



5 PCR TARGETS



2 TARGET SUBSTRATES



Individual Repertoires 1-120 of 3624

Customize displayed columns


Download JSON

Download TSV

Browse sequences from 3624 repertoires →

Stats	Repository	Study title	Diagnosis	Study group	Sequences	Lab name	Tissue	PCR target	Cell subset	Cell subset phenotype
	AIRR COVID-19	Schultheis et al., N...	COVID-19	Case	83,468	Binder Lab, Departme..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	171,810	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	496,996	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	160,409	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	294,339	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	333,999	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	296,487	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	340,151	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	360,309	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	413,604	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	128,286	Adaptive Biotechnolo..	blood	TRB		

Using the iReceptor Gateway: Filtered Results



Search ▾

1. Repertoire Metadata

2. Sequences

Help ▾

admin ▾

kira_neller ▾

1. Repertoire Metadata Search

Filter by study/subject/sample and choose repertoires to analyze relevant sequence data

Filters

Full-text search ?

covid-19

Filter by study +

Filter by subject +

Filter by sample +

More filters +

Active filters

✕ Full-text search

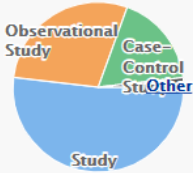
Remove all filters

☆ Bookmark this search


Search results statistics

1,074,424,168 sequences (3624 repertoires) returned from 3 remote repositories, 15 research labs and 21 studies.

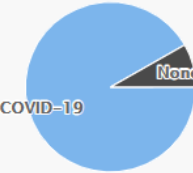
5 STUDY TYPES



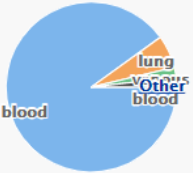
2 ORGANISMS



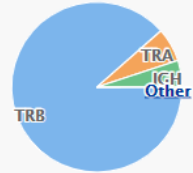
2 DIAGNOSES



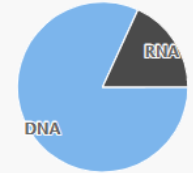
6 TISSUES



5 PCR TARGETS



2 TARGET SUBSTRATES













Individual Repertoires 1-120 of 3624

Customize displayed columns

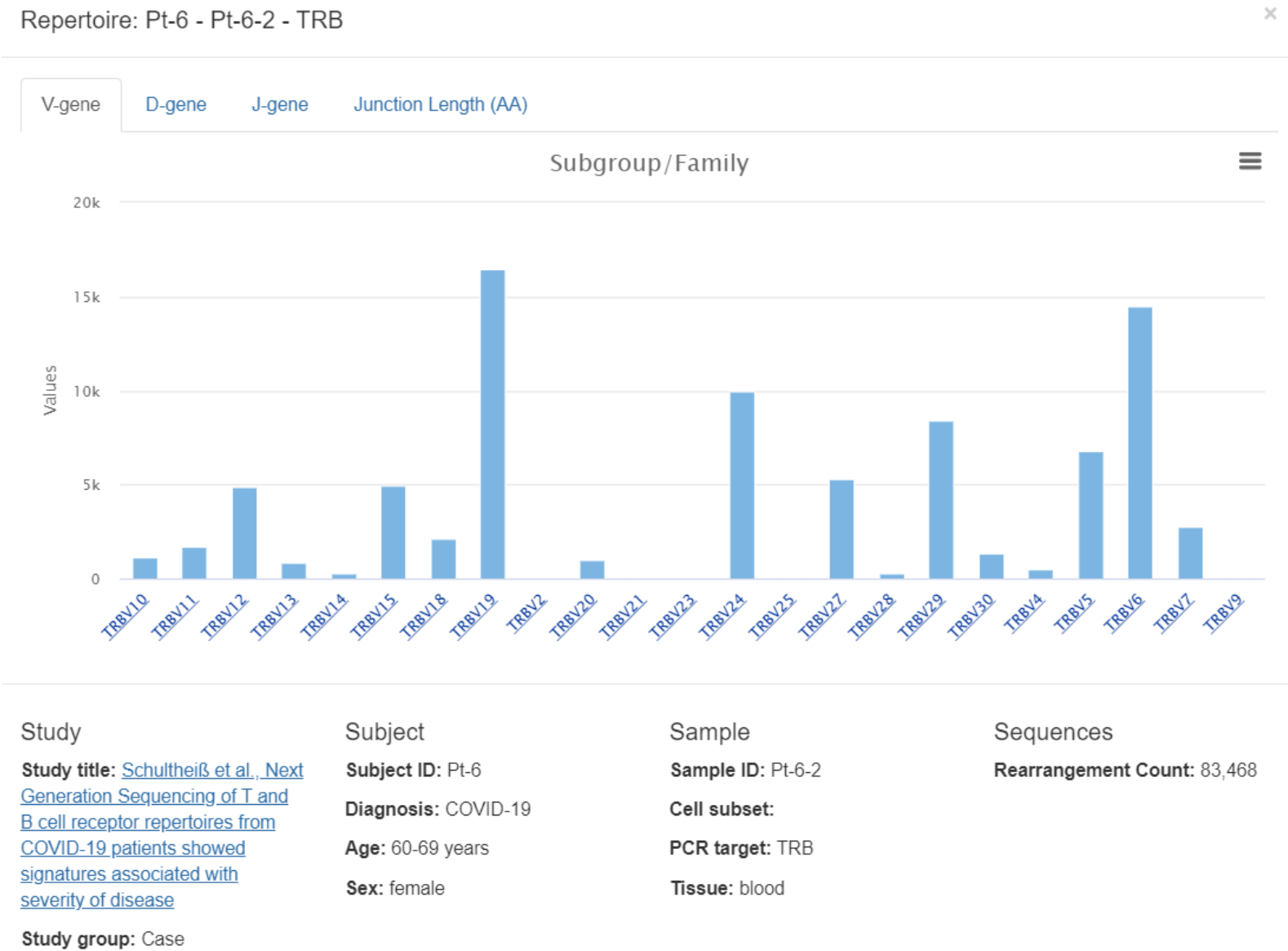
Download JSON

Download TSV


Browse sequences from 3624 repertoires →

Stats	Repository	Study title	Diagnosis	Study group	Sequences	Lab name	Tissue	PCR target	Cell subset	Cell subset phenotype
	AIRR COVID-19	Schultheis et al., N...	COVID-19	Case	83,468	Binder Lab, Departme..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	171,810	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	496,996	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	160,409	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	294,339	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	333,999	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	296,487	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	340,151	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	360,309	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	413,604	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	128,286	Adaptive Biotechnolo..	blood	TRB		

Using the iReceptor Gateway: Statistics



Using the iReceptor Gateway: Browsing Sequences



Search ▾

1. Repertoire Metadata

2. Sequences

Help ▾

admin ▾

kira_neller ▾

1. Repertoire Metadata Search

Filter by study/subject/sample and choose repertoires to analyze relevant sequence data

Filters

Full-text search ?

covid-19

Filter by study +

Filter by subject +

Filter by sample +

More filters +

Active filters

✕ Full-text search

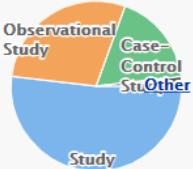
Remove all filters

☆ Bookmark this search


Search results statistics

1,074,424,168 sequences (3624 repertoires) returned from 3 remote repositories, 15 research labs and 21 studies.

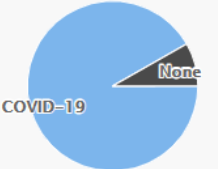
5 STUDY TYPES



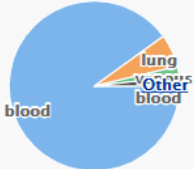
2 ORGANISMS



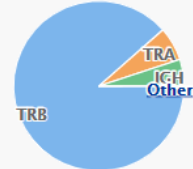
2 DIAGNOSES



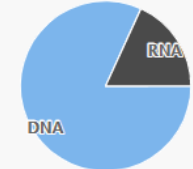
6 TISSUES



5 PCR TARGETS



2 TARGET SUBSTRATES



Individual Repertoires 1-120 of 3624

Customize displayed columns


Download JSON

Download TSV

Browse sequences from 3624 repertoires →

Stats	Repository	Study title	Diagnosis	Study group	Sequences	Lab name	Tissue	PCR target	Cell subset	Cell subset phenotype
	AIRR COVID-19	Schultheiß et al., N...	COVID-19	Case	83,468	Binder Lab, Departme..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	171,810	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	496,996	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	160,409	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	294,339	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	333,999	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	296,487	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	340,151	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	360,309	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	413,604	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-..	COVID-19	Case	128,286	Adaptive Biotechnolo..	blood	TRB		

Using the iReceptor Gateway: Browsing Sequences



Search ▾

1. Repertoire Metadata

2. Sequences

Help ▾

admin ▾

kira_neller ▾

1. Repertoire Metadata Search

Filter by study/subject/sample and choose repertoires to analyze relevant sequence data

Filters

Full-text search ?

covid-19

Filter by study +

Filter by subject +

Filter by sample +

More filters +

Active filters

✕ Full-text search

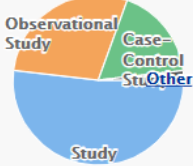
Remove all filters

☆ Bookmark this search


Search results statistics

1,074,424,168 sequences (3624 repertoires) returned from 3 remote repositories, 15 research labs and 21 studies.

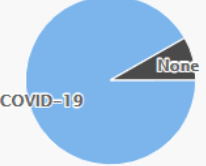
5 STUDY TYPES



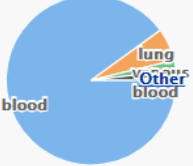
2 ORGANISMS



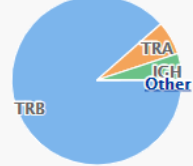
2 DIAGNOSES



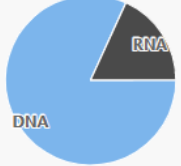
6 TISSUES



5 PCR TARGETS



2 TARGET SUBSTRATES



Individual Repertoires 1-120 of 3624 [Customize displayed columns](#)


↓ JSON

↓ TSV

Browse sequences from 3624 repertoires →

Stats	Repository	Study title	Diagnosis	Study group	Sequences	Lab name	Tissue	PCR target	Cell subset	Cell subset phenotype
	AIRR COVID-19	Schultheiß et al., N...	COVID-19	Case	83,468	Binder Lab, Departme..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-...	COVID-19	Case	171,810	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-...	COVID-19	Case	496,996	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-...	COVID-19	Case	160,409	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-...	COVID-19	Case	294,339	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-...	COVID-19	Case	333,999	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-...	COVID-19	Case	296,487	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-...	COVID-19	Case	340,151	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-...	COVID-19	Case	360,309	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-...	COVID-19	Case	413,604	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-...	COVID-19	Case	128,286	Adaptive Biotechnolo..	blood	TRB		

Using the iReceptor Gateway: Browsing Sequences



Search ▾

1. Repertoire Metadata

2. Sequences

Help ▾

admin ▾

kira_neller ▾

1. Repertoire Metadata Search

Filter by study/subject/sample and choose repertoires to analyze relevant sequence data

Filters

Full-text search ?

covid-19

Filter by study +

Filter by subject +

Filter by sample +

More filters +

Active filters

✕ Full-text search

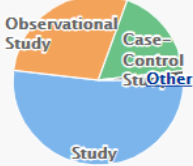
Remove all filters

☆ Bookmark this search


Search results statistics

1,074,424,168 sequences (3624 repertoires) returned from 3 remote repositories, 15 research labs and 21 studies.

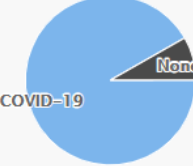
5 STUDY TYPES



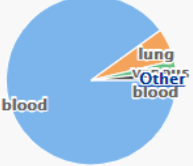
2 ORGANISMS



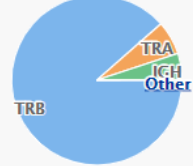
2 DIAGNOSES



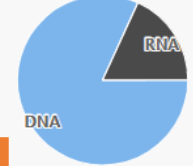
6 TISSUES



5 PCR TARGETS



2 TARGET SUBSTRATES



Individual Repertoires 1-120 of 3624 [Customize displayed columns](#)


↓ JSON

↓ TSV

Browse sequences from 3624 repertoires →

Stats	Repository	Study title	Diagnosis	Study group	Sequences	Lab name	Tissue	PCR target	Cell subset	Cell subset phenotype
	AIRR COVID-19	Schultheiß et al., N...	COVID-19	Case	83,468	Binder Lab, Departme..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-...	COVID-19	Case	171,810	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-...	COVID-19	Case	496,996	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-...	COVID-19	Case	160,409	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-...	COVID-19	Case	294,339	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-...	COVID-19	Case	333,999	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-...	COVID-19	Case	296,487	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-...	COVID-19	Case	340,151	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-...	COVID-19	Case	360,309	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-...	COVID-19	Case	413,604	Adaptive Biotechnolo..	blood	TRB		
	AIRR COVID-19	Nolan et al., COVID-...	COVID-19	Case	128,286	Adaptive Biotechnolo..	blood	TRB		

Using the iReceptor Gateway: Downloading Sequences



Search ▾

1. Repertoire Metadata

2. Sequences

Help ▾

admin ▾

kira_neller ▾

2. Sequence Search

Filter by sequence and sequence annotation feature

Filters

Filter by VDJ +

Filter by Junction AA +

Advanced filters +

Active filters

Repertoire Metadata filters: Full-text search [Go back to Repertoire Metadata Search](#)

Search results statistics

83,468 sequences (1 repertoire) returned from [1 remote repository](#), [1 research lab](#) and [1 study](#).

1 STUDY TITLE

Schultheiß et al

1 SUBJECT ID

Pt-6

1 SAMPLE ID

Pt-6-2

1 DIAGNOSIS

COVID-19

1 TISSUE

blood

1 PCR TARGET

TRB

Bookmark this search

Individual Sequences


1-10 of 83 thousand

[Customize displayed columns](#)

[Download all 83,468 sequences](#)

V Gene With Allele ?	D Gene With Allele ?	J Gene With Allele ?	Junction/CDR3 AA ?	Junction Length (AA) ?
TRBV5-1*00	TRBD2*00	TRBJ2-1*00	CARCWTSGGVNDDQFF	16
TRBV30*00	TRBD2*00	TRBJ2-2*00	CAWAGGQGGELFF	13
TRBV11-1*00	TRBD1*00	TRBJ1-1*00	CASSLEGTGFEAFF	14
TRBV18*00		TRBJ2-3*00	CASSVGG_TDTQYF	14
TRBV29-1*00	TRBD1*00	TRBJ1-1*00	CSALPGQLEAFF	12
TRBV27*00	TRBD2*00	TRBJ2-3*00	CASSLGTSGGDTQYF	15
TRBV6-5*00	TRBD1*00	TRBJ1-2*00	CASSPVTGTGHYGYTF	16
TRBV19*00		TRBJ1-1*00	CASSNMGSgteAFF	14
TRBV29-1*00		TRBJ2-1*00	CSVEPTRGWGVPYNEQFF	18
TRBV6-1*00		TRBJ1-6*00	CASSEETDWPNSPLHF	16

Using the iReceptor Gateway: Downloading Sequences



Search ▾

1. Repertoire Metadata

2. Sequences

Help ▾

admin ▾

kira_neller ▾

2. Sequence Search

Filter by sequence and sequence annotation feature

Filters

Filter by VDJ +

Filter by Junction AA +

Advanced filters +

Active filters

Repertoire Metadata filters: Full-text search [Go back to Repertoire Metadata Search](#)

Search results statistics

83,468 sequences (1 repertoire) returned from [1 remote repository](#), [1 research lab](#) and [1 study](#).

1 STUDY TITLE

Schultheiß et al

1 SUBJECT ID

Pt-6

1 SAMPLE ID

Pt-6-2

1 DIAGNOSIS

COVID-19

1 TISSUE

blood

1 PCR TARGET

TRB

Bookmark this search

Individual Sequences

1-10 of 83 thousand

[Customize displayed columns](#)

Download all 83,468 sequences

V Gene With Allele ?	D Gene With Allele ?	J Gene With Allele ?	Junction/CDR3 AA ?	Junction Length (AA) ?
TRBV5-1*00	TRBD2*00	TRBJ2-1*00	CARCWTSGGVNDDQFF	16
TRBV30*00	TRBD2*00	TRBJ2-2*00	CAWAGGQGGELFF	13
TRBV11-1*00	TRBD1*00	TRBJ1-1*00	CASSLEGTGFEAFF	14
TRBV18*00		TRBJ2-3*00	CASSVGG_TDTQYF	14
TRBV29-1*00	TRBD1*00	TRBJ1-1*00	CSALPGQLEAFF	12
TRBV27*00	TRBD2*00	TRBJ2-3*00	CASSLGTSGGDTQYF	15
TRBV6-5*00	TRBD1*00	TRBJ1-2*00	CASSPVTGTGHYGYTF	16
TRBV19*00		TRBJ1-1*00	CASSNMGSgteAFF	14
TRBV29-1*00		TRBJ2-1*00	CSVEPTRGWGVPYNEQFF	18
TRBV6-1*00		TRBJ1-6*00	CASSEETDWPNSPLHF	16

Using the iReceptor Gateway: Sequence Search



Search ▾

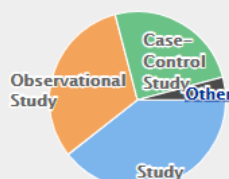
Help ▾

admin ▾

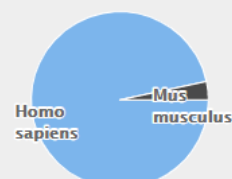
kira_neller ▾

4 billion sequences and **6110 repertoires** are currently available, from [4 remote repositories](#), [49 research labs](#) and [60 studies](#).

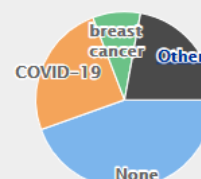
5 STUDY TYPES



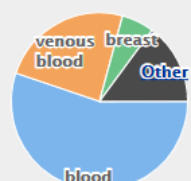
2 ORGANISMS



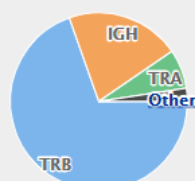
44 DIAGNOSES



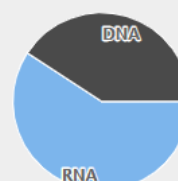
24 TISSUES



5 PCR TARGETS



2 TARGET SUBSTRATES



Repertoire Metadata Search

Find interesting sequences and sequence annotations by exploring study, subject, and sample metadata

[Browse Repertoire Metadata →](#)

Sequence Quick Search

Find sequences through all repositories with a specific Junction/CDR3 AA substring.

Junction/CDR3 AA ?

Cell subset ?

Any ▾

Organism ?

Any ▾

[Search](#)

About iReceptor

Our Mission

iReceptor facilitates the curation, analysis and sharing of antibody/B-cell and T-cell receptor repertoires (Adaptive Immune Receptor Repertoire or AIRR-seq data) from multiple labs and institutions. We are committed to providing a platform for researchers to increase the value of their data through sharing with the community. This will greatly increase the amount of data available to answer complex questions about the adaptive immune response, accelerating the development of vaccines, therapeutic antibodies against autoimmune diseases, and cancer immunotherapies.

Getting Started

- [Displaying repertoire statistics](#)
- [Finding COVID-19 data](#)
- [Finding B-cell or T-cell repertoires](#)
- [Finding sequences using a specific V-Gene](#)
- [Downloading sequences](#)
- [Frequently Asked Questions](#)

See the [iReceptor Gateway documentation](#) for more.

Using the iReceptor Gateway: Sequence Search



Search ▾

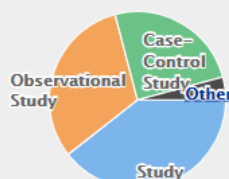
Help ▾

admin ▾

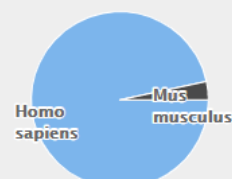
kira_neller ▾

4 billion sequences and **6110 repertoires** are currently available, from [4 remote repositories](#), [49 research labs](#) and [60 studies](#).

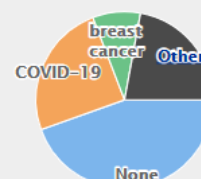
5 STUDY TYPES



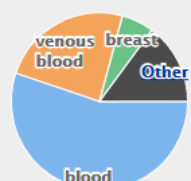
2 ORGANISMS



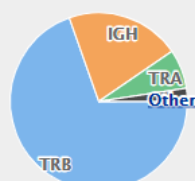
44 DIAGNOSES



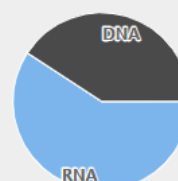
24 TISSUES



5 PCR TARGETS



2 TARGET SUBSTRATES



Repertoire Metadata Search

Find interesting sequences and sequence annotations by exploring study, subject, and sample metadata

[Browse Repertoire Metadata →](#)

Sequence Quick Search

Find sequences through all repositories with a specific Junction/CDR3 AA substring.

Junction/CDR3 AA ?

Cell subset ?

Any ▾

Organism ?

Any ▾

[Search](#)

About iReceptor

Our Mission

iReceptor facilitates the curation, analysis and sharing of antibody/B-cell and T-cell receptor repertoires (Adaptive Immune Receptor Repertoire or AIRR-seq data) from multiple labs and institutions. We are committed to providing a platform for researchers to increase the value of their data through sharing with the community. This will greatly increase the amount of data available to answer complex questions about the adaptive immune response, accelerating the development of vaccines, therapeutic antibodies against autoimmune diseases, and cancer immunotherapies.

Getting Started

- [Displaying repertoire statistics](#)
- [Finding COVID-19 data](#)
- [Finding B-cell or T-cell repertoires](#)
- [Finding sequences using a specific V-Gene](#)
- [Downloading sequences](#)
- [Frequently Asked Questions](#)

See the [iReceptor Gateway documentation](#) for more.

Using the iReceptor Gateway: Sequence Search

Sequence Quick Search

Find sequences through all repositories with a specific Junction/CDR3 AA substring.

Junction/CDR3 AA ⓘ

Cell subset ⓘ

Organism ⓘ

Using the iReceptor Gateway: Sequence Search



Search

Sequence Quick Search

Help

admin

kira_neller

Sequence Quick Search

Filter by sequence and sequence annotation features (e.g. Junction)

Filters

Junction/CDR3 AA

ARDLYYYGMDV

Sample level filters

Organism

Any

Cell subset

Any

Apply filters

Active filters

Junction/CDR3 AA

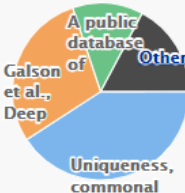
Remove all filters

Bookmark this search

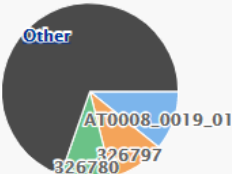
Search results statistics

906 sequences (89 repertoires) returned from 4 remote repositories, 11 research labs and 12 studies.

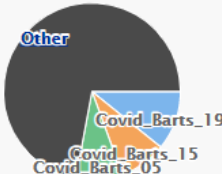
12 STUDY TITLES



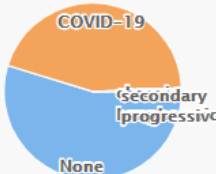
39 SUBJECT IDS



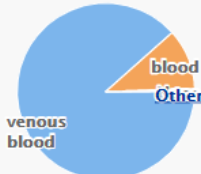
89 SAMPLE IDS



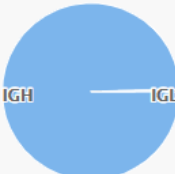
4 DIAGNOSES



5 TISSUES



2 PCR TARGETS



Individual Sequences 1-31 of 906

Customize displayed columns

Download all 906 sequences

V Gene With Allele	D Gene With Allele	J Gene With Allele	Junction/CDR3 AA	Junction Length (AA)
IGHV1-8*01		IGHJ6*02	CARDLYYYGMDVW	13
IGHV3-53*02, or IGHV3-53*01		IGHJ6*02	CARDLYYYGMDVW	13
IGHV3-53*02, or IGHV3-53*01		IGHJ6*02	CARDLYYYGMDVW	13
IGHV3-74*01		IGHJ6*02	CARDLYYYGMDVW	13
IGHV3-53*02, or IGHV3-53*01		IGHJ6*02	CARDLYYYGMDVW	13
IGHV3-53*02, or IGHV3-53*01		IGHJ6*02	CARDLYYYGMDVW	13
IGHV3-53*02, or IGHV3-53*01		IGHJ6*02	CARDLYYYGMDVW	13
IGHV3-21*04		IGHJ6*02	CARDLYYYGMDVW	13
IGHV3-64*01		IGHJ6*02	CARDLYYYGMDVW	13
IGHV3-53*04		IGHJ6*02	CARDLYYYGMDVW	13

Part III:
AIRR Data Commons Tool Demos

AIRR-seq Data Curation

Curating an AIRR-seq study for the ADC

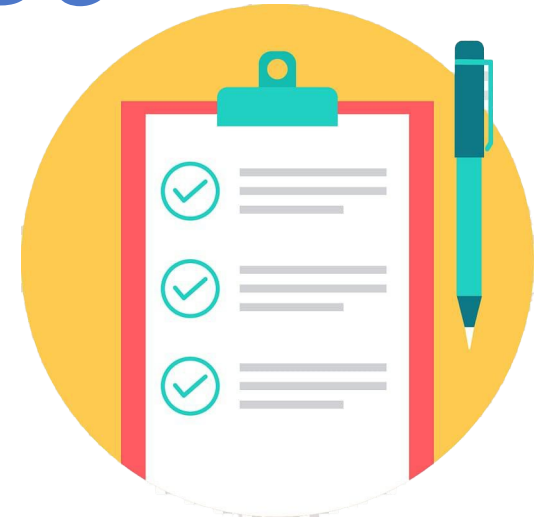
What you need...

Rearrangement annotation files

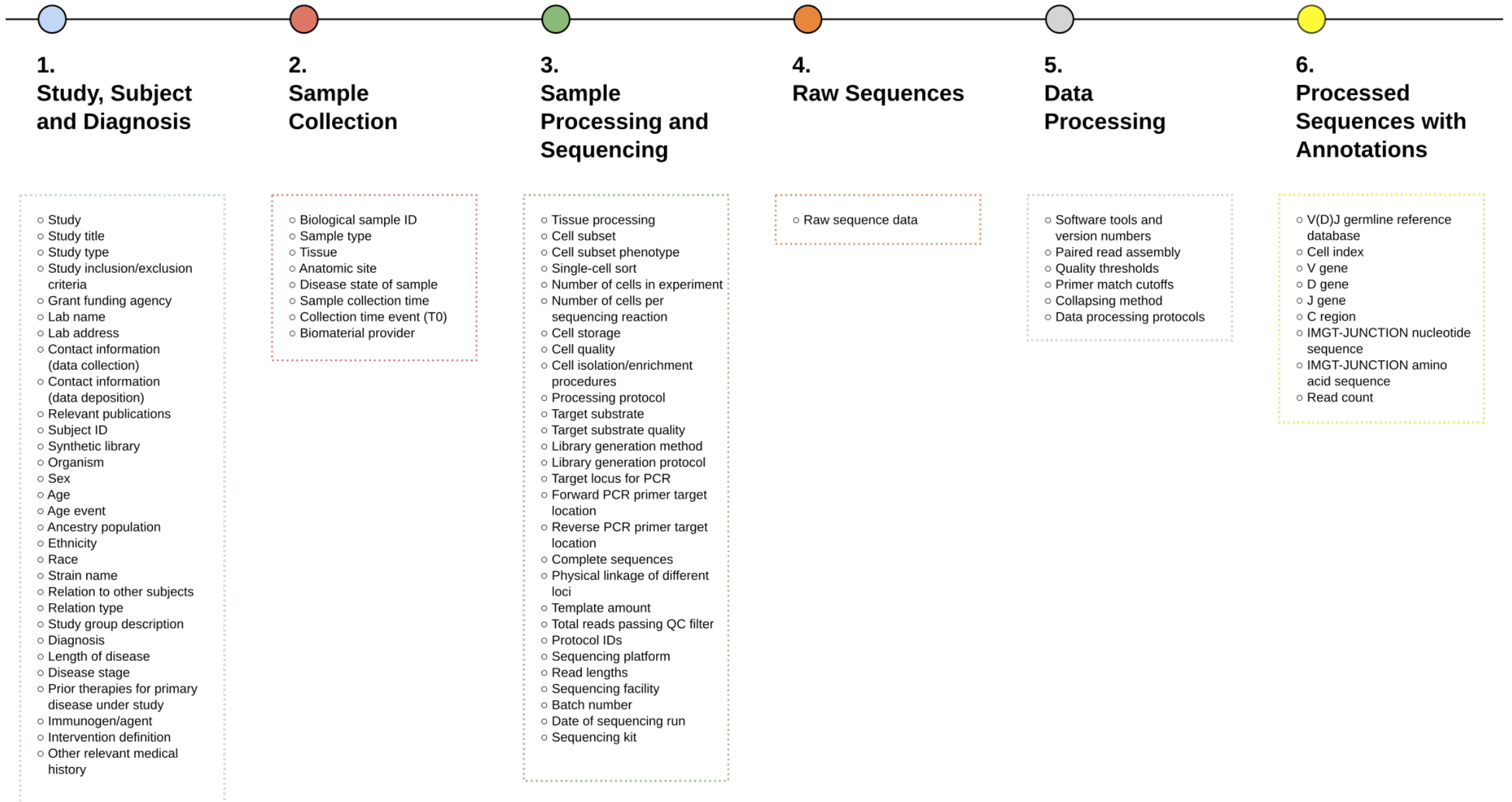
- One per repertoire, in AIRR.tsv format
- <https://docs.airr-community.org/en/stable/datarep/format.html>

Metadata

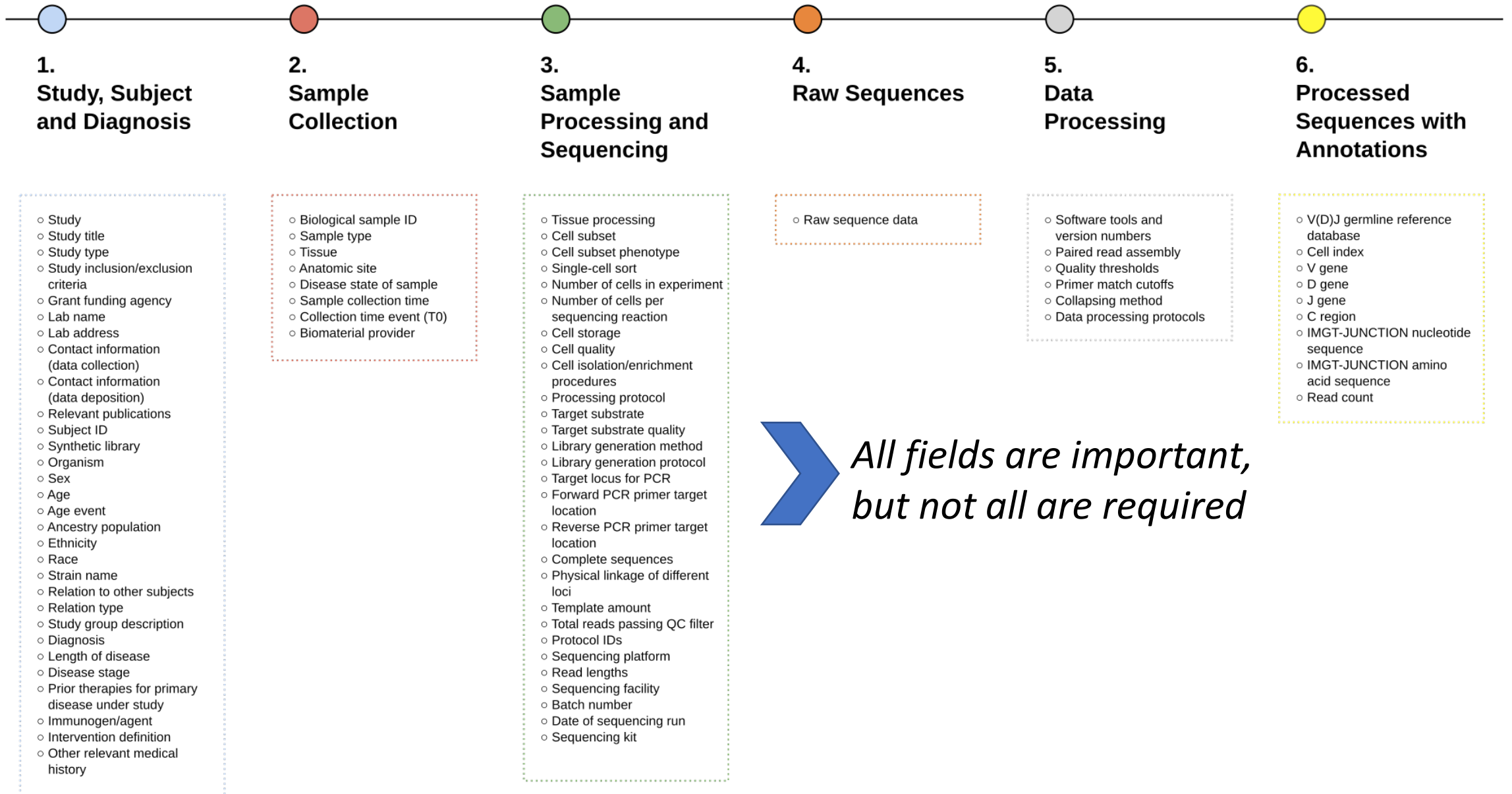
- One file for the study, following MiAIRR Reporting Standards
- https://docs.airr-community.org/en/stable/miairr/data_elements.html



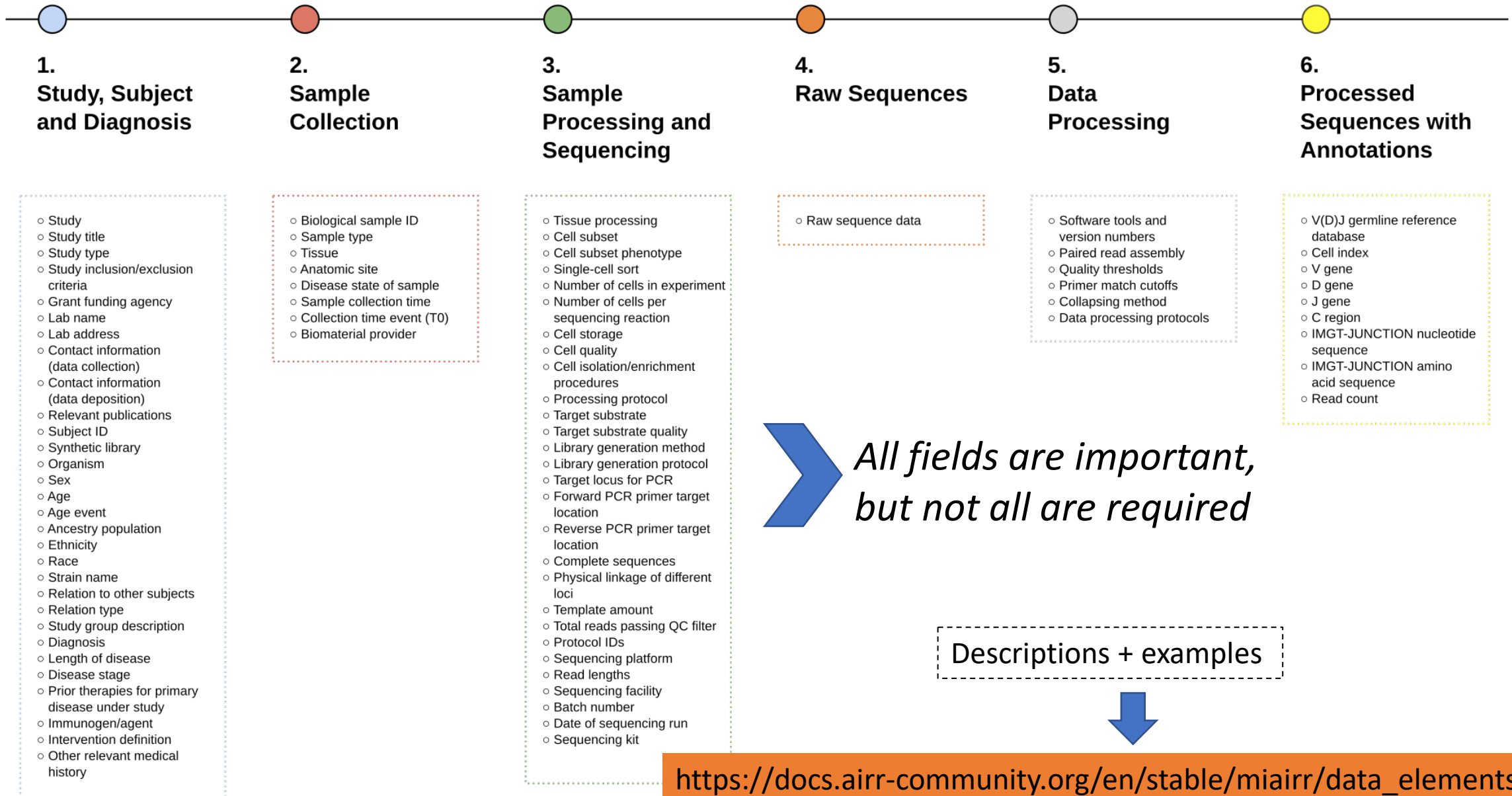
Curating an AIRR-seq study: MiAIRR



Curating an AIRR-seq study: MiAIRR



Curating an AIRR-seq study: MiAIRR



Curating an AIRR-seq Study: Resources

- <https://github.com/sfu-ireceptor/dataloading-curation>

☰ README.md

iReceptor Data Curation

This GIT repository contains example files and documentation for loading data into iReceptor repositories. Examples for metadata as well as rearrangement files for a number of widely used annotation tools are provided. The README files in each of the subfolders contain more documentation.

For more information on metadata curation, please refer to:

- [The iReceptor Metadata documentation](#)
- [The iReceptor metadata example](#)
- [The AIRR repertoire metadata example](#)

For more details on rearrangement curation, please refer to:

- [The test data set documentation](#)
- [The AIRR rearrangement format \(including igblast\) example](#)
- [The MiXCR rearrangement format example](#)
- [The IMGT V-QUEST rearrangement example](#)

Curating an AIRR-seq Study: Resources

- <https://github.com/sfu-ireceptor/dataloading-curation>

☰ README.md

iReceptor Data Curation

This GIT repository contains example files and documentation for loading data into iReceptor repositories. Examples for metadata as well as rearrangement files for a number of widely used annotation tools are provided. The README files in each of the subfolders contain more documentation.

For more information on metadata curation, please refer to:

- [The iReceptor Metadata documentation](#)
- [The iReceptor metadata example](#)



	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	study_id	study_title	study_type	study_type_id	study_description	inclusion_exclusion_criteria	grants	collected_by	lab_name	lab_address	submitted_by	pub_ids	keywords_study
2	PRJNA188191	Mining the antibody	Study	NCIT:C63536	HIV+ Study			pdkwong@nih.gov, Is	Kwong Lab	National Institute of Allergy and Infectious Diseases	P.D. Kwong, J. Zhu	PMID: 23536288	contains_ig
3	PRJNA206548	Immunoglobulin gene	Case-Control Study	NCIT:C15197	Cancer Study			ramit.mehr@biu.ac.il	Ramit Mehr's Computational Immunology Lab	Bar Ilan University	M. Michaeli, H. Tadmor	PMID: 24917868	contains_ig
4	PRJNA321261	Identification of shared	Case-Control Study	NCIT:C15197	Cancer Study	Inclusion breast cancer	Department of Defense	kapplerj@njhealth.org	Department of Immunology	Department of Immunology	D. J. Munson, C.A. Janeway	PMID: 27307436	contains_tcr
5	PRJNA312319	Tracking T-cell immune	Case-Control Study	NCIT:C15197	Cancer Study	HST transplant	Russian Science Foundation	chudakovdm@mail.ru	Chudakov Lab Shemya	Russian Academy of Sciences	I.V. Zvyagin, I.Z. Nizhnik	PMID: 27811849	contains_tcr
6	PRJNA316033	T cell Beta chain rearr	Study	NCIT:C63536	Cancer Study	Patients with esophageal	Natural Science Foundation of China	keyang@bjmu.edu.cn	Key Laboratory of Cancer	Key Laboratory of Cancer	Z. Chen, C. Zhang	PMID: 27171315	contains_tcr
7	PRJNA356992	Multi-omics study of	Study	NCIT:C63536	Cancer Study	primary liver carcinoma	National Key Research and Development	chengshj@263.net.cn	State Key Lab of Molecular	National Cancer Centre	L. Shi, J. Wu, K. Zhang	PMID: 28422742	contains_tcr
8	string	string	ontology	ontology	string	string	string	string	string	string	string	string	array
9	Unique ID assigned by study repository	Descriptive study title	Type of study design	Type of study design (Ontology)	Generic study description	List of criteria for inclusion/exclusion	Funding agencies and grant numbers	Full contact information	Department of data collection	Institution and institution address	Full contact information	Publications describing the study	Keywords describing the study

Curating an AIRR-seq Study: Resources

- <https://github.com/sfu-ireceptor/dataloading-curation>

The iReceptor Data Curation process

The iReceptor team follows a relatively strict data curation process. This process is documented on the [iReceptor Curation page](#). We do not discuss this process in detail here, but instead suggest simple processes that can make data curation easier to manage.

The iReceptor curation process is focused around the curation of data for a single study. As such, we recommend that all data that is being curated for a specific study be stored in a single directory. As an example, we will use one of the [IMGT example data sets](#).

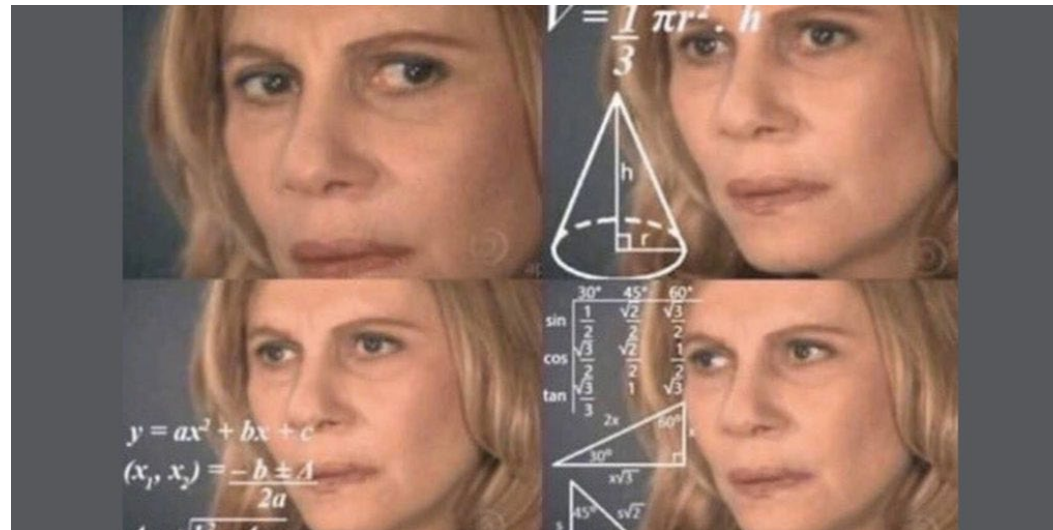
Curating an AIRR-seq Study: Resources

- <https://github.com/sfu-ireceptor/dataloading-curation>

The iReceptor Data Curation process

The iReceptor team follows a relatively strict data curation process. This process is documented on the [iReceptor Curation page](#). We do not discuss this process in detail here, but instead suggest simple processes that can make data curation easier to manage.

The iReceptor curation process is focused around the curation of data for a single study. As such, we recommend that all data that is being curated for a specific study be stored in a single directory. As an example, we will use one of the [IMGT example data sets](#).



Need help? Email us:
support@ireceptor.org
kira_neller@sfu.ca

Curating an AIRR-seq study: Common Pitfalls

Issue	Why is this a problem?

Curating an AIRR-seq study: Common Pitfalls

Issue	Why is this a problem?
Not providing annotations in an AIRR TSV format (or a format that can be converted to AIRR TSV).	Non AIRR-compliant data cannot be loaded into an ADC repository. Some repository data loaders (e.g. iReceptor Turnkey) accept other annotation tool formats (MiXCR, IMGT Vquest, igblast/AIRR TSV).

Curating an AIRR-seq study: Common Pitfalls

Issue	Why is this a problem?
Not providing annotations in an AIRR TSV format (or a format that can be converted to AIRR TSV).	Non AIRR-compliant data cannot be loaded into an ADC repository. Some repository data loaders (e.g. iReceptor Turnkey) accept other annotation tool formats (MiXCR, IMGT Vquest, igblast/AIRR TSV).
Only providing processed sequences (e.g. CDR3s), not raw sequences	Hinders data re-use and verification

Curating an AIRR-seq study: Common Pitfalls

Issue	Why is this a problem?
Not providing annotations in an AIRR TSV format (or a format that can be converted to AIRR TSV).	Non AIRR-compliant data cannot be loaded into an ADC repository. Some repository data loaders (e.g. iReceptor Turnkey) accept other annotation tool formats (MiXCR, IMGT Vquest, igblast/AIRR TSV).
Only providing processed sequences (e.g. CDR3s), not raw sequences	Hinders data re-use and verification
Using incorrect formatting for miAIRR fields in metadata file	You shall not pass! The data loader will not load your data. Be mindful of different formatting types (string, ontology, integer, controlled vocabulary, boolean,)

Curating an AIRR-seq study: Common Pitfalls

Issue	Why is this a problem?
Not providing annotations in an AIRR TSV format (or a format that can be converted to AIRR TSV).	Non AIRR-compliant data cannot be loaded into an ADC repository. Some repository data loaders (e.g. iReceptor Turnkey) accept other annotation tool formats (MiXCR, IMGT Vquest, igblast/AIRR TSV).
Only providing processed sequences (e.g. CDR3s), not raw sequences	Hinders data re-use and verification
Using incorrect formatting for miAIRR fields in metadata file	You shall not pass! The data loader will not load your data. Be mindful of different formatting types (string, ontology, integer, controlled vocabulary, boolean,)
Not including all data processing details (tools and versions, QC/filtering steps, reference genomes, etc.)	Hinders data re-use and verification

Curating an AIRR-seq study: Common Pitfalls

Issue	Why is this a problem?
Not providing annotations in an AIRR TSV format (or a format that can be converted to AIRR TSV).	Non AIRR-compliant data cannot be loaded into an ADC repository. Some repository data loaders (e.g. iReceptor Turnkey) accept other annotation tool formats (MiXCR, IMGT Vquest, igblast/AIRR TSV).
Only providing processed sequences (e.g. CDR3s), not raw sequences	Hinders data re-use and verification
Using incorrect formatting for miAIRR fields in metadata file	You shall not pass! The data loader will not load your data. Be mindful of different formatting types (string, ontology, integer, controlled vocabulary, boolean,)
Not including all data processing details (tools and versions, QC/filtering steps, reference genomes, etc.)	Hinders data re-use and verification
Not using the same subject/sample identifiers as in the published paper	Others cannot (easily) compare the dataset to the paper

Curating an AIRR-seq study: Common Pitfalls

Issue	Why is this a problem?
Not providing annotations in an AIRR TSV format (or a format that can be converted to AIRR TSV).	Non AIRR-compliant data cannot be loaded into an ADC repository. Some repository data loaders (e.g. iReceptor Turnkey) accept other annotation tool formats (MiXCR, IMGT Vquest, igblast/AIRR TSV).
Only providing processed sequences (e.g. CDR3s), not raw sequences	Hinders data re-use and verification
Using incorrect formatting for MiAIRR fields in metadata file	You shall not pass! The data loader will not load your data. Be mindful of different formatting types (string, ontology, integer, controlled vocabulary, boolean)
Not including all data processing details (tools and versions, QC/filtering steps, reference genomes, etc.)	Hinders data re-use and verification
Not using the same subject/sample identifiers as in the published paper	Others cannot (easily) compare the dataset to the paper
Not providing all available metadata	Though your study may not have considered certain variables (e.g. age), others may need these for data re-use

Part III:
AIRR Data Commons Tool Demos

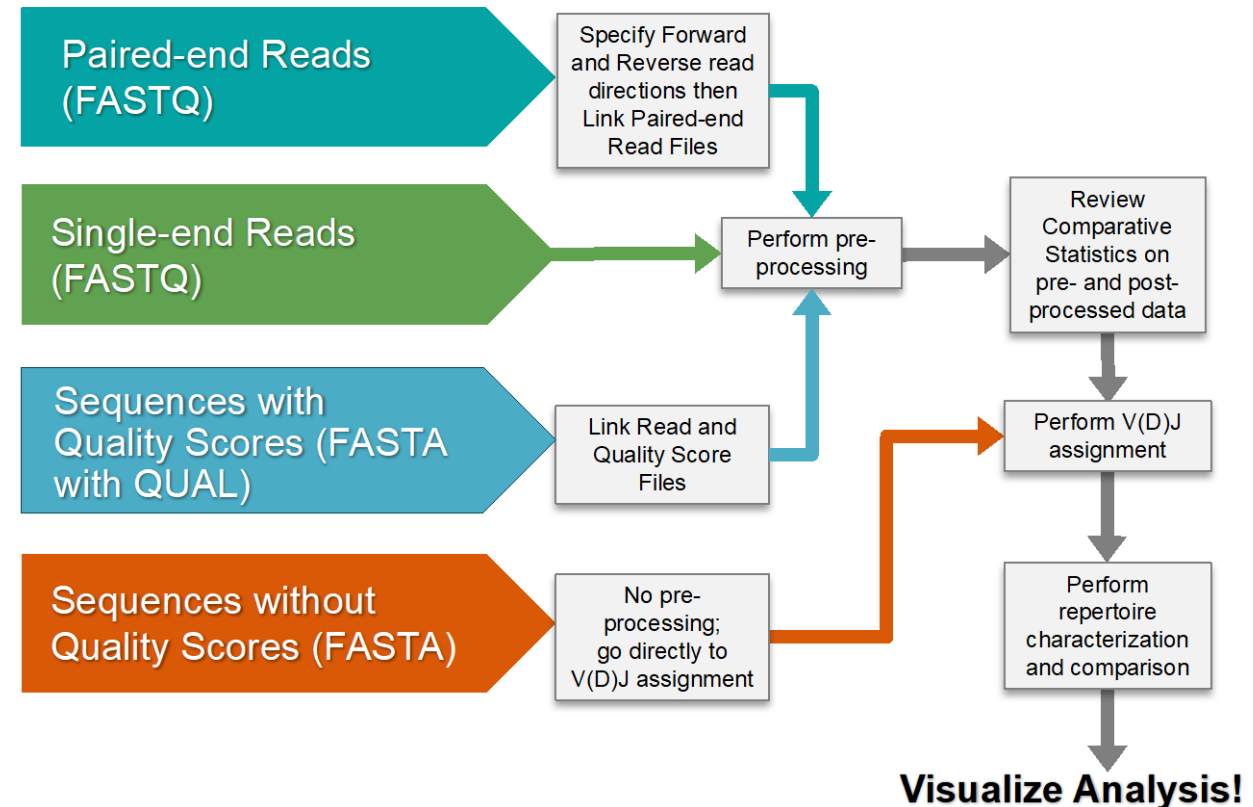
VDJServer Analysis Portal



VDJServer: A Cloud-Based Analysis Portal and Data Commons for Immune Repertoire Sequences and Rearrangements

- NIAID-recommended service for immune repertoire analysis (R01 AI097403)
- Web browser interface: <https://vdjserver.org>
- Free Account
- Provides Immcantation tool suite and other specialized analysis tools
- MiAIRR study metadata entry
- Immune repertoire analysis
 - Pre-processing and quality control
 - Pre-processing visualizations
 - V(D)J gene assignment
 - Repertoire characterization and comparison
 - Analysis visualizations
- Community Data Portal
- Tool Ecosystem

Workflow Overview: What type of read level data do you have?



VDJServer Community Data Portal

- Users can publish their studies making the data, metadata, analyses and visualizations available to the public.
 - Users can privately share project with other VDJServer users.
 - Users can publicly share by publishing to the Community Data Portal.
 - Not necessary to run analyses on VDJServer in order to share data.
- Sharing data can be as easy as 1, 2, 3:
 1. Create project
 2. Upload files to project
 3. Publish project

VDJServer Community Data Portal

- VDJServer is a data repository in the AIRR Data Commons
 - 20+ studies, 2.5B rearrangement records
 - ADC API: <https://vdjserver.org/airr/v1>
 - Almost all studies are processed from SRA/ENA raw read data with VDJServer's analysis pipeline for consistency.
- VDJServer V2 brings new capabilities
 - MiAIRR data entry (AIRR Repertoire metadata) for studies
 - Integrated with iReceptor Plus platform
 - Users can perform comparative analysis between their private data and data queried from the AIRR Data Commons
 - Tracking the draft AIRR standards
 - Clones, trees, single cell, data processing, germline gene sets

VDJServer Community Data Portal

- Publish study in the AIRR Data Commons with VDJServer
 - Additional steps beyond publishing to the AIRR Community Data Portal
 - Data needs to be compliant with AIRR Standards
 - VDJServer V1 projects require conversion
 - VDJServer V2 projects will be compliant automatically
 - Data needs to be processed through VDJServer's analysis pipeline.
- Based upon the size of the data, publishing can take hours or days
 - VDJServer admin performs the database load
 - Contact us to publish your study: vdjserver@utsouthwestern.edu

VDJServer Community Data Portal

- Video demo: <https://youtu.be/DgqKP-L75CM>

Useful References

- Websites
 - AIRR Community: www.airr-community.org
 - iReceptor Website: www.ireceptor.org
 - iReceptor Scientific Gateway: gateway.ireceptor.org
 - VDJServer: <https://vdjserver.org/>
 - iReceptor Plus: www.ireceptorplus.org
- GitHub
 - <https://github.com/sfu-ireceptor>
 - <https://github.com/airr-community>
- Publications
 - iReceptor paper: Immunological Reviews – (DOI: 10.1111/imr.12666)
 - AIRR Community paper: Frontiers in Immunology – (DOI: 10.3389/fimmu.2017.01418)
 - AIRR Data Standards paper: Nature Immunology – (DOI: 10.1038/ni.3873)
 - AIRR Data Rep paper: Frontiers in Immunology – (DOI: 10.3389/fimmu.2018.02206)

Who to contact, How to Access

- AIRR Community:
 - Contact: comm@airrc.antibodysociety.org
 - AIRR Data Commons: <https://www.antibodysociety.org/the-airr-community/airr-data-commons/>
 - AIRR documentation: <http://docs.airr-community.org>
- iReceptor:
 - Contact: support@ireceptor.org
 - iReceptor Gateway account/access (<https://gateway.ireceptor.org>)
 - Download the iReceptor Turnkey repository (<https://github.com/sfu-ireceptor/turnkey-service-php>)
- VDJServer:
 - Contact: vdjserver@utsouthwestern.edu
 - VDJServer portal account: <https://vdjserver.org>

Acknowledgements

- Colleagues in the AIRR Community
- Collaborators across iReceptor Plus
- Funders
 - iReceptor: CANARIE, Canada Foundation for Innovation, CIHR, BC Knowledge Development Fund
 - VDJServer: National Institute of Allergy and Infectious Diseases (#1R01A1097403)
 - iReceptor Plus: EU Horizon 2020 Research and Innovation Programme (#825821)
- Platform providers
 - Compute Canada
 - Texas Advanced Computing Centre