AIRR Community Meeting VI: Exploring New Frontiers

May 17-19, 2022,
La Jolla, San Diego (CA), United States of America

Gold sponsors

Silver sponsors

Bronze sponsors

Institutional supports

Table of contents

Welcome from the AIRR-C Meetings Sub-Committee and the AIRR-C Executive Chair 2

Sponsors 4

The AIRR Community online 5

The AIRR Community on social media 5

The AIRR Community Working Groups and Sub-Committees contact information 6

Agenda at glance 7

Full agenda 8

Attending Virtually 11

Invited Speakers 12

Mentoring session 15

Industry Networking Reception 16

Travel Award Recipients 17

Lightning Demonstrations 18

Deep Dive Tutorials 21

Posters 25

AIRR-C Committee Sub-committees and Working Groups Reports 51

AIRR Community Meeting VI Attendees 95

Welcome from the AIRR-C Meetings Sub-Committee and the AIRR-C Executive Chair

Welcome to the sixth meeting of the Adaptive-Immune Receptor Repertoire (**AIRR**) Community of The Antibody Society, hosted at Hilton La Jolla Torrey Pines, La Jolla, California - AIRR Community Meeting VI: Exploring New Frontiers!

The AIRR Community is committed to developing standards and/or recommendations for:

i) Generating, analyzing, curating and sharing AIRR-seq data;
ii) Using and validating tools for analyzing AIRR-seq data;
iii) Relating AIRR-seq data to other "big data" types, such as microarray, flow cytometric, and single-cell gene-expression data; and
iv) Legal and ethical matters associated with the use and sharing of AIRR-seq data derived from human sources.

Through its Working Groups and meetings, the AIRR Community has developed and published recommendations and action plans to maximize the generation, use and sharing of AIRR-seq data within the scientific community for the benefit of humanity.

AIRR Community meetings are the premier event for research on adaptive immune-receptor repertoires. They are also the primary location where the AIRR Community's Working Groups and Sub-committees come together in one location to discuss how to push standardization in AIRR-sequencing (AIRR-seq) data and analysis forward. All attendees are welcome to participate in the General Assembly on Wednesday morning, when work plans, submitted by the Working Groups and Sub-committees for the coming year and meeting interval, will be ratified.

This years' meeting aims at ***Exploring New Frontiers*** through two themed "Challenge Sessions" meant to:

(i) ***The AIRRC challenge session***: Initiate and implement a strategic plan for the AIRR Community that integrates the Working Groups' activities toward the central goal of universally accepted AIRR-seq data standards;
(ii) ***The scientific challenge session - Systems Immunology***: Introduce the Community to multi-dimensional systems approaches for characterizing immune responses and how AIRR-seq data can benefit such approaches.

Through original activities, the AIRRC VI Meeting provided opportunities for investigators and trainees to network (***Mentoring session***), to participate in AIRR Community Working Groups and Sub-committee meetings (***Report session***), and to learn in workshops reviewing the immune system and AIRR-seq data generation and analysis (Software and tools – lightening demonstration & Software and tools – Deep Dive Tutorial sessions).

Last but not least, the AIRR-C Basic and Biomedical scientific sessions gathered an outstanding line up, including keynote lectures by Dennis Burton, Gunila Karlsson Hedestam, Shane Crotty and Atul Butte, invited presentations and short presentations chosen from the submitted poster abstracts.

For this AIRR-C VI meeting, the AIRR-C was awarded a NIAID/NIH travel grant to help support early career researchers and students to attend the AIRR Community VI "Exploring New Frontiers" Meeting, targeting 1/ Early Career Travel Awards and 2/ Local Access Funds. In both cases, applicants needed to be a junior faculty members or students/post-docs, with a priority given to scientists belonging to a minority-serving institution in the USA or a research institution

from low-to-middle income countries (award 1) or to a regional minority-serving institutions (award 2).

All meeting documents and video recordings can be found on the AIRR Community YouTube channel here:
https://www.antibodysociety.org/the-AIRRCommunity/meetings/AIRRCommunity-meeting-vi-exploring-new-frontiers/

We look forward to your participation and invite you to become an active member of the AIRR Community, if you aren't one already! Consider joining a Working Group or Subcommittee. (Use the links to learn more about them.) It's a great way to get involved in an interdisciplinary network of basic and biomedical scientists, bioinformaticians, ethicists and legal experts from academia and industry who are all working toward a greater good!

Lindsay Cowell (University of Texas Southwestern Medical Center), Chair, AIRR-C Executive Sub-committee
Victor Greiff (University of Oslo), Chair-Elect, AIRR-C Executive Sub-committee

And the Members of the Meeting Sub-committee:
Justin Barton (University of London)
Pam Borghardt (Simon Fraser University, Co-leader & Meeting Manager)
Encarnita Mariotti-Ferrandiz (Sorbonne Université, Co-leader)
Corey Watson (University of Louisville, USA)

Thanks to our amazing volunteers!
Kira Neller (Simon Fraser University)
Kenz Le Gouge (Sorbonne Université)
Paul Stys (Sorbonne Université)

Sponsors

The AIRR Community warmly thanks all of our meeting sponsors for their generous financial support. It is their contribution that made this meeting possible, laying the foundation for an exciting and stimulating scientific exchange.

Please take a moment and inform yourself about the ground-breaking work and services our sponsors provide and make sure to engage with them during the Industry Networking session and in our ***Science, Tools & Technology lunch session*** on Thursday, May 19th.

You will find below a list of the meeting sponsors and access to their websites.

Sponsor	**Website**	**Level**
Simon Fraser University	www.sfu.ca	Institutional Award
National Health Institute	www.nih.gov	Institutional Award
Enpicom	enpicom.com	Gold
Illumina	www.illumina.com	Gold
Omniscope	www.omniscope.ai	Gold
Roche	www.roche.com	Gold
Takara	www.takarabio.com	Gold
Alchemab	www.alchemab.com	Silver
The Antibody Society	www.antibodysociety.org	Silver
Distributed bio	www.distributedbio.com	Silver
INESC TEC	www.inesctec.pt/en	Silver
PipeBio	pipebio.com	Silver
Abterra bio	abterrabio.com	Bronze
Adaptive biotechnologies	www.adaptivebiotech.com	Bronze
Agora Partners	agora-partners.com	Bronze
BISC Global, Inc.	www.biscglobal.com	Bronze
Interteam	interteam.co.il	Bronze
Medgenome	research.medgenome.com	Bronze

The AIRR Community online

know more about the AIRRC

join the AIRRC

access AIRR Data Commons

On AIRR an AIRRC podcast

AIRRC youtube channel

AIRRC merchandizing

The AIRR Community on social media

https://twitter.com/airr_community

https://www.linkedin.com/company/the-AIRRCommunity/

The AIRR Community Working Groups and Sub-Committees contact information

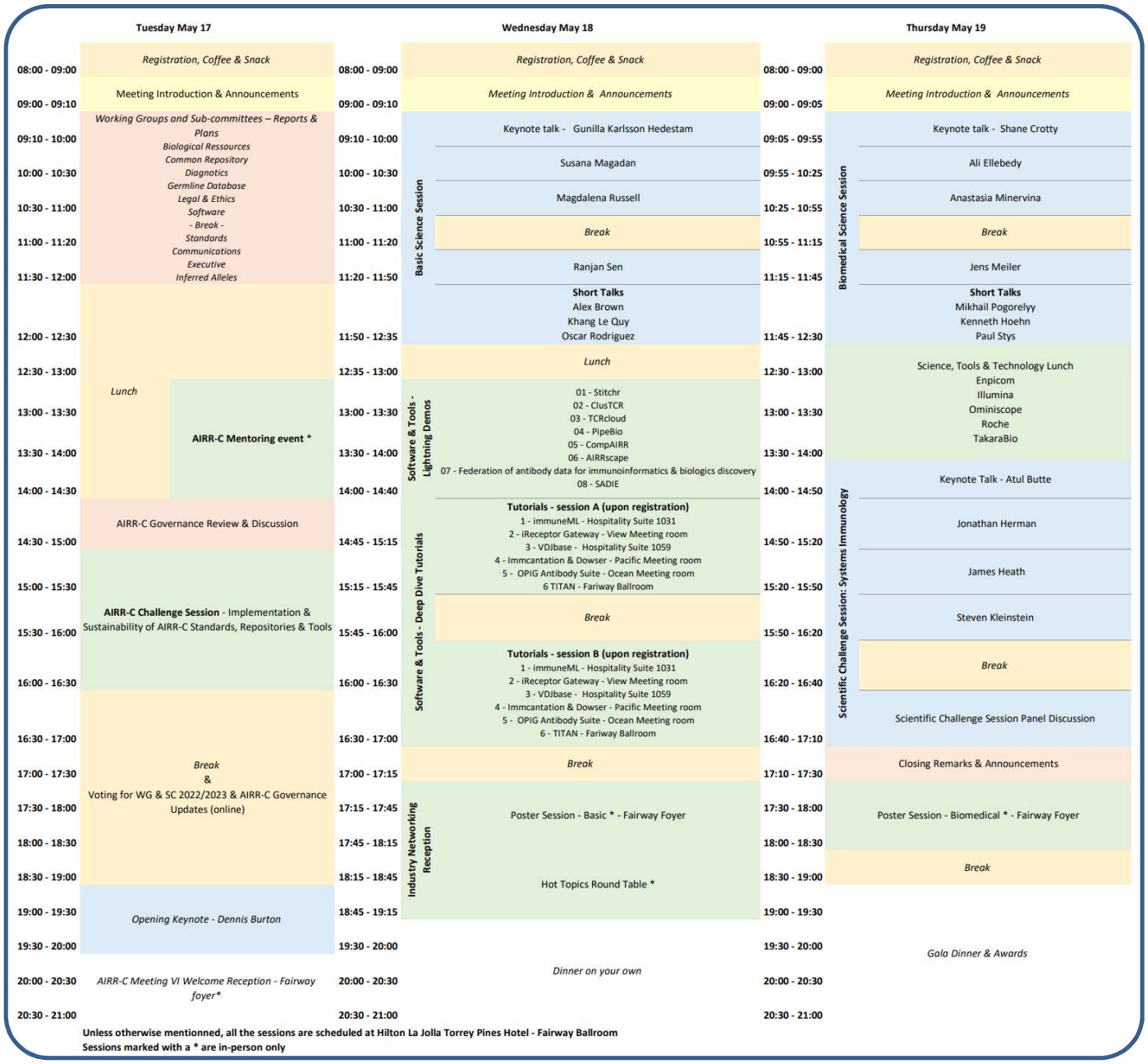
Communications SC: communications@AIRRCommunity.org
Inferred Allele Review SC: iarc@AIRRCommunity.org
Meetings SC: meetings@AIRRCommunity.org

Biological Resources WG: biological-resources@AIRRCommunity.org
Common Repository WG: common-repository@AIRRCommunity.org
Data Representation WG: datarep@AIRRCommunity.org
Diagnostics WG: diagnostics@AIRRCommunity.org
Germline WG: germline-database@AIRRCommunity.org
Legal & Ethics WG: legal-ethics@AIRRCommunity.org
Software WG: software@AIRRCommunity.org
Standards WG: standards@AIRRCommunity.org

Agenda at glance

May 17th - 19th 2022

Hilton La Jolla Torrey Pines Hotel – Fairway Ballroom



Tuesday May 17

08:00 - 09:00	*Registration, Coffee & Snack*
09:00 - 09:10	Meeting Introduction & Announcements
09:10 - 10:00	*Working Groups and Sub-committees – Reports & Plans*
10:00 - 10:30	*Biological Ressources*
10:30 - 11:00	*Common Repository*
11:00 - 11:20	*Diagnotics*
11:30 - 12:00	*Germline Database*
	Legal & Ethics
	Software
	- Break -
	Standards
	Communications
	Executive
	Inferred Alleles
12:00 - 12:30	*Lunch*
12:30 - 13:00	
13:00 - 13:30	**AIRR-C Mentoring event ***
13:30 - 14:00	
14:00 - 14:30	
14:30 - 15:00	AIRR-C Governance Review & Discussion
15:00 - 15:30	**AIRR-C Challenge Session** - Implementation & Sustainability of AIRR-C Standards, Repositories & Tools
15:30 - 16:00	
16:00 - 16:30	
16:30 - 17:00	
17:00 - 17:30	*Break* & Voting for WG & SC 2022/2023 & AIRR-C Governance Updates (online)
17:30 - 18:00	
18:00 - 18:30	
18:30 - 19:00	
19:00 - 19:30	*Opening Keynote - Dennis Burton*
19:30 - 20:00	
20:00 - 20:30	*AIRR-C Meeting VI Welcome Reception - Fairway foyer**
20:30 - 21:00	

Wednesday May 18

08:00 - 09:00	*Registration, Coffee & Snack*	
09:00 - 09:10	*Meeting Introduction & Announcements*	
09:10 - 10:00	Keynote talk - Gunilla Karlsson Hedestam	Basic Science Session
10:00 - 10:30	Susana Magadan	
10:30 - 11:00	Magdalena Russell	
11:00 - 11:20	*Break*	
11:20 - 11:50	Ranjan Sen	
11:50 - 12:35	**Short Talks** Alex Brown Khang Le Quy Oscar Rodriguez	
12:35 - 13:00	*Lunch*	
13:00 - 13:30	01 - Stitchr 02 - ClusTCR 03 - TCRcloud 04 - PipeBio 05 - CompAIRR 06 - AIRRscape 07 - Federation of antibody data for immunoinformatics & biologics discovery 08 - SADIE	Software & Tools - Lightning Demos
13:30 - 14:00		
14:00 - 14:40		
14:45 - 15:15	**Tutorials - session A (upon registration)** 1 - immuneML - Hospitality Suite 1031 2 - iReceptor Gateway - View Meeting room 3 - VDJbase - Hospitality Suite 1059 4 - Immcantation & Dowser - Pacific Meeting room 5 - OPIG Antibody Suite - Ocean Meeting room 6 TITAN - Fariway Ballroom	Software & Tools - Deep Dive Tutorials
15:15 - 15:45		
15:45 - 16:00	*Break*	
16:00 - 16:30	**Tutorials - session B (upon registration)** 1 - immuneML - Hospitality Suite 1031 2 - iReceptor Gateway - View Meeting room 3 - VDJbase - Hospitality Suite 1059 4 - Immcantation & Dowser - Pacific Meeting room 5 - OPIG Antibody Suite - Ocean Meeting room 6 - TITAN - Fariway Ballroom	
16:30 - 17:00		
17:00 - 17:15	*Break*	
17:15 - 17:45	Poster Session - Basic * - Fairway Foyer	Industry Networking Reception
17:45 - 18:15		
18:15 - 18:45	Hot Topics Round Table *	
18:45 - 19:15		
19:30 - 20:00	*Dinner on your own*	
20:00 - 20:30		
20:30 - 21:00		

Thursday May 19

08:00 - 09:00	*Registration, Coffee & Snack*	
09:00 - 09:05	*Meeting Introduction & Announcements*	
09:05 - 09:55	Keynote talk - Shane Crotty	Biomedical Science Session
09:55 - 10:25	Ali Ellebedy	
10:25 - 10:55	Anastasia Minervina	
10:55 - 11:15	*Break*	
11:15 - 11:45	Jens Meiler	
11:45 - 12:30	**Short Talks** Mikhail Pogorelyy Kenneth Hoehn Paul Stys	
12:30 - 13:00	Science, Tools & Technology Lunch Enpicom Illumina Ominiscope Roche TakaraBio	
13:00 - 13:30		
13:30 - 14:00		
14:00 - 14:50	Keynote Talk - Atul Butte	Scientific Challenge Session: Systems Immunology
14:50 - 15:20	Jonathan Herman	
15:20 - 15:50	James Heath	
15:50 - 16:20	Steven Kleinstein	
16:20 - 16:40	*Break*	
16:40 - 17:10	Scientific Challenge Session Panel Discussion	
17:10 - 17:30	Closing Remarks & Announcements	
17:30 - 18:00	Poster Session - Biomedical * - Fairway Foyer	
18:00 - 18:30		
18:30 - 19:00	*Break*	
19:00 - 19:30		
19:30 - 20:00	*Gala Dinner & Awards*	
20:00 - 20:30		
20:30 - 21:00		

Unless otherwise mentionned, all the sessions are scheduled at Hilton La Jolla Torrey Pines Hotel - Fairway Ballroom
Sessions marked with a * are in-person only

Full agenda

May 17th - 19th 2022

Hilton La Jolla Torrey Pines Hotel – Fairway Ballroom

To facilitate global engagement and increase accessibility parts of the AIRR Community Meeting will be live streamed and recorded for later viewing

Monday, May 16th 2022		
Start	**End**	**Activities**
8:00	9:00	*Registration open*
8:00	16:00	Prep meetings for Working Groups and Sub-committees and Pre-meetings for International Partners
15:00	16:00	*Registration open*
16:00	17:00	Informal Social time
		Dinner on your own

Tuesday, May 17th 2022		
Start	**End**	**Activities**
8:00	9:00	*Registration open & coffee*
9:00	12:30	**Introduction + Announcements** by AIRRC Chair, Executive Sub-committee - Lindsay Cowell, University of Texas Southwestern Medical Center (USA) **Working Groups and Sub-committee – Reports & Plans** Biological Resources WG Common Repository WG Diagnostics WG Germline Database WG Legal and Ethics WG Software WG Standards WG Communications SC Executive SC Inferred Allele SC Meetings SC **Voting for adoption of plans**
12:30	13:00	*Lunch*
13:30	14:30	**Student mentoring session** Moderator, Encarnita Mariotti-Ferrandiz, Sorbonne Univ. (France)
14:30	15 :00	**AIRRC Governance Review & Discussion:** Christian Busse, DKFZ (Germany) & Lindsey Cowell, University of Texas Southwestern Medical Center (USA) Introducing and voting the governance updates of the AIRRC
15:00	16:00	**AIRRC Challenge Session:** Moderator, Victor Greiff, University of Oslo (Norway) Implementation & Sustainability of AIRRC Standards, Repositories & Tools
		Informal Social time
19:00	20:00	**Opening Keynote:** Dennis Burton, Scripps Research (USA) Introduction by Gunilla Karlsson Hedestam, Karolinska Institutet (Sweden)
20:00	21:00	**AIRRC Meeting VI Welcome Reception**

Wednesday, May 18th 2022		
Start	**End**	**Activities**
8:00	8:30	*Registration open*
8:30	9:00	*Coffee*
9:00	9:10	**Introduction to the Meeting + Announcements** by AIRRC Chair, Executive Sub-committee - Lindsay Cowell, University of Texas Southwestern Medical Center (USA)
9:10	12:35	**Basic Science Session**, Steve Kleinstein, Yale University (USA), Moderator
9:10	10:00	**Keynote:** Gunilla Karlsson Hedestam, Karolinska Institutet (Sweden) (40-min talk + 10-min Q&A)
10:00	10:30	**Invited Talk1:** Susana Magadan, Univ. of Vigo (Spain) and Univ. of New Mexico (USA)
10:30	11:00	**Invited Talk2:** Magdalena L Russell, Fred Hutch, University of Washington (USA)
11:00	11:20	*Break*
11:20	11:50	**Invited Talk3:** Ranjan Sen, National Institute on Aging, NIH (USA)
11:50	12:35	**Contributed Talks** *Alex Brown, University of Colorado & National Jewish Health (USA)* *Khang Le Quy, University of Oslo (NO)* *Oscar Rodriguez, University of Louisville School of Medicine (UAS)*
13:00	14:30	*Registration open*
13:00	14:30	*Lunch* + **Software & Tool - Lightning Demos** Moderator : Justin Barton, University of London, (UK) Lightning Demo 01 - Stitchr Lightning Demo 02 - ClusTCR Lightning Demo 03 - TCRcloud Lightning Demo 04 - PipeBio Lightning Demo 05 - CompAIRR Lightning Demo 06 - AIRRscape Lightning Demo 07 - Federation of antibody data for immunoinformatics & biologics discovery Lightning Demo 08 - SADIE
14:45	17:00	**Software & Tool – Deep Dive Tutorials**, Moderator : Justin Barton, University of London, (UK) Pre-registration is required via registration site; attendees can participate in 2 x 60-min sessions separated by 15 mins break) *Tutorial 1 - immuneML* *Tutorial 2 - iReceptor Gateway* *Tutorial 3 - VDJbase* *Tutorial 4 - Immcantation & Dowser* *Tutorial 5 - OPIG Antibody Suite* *Tutorial 6 - TITAN*
17:15	19:15	**Industry Networking Reception & Poster Session I**, Encarnita Mariotti-Ferrandiz, Sorbonne Université (France) & Andrew Farmer, Takara Bio (USA) Moderators.
		Dinner on your own

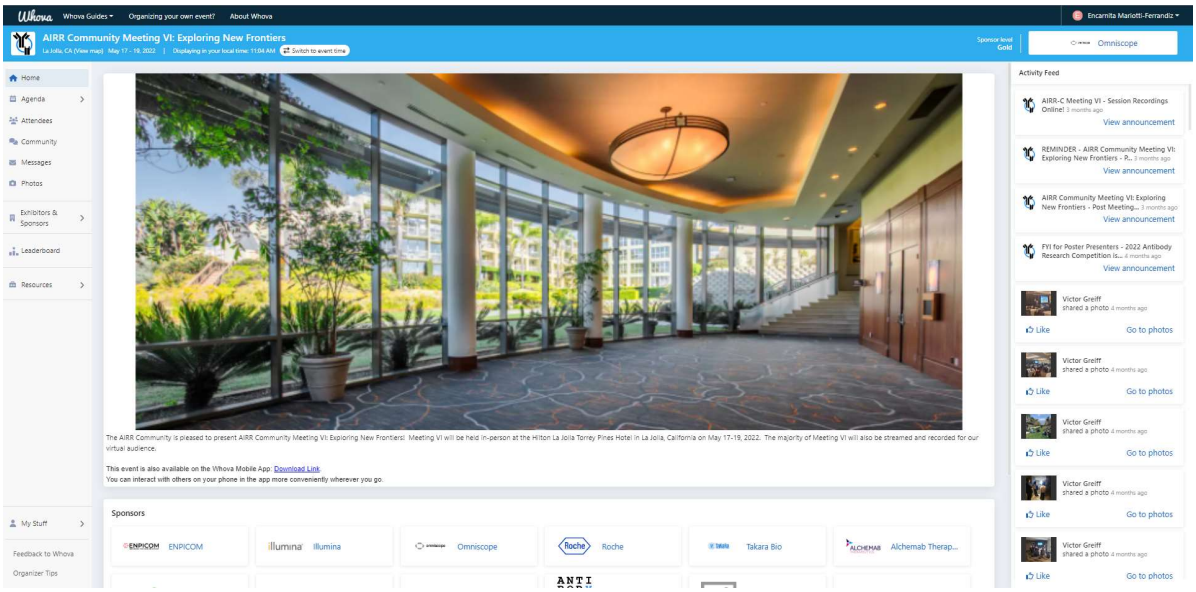
Thursday, May 19th 2022		
Start	**End**	**Activities**
8:00	8:30	*Registration open*
8:30	9:00	*Coffee*
9:00	9:05	**Introduction + Announcements** by AIRRC Executive Director, Pam Borghardt
9:05	12:30	**Biomedical Science Session** Christian Busse, DKFZ, (Germany) Moderator
9:05	9:55	**Keynote:** Shane Crotty, La Jolla Institute for Immunology (USA)
9:55	10:25	**Invited Talk1:** Ali Ellebedy, Washington Univ. School of Medicine, St. Louis (USA)

10:25	10:55	**Invited Talk2:** Anastasia Minervina, St. Jude Children's Research Hospital (USA)
10:55	11:15	*Break*
11:15	11:45	**Invited Talk3:** Jens Meiler, PhD, Vanderbilt University (USA)
11:45	12:30	**Contributed Talks** *Kenneth Hoehn, Yale School of Medicine (USA)* *Mikhail Pogorelyy, St. Jude Children's Research Hospital (USA)* *Paul Stys, Sorbonne Université (FR)*
12:30	14:00	*Science, Tools & Technology Lunch - Gold Sponsor Presentations* Enpicom, Illumina, Omniscope, Roche, Takara Bio
14:00	17:00	**Scientific Challenge Session: Systems Immunology**, Encarnita Mariotti-Ferrandiz, Sorbonne Université (France) Moderator
14:00	14:50	**Keynote:** Atul Butte, University of California San Francisco (USA)
14:50	15:20	**Invited Talk1:** Jonathan Herman, Massachusetts General Hospital (USA)
15:20	15:50	**Invited Talk2:** James Heath, Institute for Systems Biology, Seattle (USA)
15:50	16:20	**Invited Talk3:** Steve Kleinstein, Yale University (USA)
16:20	16:40	Break
16:40	17:10	**Panel Discussion with the Challenge Session Speakers** With prepared questions and questions from the audience
17:10	17:30	**Closing Remarks & Announcements**
17:30	18:30	**Poster Session II**, Victor Greiff, University of Oslo (NO) Moderator
19:00	21:00	*GALA DINNER & AWARDS*

Friday, May 20th 2022		
Start	**End**	**Activities**
8:00	12:00	Working Groups and Sub-committees Ad hoc Wrap Up and Planning Meetings

Attending Virtually

The AIRRC Meetings Sub-committee is pleased to provide a Virtual Attendance option for AIRR Community Meeting VI through the Whova Application.



AIRR Community Meeting VI was streamed live through the Whova platform. Virtual attendees were able to follow all of the scientific sessions seeing the slides and listening to the speakers. They also had the chance to ask questions via live chat during the Q&A portion of each session.

All recorded sessions will be kept available on the platform after the meeting.
To access the platform, use your credentials received upon registration.

Invited Speakers

Session	Name	Affiliation	Bio
Basic Science	Gunilla Karlsson Hedestam	Karolinska Institutet	Gunilla Karlsson Hedestam is a Professor at the Karolinska Institute in Stockholm, Sweden, where her group works on anti-viral immunity and immunogenetics. With the computational tool, IgDiscover, her group has revealed a great degree of diversity in the genes that encode adaptive immune receptors in both humans and macaques. Central questions in her research are how this germline gene variation has evolved and how it shapes immune responses to vaccines and infections. In several projects, her group combines monoclonal antibody isolation and deep antibody repertoire studies to define the evolution of antigen-specific B cell responses. She received her BSc from Uppsala University in 1990 and a PhD from the University of Oxford in 1993. Between 1994 and 1998, she was a post-doctoral fellow at the Dana-Farber Cancer Institute at Harvard Medical School in Boston. She became an Associate Professor at Karolinska Institutet in 2004 and a tenured Professor in 2012. She is the deputy Chair of the Committee for the Nobel Prize in Physiology or Medicine, and a member of the Royal Swedish Academy of Sciences.
Basic Science	Ranjan Sen	National Institute on Aging, NIH (USA)	Dr. Sen received his Ph.D., degree in chemistry from Columbia University in 1982. He made the transition to molecular biology as a postdoctoral fellow in David Baltimore's laboratory at M.I.T. and the Whitehead Institute. During this stage he developed his current interests in gene regulation. In 1987 Dr. Sen was appointed Assistant Professor in the Department of Biology and Rosenstiel Research Center at Brandies University. He earned tenure in 1991 and was promoted to Professor of Biology in 1998. He moved to his present position as Chief, Laboratory of Molecular Biology and Immunology, National Institute on Aging in 2003.

Basic Science	Susana Magadan	University of Vigo and University of New Mexico (USA)	I am an Associate professor (tenure track) at University of Vigo (Spain). During my PhD studies in Biology, I worked in a project to generate fully human monoclonal antibodies directed to human leukemia cells using transgenic mice. Then, I continued my career in a private university in Portugal (ISAVE), where I worked in applied immunology and started new researcher line about genomic characterization of the adaptive immune components in different low vertebrates (teleost and reptils). In the last ten years, I have been working in different international and national research centers. I worked with Dr. Pierre Boudinot at the Institute National de la Recherche Agronomique (INRA, Jouy en Josas, France) and with Dr. Salinas (University of New Mexico, USA) where I developed different deep sequencing protocols of Adaptive Immune Receptor Repertoire sequencing (AIRRseq) to characterize the complexity of adaptive immune response in teleost and to identify B/T cell molecular markers of specific immune response against pathogens and vaccines, and to get knowledge in the development of adaptive immune memory in teleost. My scientific career has been awarded by competitive pre-doctoral and post-doctoral grants, including a Co-fund Talent Attraction Marie Curie Postdoctoral Fellowship (FellowSea) in 2016, to develop a research project at the University of Vigo focused on the study of adaptive immune response in teleost at mucosal and systemic level. She also started a new line of research, in collaboration with groups of Spanish National Research Council (CSIC), to unveil and characterize the sensitization phase in fish food allergy. This is a multidisciplinary project where we integrate different approaches such as animal models, Proteomics and Systems Biology. The results of her work are reflected in 2 patents, over 40 peer-reviewed JCR publications, 8 book chapters in renowned national and international publishers (ORCID: 0000-0003-2968-0102). As a faculty member, I am involved in teaching activities within the degree on Chemistry and Biology at the University of Vigo and, in four Masters: in Nutrition, in Advanced Biotechnology, in Genomics and Genetics, and in Aquaculture, as well as training students. I am also recognized as Adjunct Assistant Professor at the faculty of Biology of University of New Mexico (UNM, USA).

Biomedical Science	Ali Ellebedy	Washington Univ. School of Medicine, St. Louis (USA)	Ali Ellebedy is a viral immunologist. He was born in Egypt. He graduated with a B.S. in pharmaceutical sciences from Cairo University in 2004. It was during his time at pharmacy school that he was first exposed to- and became fascinated with immunology. In 2006, he moved to the US where he studied influenza virus vaccines at St Jude Children's Research Hospital in Memphis, TN for his Ph.D. studies. He earned his Ph.D. in 2011 and moved to Emory University in Atlanta, GA, where he was a postdoctoral fellow in the laboratory of Rafi Ahmed. At Emory, he studied human B cell responses to influenza. In 2017, Ali joined the Department of Pathology and Immunology at Washington University School of Medicine in St. Louis as an Assistant Professor. His team studies the factors determining the ultimate fate of B cells once they are engaged via virus infection or vaccination. He is a tenured associate professor and co-director of the Center for Vaccines and Immunity to Microbial Pathogens at Washington University School of Medicine

Mentoring session

Mentoring event:

The AIRRC proposed for the first time a mentoring plan for undergraduate students. The aim of this mentoring plan were 1) to introduce the next generation of AIRR researchers to the challenges of the field and 2) to identify needs and challenges in their education path.

For the whole duration of the meeting, a mentor will accompany one student volunteer for mentoring in the understanding of the field from the biological and computational aspects towards application in basic and applied research. This will happen by spontaneous and self-organization discussions between the mentor and the student. At the end of the meeting, the whole group of students may draft a first version of a manuscript to highlight what they learned, and what are the challenges in their education curriculum to gain an interdisciplinary background supporting their future career as leaders in systems immunology and AIRR research. The manuscript will be submitted, upon revision by the mentor and the session moderator, for publication in a peer-review journal (e.g. Frontiers in systems biology, possible Research Topic: Education in Systems Biology 2022; Emerging Talents in Systems Biology: Integrative Systems Immunology 2022).

Mentoring session description:

The mentoring session was organized over lunch. This was an introductory session that will gather students and mentors. Students and mentors will meet at a spot to be announced on site, grab lunch and start the session. After a general introduction of the mentoring aims by the session moderator, students and mentors will introduce themselves (2 to 3 minutes), with a focus on their field of training, education level (for students) and research expertise. The pairs of "mentor/student" will be announced at this session.

Mentoring session organization:

This session lasted 1h.

Moderator (Encarnita Mariotti-Ferrandiz): Introduction of the mentoring goals (10 minutes)

Round table of student self-introduction (20 minutes): 1 "slide" self presentation, education, how did they learn about AIRR & interest in the AIRRC field (biology, computational aspects, bioinformatics….)

Round table of mentors self-introduction (20 minutes): self introduction (expertise, training background)

Forming pairs (5 minutes).

Conclusion (5 minutes)

Mentoring session participants:

16 senior researchers (including post-docs, associate professors and professors) offered their participation as mentors.

23 students (PhD candidate or graduate) applied as mentees.

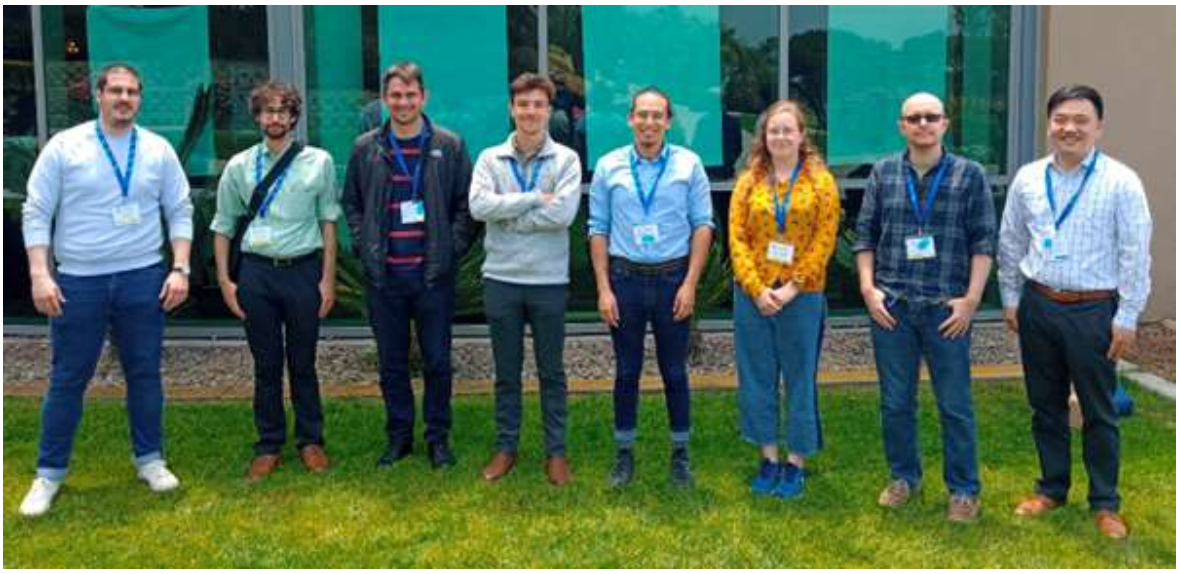
Industry Networking Reception

AIRR Community Meeting VI: Exploring New Frontiers Industry Networking Reception - May 18th, 2022 Hot Topics in the Field - Round Table Session			
Table #	**Topic**	**Co-Moderator - AIRR C**	**Co-Moderator - Industry**
1	Data Sharing & Data Standards	Brian Corrie - SFU/AIRRC	Holger Heyn - Omniscope
2	AIRR Molecular Biology Methods & Controls	Encarnita Mariotti-Ferrandiz - Sorbonne Universite/AIRRC	Andrew Farmer - Takara Bio
3	Diagnostics & Antibody Discovery	Victor Greiff - University of Oslo/AIRRC	Jake Galson - Alchemab
4	AIRR Analysis Software (from annotation to modeling)	Jason Vander Heiden - Genentech/AIRRC	Konrad Krawczyk - Natural Antibody
5	Privacy Preservation	Artur Rocha - INESC TEC/AIRRC	Henk Jan van den Ham - Enpicom

Travel Award Recipients

The AIRR Community was successful in their NIH Scientific Conference Grant proposal and is delighted to recognize the Meeting VI NIH travel award recipients.

- Alexander Brown – USA
- Lauren Overend – UK
- Edward Lee – USA
- James Heather – USA
- Sebastiaan Valkiers – Belgium
- Brennan Abanades Kenyon – UK
- Oscar Rodriguez – USA
- Paul Stys – France



Lightning Demonstrations

Title	Presenter	Abstract
Lightning Demo 01 - Stitchr	Jamie Heather	Stitchr accurately generates full-length, properly spliced, coding nucleotide sequences for T cell receptors, given only the minimal V gene, J gene, and CDR3 sequences that are typically recorded. It can be operated via a graphical user interface or the command for a few sequences (e.g. for generating expression constructs), or via the command line for high-throughput volumes of TCRs provided in a tab-separated file as part of an analysis pipeline. Modification of the reference genes also allows the introduction of novel sequences, e.g. for rational TCR engineering. Stitchr offers a fast, reliable, and repeatable way to generate TCR sequences, which should be of use to many across the fields of TCR research.
Lightning Demo 02 - ClusTCR	Sebastiaan Valkiers	The T-cell receptor (TCR) determines the specificity of a T-cell towards an epitope. Although the rules for epitope recognition remain largely elusive, it is well established that TCRs with similar sequences have a high probability of targeting the same epitope. Based on this observation, TCR sequences can be grouped into clusters of high sequence similarity, which are typically reflective of shared epitope-specificity. Similarity-based clustering thereby provides an overview of the potential antigens an individual's repertoire can possibly target. ClusTCR is a clustering application that can evaluate the sequence similarity of millions of TCR sequences, without the need for high performance computing infrastructures. ClusTCR implements the Faiss library to rapidly assign TCR sequences to large superclusters of approximate similarity. Subsequently, the algorithm applies a slower, more accurate method to divide the superclusters into epitope-specific subgroups. This tool provides the much needed functionality for interpreting the growing amount of highly complex adaptive immune receptor repertoire (AIRR) profiles. In addition to clustering, ClusTCR provides analysis tools to explore clustering results. These include the identification of consensus motifs, average cluster generation probability, or calculation of numerical representations for TCR clusters. Finally, the batch clustering functionality allows for the evaluation of shared cluster patterns across multiple unique repertoires. ClusTCR is available as a python package, and can be installed from the conda repository. Detailed instructions and examples on the use of ClusTCR can be found in the documentation:https://svalkiers.github.io/clusTCR

Lightning Demo 03 - TCRcloud	Eric de Sousa	TCRcloud is a tool to create visualizations of TCR repertoires. TCRcloud is implemented in Python and is free open-source software distributed under MIT license and can be installed via the Python Package Index (PyPI). TCRcloud is compatible with all four chains of the TCR, uses the Adaptive Immune Receptor Repertoire (AIRR) Data Commons (ADC) API and can be used for three tasks: to download rearrangement files directly from 6 ADC compliant repositories, to create word clouds representing the TCR repertoire and to create radar charts with several common diversity metrics.
Lightning Demo 04 - PipeBio		PipeBio's cloud based bioinformatics sequence analysis tools offer advanced graphical interactive filtering and a powerful computational engine designed for the large volumes of NGS and Sanger datasets common in today’s biologic workflows. The software is flexible to support annotation, clustering, differential enrichment, hit picking etc of traditional antibodies but also peptide sequences such as affibodies and more.
Lightning Demo 05 - CompAIRR	Lonneke Scheffer	Identifying identical or similar AIR sequences across individuals is a key step in AIRR analysis for revealing convergent immune response patterns that may be exploited for diagnostics and therapy. Existing methods for quantifying AIRR overlap do not scale with increasing dataset numbers and sizes. To address this limitation, we developed CompAIRR, which enables ultra-fast computation of AIRR overlap, based on either exact or approximate sequence matching. CompAIRR improves computational speed 1000-fold relative to the state of the art and uses only one-third of the memory: on the same machine, the exact pairwise AIRR overlap of 10^4 AIRRs with 10^5 sequences is found in ∼17 minutes, while the fastest alternative tool requires 10 days. CompAIRR has been integrated with the machine learning ecosystem immuneML to speed up various commonly used AIRR-based machine learning applications.
Lightning Demo 06 - AIRRscape	Eric Waltari	R Shiny tool to visualize antibody lists or repertoires. Repertoires are first displayed as interactive and explorable bins of germline V-gene, germline J-gene, and CDR3 length, providing a high-level view of the entire repertoire. Interesting subsets of repertoires can be quickly identified and selected, and then network topologies of CDR3 motifs can be generated for further exploration or to examine convergence among motifs.

Lightning Demo 07 - Federation of antibody data for immunoinformatics & biologics discovery	Konrad Krawczyk	At NaturalAntibody we are developing computational methods to facilitate development of antibody therapeutics. In order to get the best possible statistical view of antibodies we collate data on these molecules from multiple public sources, notwithstanding BCR sequences from AIRR- datasets. Via automated annotation we identified close to 250 BCR-containing repositories which are available via our federated system alongside patent, structure, GenBank and scientific publication data. We will demonstrate how one can use our platform to discover publicly available information on antibodies via sequences & text-based retrievals, focusing on the AIRR-datasets.
Lightning Demo 08 - SADIE	Jordan Willis	SADIE is the Sequencing Analysis and Data library for Immunoinformatics Exploration. SADIE is python library developed in accordance with the standards set by the AIRR Community. SADIE provides the following: Pre-built command line apps for popular immunoinformatics applications. A low-level API framework for immunoinformatics developers to build higher level tools. A testable, reusable, and statically typed library. A customizable and verified germline reference library streamed from public APIs. Portability ready to use out of the box. SADIE aims to meet the needs of all level of users. SADIE contains both low, mid and high level functionality for immunoinformatics tools and workflows. You can use SADIE as a framework to develop your own tools, use many of the prebuilt contributed tools, or run it in a notebook to enable data exploration. In addition, SADIE aims to port common immunoinformatics applications to python because it relies heavily on the Pandas library, the workhorse of the data science/machine learning age.

Deep Dive Tutorials

Title	Presenter	Abstract
Tutorial 1 - immuneML	Lonneke Scheffer	immuneML (https://immuneml.uio.no) is a software platform for machine learning analysis of adaptive immune receptor repertoires. The main use cases are repertoire classification (immune state prediction) and antigen-binding prediction. immuneML allows for the development and benchmarking of machine learning models for these two tasks. The platform is extensible and open-source and is available as a Python package, Docker image, and Galaxy web server.
Tutorial 2 - iReceptor Gateway	Brian Corrie	The iReceptor Gateway is a widely used platform to access the AIRR Data Commons. The original iReceptor paper was published in 2018, with iReceptor Version 3 released in mid June 2020. In this tutorial we will cover new iReceptor Gateway capabilities (including AIRR v2.0 standards) being released as iReceptor Version 4 in mid 2022. These new capabilities include storage and searching of clone data, storage and searching of single cell data (including gene expression data) as well as advanced analysis capabilities. The iReceptor Gateway is a widely used platform to access the AIRR Data Commons. The original iReceptor paper was published in 2018, with iReceptor Version 3 released in mid June 2020. In this tutorial we will cover new iReceptor Gateway capabilities (including AIRR v2.0 standards) being released as iReceptor Version 4 in mid 2022. These new capabilities include storage and searching of clone data, storage and searching of single cell data (including gene expression data) as well as advanced analysis capabilities. In addition, we will take some time to explore the iReceptor Turnkey, an AIRR Standards compliant data repository that makes it easy to download, install and run your own AIRR Data Commons repository containing your own data.

Tutorial 3 - VDJbase	William Lees	Understanding of immune receptor genes has increased significantly in recent years. In part, this has come as a consequence of novel sequencing techniques, which now provide substantially improved genomic reference assemblies and the capability to conduct high-throughput long-read sequencing. Improved analysis has also proved important, and, in particular, has revealed important insights into allelic variation and improvements in genotyping and haplotyping. Workshop attendees will gain hands-on experience of tools available to explore and utilise this data, focussing on advanced and experimental facilities in VDJbase. They will also learn about the application of the underlying tools to AIRR-seq studies in general, in order to improve data quality and to gain greater understanding of the impact of genetic variation. The workshop will be particularly interesting to researchers interested in understanding the impact of receptor gene variation on the immune response, or in conducting AIRR-seq analyses, such as B-cell clonal analyses, for which an accurate allele call is essential. It will also be of interest to those wishing to explore allelic variation and allele discovery in humans and other species, and those interested in exploring the impact of features outside the core V,D,J coding regions, for example variation in the constant region genes, or in UTR, leader and RSS.

Tutorial 4 - Immcantation & Dowser	Kenneth Hoehn	The Immcantation suite (http://immcantation.org) is a well-established collection of open-source Python and R packages that provide methods to characterize the adaptive immune receptors that mediate immune responses, with a focus on B cell receptors. Immcantation is also available as a versioned Docker container that includes the entire set of tools, third-party dependencies, a collection of example workflows, and reference data. All Immcantation tools are compatible with AIRR Community standards, including reading/writing AIRR-format files. This demo will provide a high-level overview of the full framework, with special attention given to the most recent methods for single-cell sequencing data. This demo will provide a high-level overview of the full framework, with special attention given to the most recent methods for single-cell sequencing data .B cell lineage trees are a critical component of many AIRR-seq analyses. B cell somatic hypermutation produces clonal lineages of B cells, which can be studied using lineage trees inferred from B cell receptor (BCR) sequences. These trees are models of ancestor/descendant relationships among cells and can be used to reconstruct the series of mutations leading from the clonal germline to observed BCR sequences. Further, lineage trees that contain B cells from multiple tissue biopsies, cell types, and/or timepoints have the potential to be used to make inferences about B cell migration, differentiation, and evolution over time. Dowser is a newly developed R package for easily building, visualizing, and analyzing B cell lineage trees from bulk and single-cell BCR sequencing data. Dowser enables users to build lineage trees and reconstruct intermediate sequences using multiple methods including maximum parsimony, maximum likelihood, and IgPhyML. For visualizing trees, Dowser provides a flexible system for generating publication-quality tree figures. Beyond simply building and visualizing trees, Dowser implements newly developed statistical tests for using B cell lineage trees to understand different aspects of immune responses. If analyzing data from multiple tissue samples, cell types, or isotypes, Dowser can be used to analyze patterns B cell migration, differentiation, and class switching, respectively. If using data from multiple timepoints from the same subject, Dowser can be used to detect ongoing evolution of B cell lineages over the sample interval. For interoperability, Dowser is designed to be used with AIRR Rearrangement-formatted data. Finally, Dowser is supported as part of the Immcantation suite, which provides a start-to-finish ecosystem of AIRR-seq analysis programs.

Tutorial 5 - OPIG Antibody Suite	Eve Richardson ; Sarah Robinson; Brennan Abanades Kenyon	In this tutorial, we will explore computational techniques for B-cell receptor/antibody analysis through a mock in silico drug discovery case study, starting from convalescent patient repertoires that represent a trove of potential human-expressible therapeutic antibodies. We will identify the convergent (seen across multiple individuals) disease-responding antibodies, analyse them for likelihood of pathogen complementarity, and finally assess them for developability concerns such as stability and humanness. In this tutorial, we will explore computational techniques for B-cell receptor/antibody analysis through a mock in silico drug discovery case study, starting from convalescent patient repertoires that represent a trove of potential human-expressible therapeutic antibodies. We will identify the convergent (seen across multiple individuals) disease-responding antibodies, analyse them for likelihood of pathogen complementarity, and finally assess them for developability concerns such as stability and humanness.
Tutorial 6 - TITAN	Anna Weber	Reliable prediction of T cell receptor (TCR) specificity and understanding the mechanisms underlying the TCR-pMHC interaction is both a daunting and highly relevant challenge. To train our model TITAN Tcr epITope bimodal Attention Networks and this task, we leveraged machine learning techniques from transfer learning to interpretability to achieve a state-of-the-art prediction performance while also giving insights into the decision process of the model. In the tutorial you will get a solid overview of the methods we employed and most importantly: learn how to apply the model hands-on with the code provided on our GitHub. After the workshop, you will have a fully trained TITAN running on your computer, ready to make predictions or to be retrained on any TCR specificity dataset you might have.You want to gain even more biological insight about the complex interaction of TCR and pMHC? Then you will also be interested to hear about our new method DECODE (conditionally accepted at ISMB2022), an interpretable pipeline that can applied to any sequence-based TCR specificity prediction model. Using TITAN as an example, you will learn how to apply the DECODE pipeline to explore the rules based on which a model makes its predictions.

Posters

Number	Title	Name	Description
List of Posters – Session 1			
101	MHC alleles differentially affect TCR repertoire diversity and alloreactivity	Alex Brown	In humans and mice, certain MHC alleles confer susceptibility to autoimmune diseases, while other alleles are protective; even when co-expressed with the disease-promoting allele. We suspect that protection is driven by changes in TCR repertoire, yet little is known about how the inheritance of one MHC allele versus two MHC alleles affects TCR repertoire. Since current technologies are unable to capture and sequence TCR α/β chain-pairs at repertoire levels, we generated mice with a limited TCR repertoire which express a single transgenic TCRβ chain but were heterozygous for TCRα and MHC-II homozygous (MHC/hom: I-Ab, I-Af, I-Ag7 or I-As) or MHC-II heterozygous (MHC/het: I-Abxf, I-Abxg7, I-Abxs or I-Afxs). One might predict that MHC/het animals would contain T cells bearing almost all the TCR α/β pairs that are present in either MHC/hom parent. Surprisingly, MHC/het mice lack up to half of the TCR sequences that can be found in a given parent. To test that TCRs missing in MHC/het animals can appear in the presence of both parental MHC/hom class II molecules, we developed tetraparental mice (derived from chimerism two MHC/hom counterpart blastocysts). Tetraparentals do not express mixed molecule MHC-II, but express both "parental" MHC-II on separate cells in the mouse and, in contrast to MHC/het animals, do not develop 'gaps' in their repertoire. We cloned 900 candidate TCR sequences and developed a library screening approach to test if these TCR sequences present in MHC/hom parents are absent in the MHC/het mice due to alloreactivity with the MHC-II of the other parent and/or mixed MHCII. These experiments suggest: 1.) There is an inverse relationship between number of MHC alleles expressed and the size of the TCR repertoire. 2.) Some TCRα chains missing in MHC/hets are negatively selected due to self-reactivity on mixed MHC-II molecules or MHC-II from the other parent. 3.) Protective MHC alleles may thymically delete the autoreactive clonotypes positively selected on the risk promoting MHC allele. 4.) Positive selection and self-reactivity are primarily determined by TCR CDR3 sequence.

102	MHC and sex bias: Shaping of the systemic T cell repertoire in three-spined stickleback fish	Ana Teles	The T cell receptor (TCR) diversity necessary for the recognition of a broad antigen spectrum is determined by the interaction between TCRs and cognate ligands presented by major histocompatibility complex (MHC) molecules during T cell selection in the thymus. Evolution might have favored an optimal diversity in the copy number-variable MHC, defined by a trade-off between the benefits of presenting a larger number of foreign antigens and the disadvantage of a limited mature T cell repertoire following negative thymic selection, due to presentation of a larger number of self-antigens. However, our understanding of how the initial T cell repertoire is shaped is still very limited. Three-spined sticklebacks have a completely functional adaptive immune system and exhibit a natural level of diversity at the MHC. The small size of this wild fish allows an easier estimate of the systemic TCR diversity for each individual, ideal in eco-immunological studies. We have developed a cDNA-based 5'RACE protocol with unique molecular identifiers and a stickleback TCRß gene reference library. By characterizing the systemic TCRß repertoire diversity among male and female naive lab bred individuals belonging to different families and harboring copy number-variable MHC genotypes ranging from low to high diversity, we have directly tested the association between gender, MHC, and naive TCR diversity. Our results contribute to the understanding of the T cell selection process and the balance between the recognition of pathogenic vs self antigens during the evolution of the vertebrate adaptive immune system.
103	Analysis of inter-individual T cell repertoire diversity in the three-spined stickleback during infection as a natural model system for adaptive immunity	Artemis Efstratiou	The vertebrate adaptive immune system is based on lymphocyte recognition of pathogen-derived antigens presented by Major Histocompatibility Complex (MHC) molecules. As the engagement of a peptide-MHC complex with a suitable T cell receptor (TCR) is the critical first step in the initiation of adaptive immune responses, the availability of a diverse T cell repertoire, constituted by a pool of broad TCR specificities, is crucial. However, surprisingly little is known regarding the degree of natural variation in the inter-individual diversity and dynamics of the T cell repertoire, especially during infection. Our research thus examines the qualitative and quantitative variation and dynamics of TCRβ repertoires within and among individuals of the three-spined stickleback, an eco-evolutionary model species. Here we analyze T cell repertoires of lab-bred sticklebacks that were experimentally exposed to the cestode parasite Schistocephalus solidus, known to trigger adaptive immunity. Using NGS sequencing and advanced bioinformatics tools, we investigate TCRß repertoire size and diversity in relation to infection treatment, family background and individual MHC diversity. Preliminary analyses indicate substantial variation of TCRβ repertoires among individuals and across infection status, as expected for a species with a natural level of genetic diversity. Interestingly, infected individuals appear to exhibit higher inter-repertoire overlap than control ones. The existence of public, expanded clonotypes shared by all infected individuals further

			hints at convergent antigen-specific T cell responses. Yet - surprisingly- most of the top public clonotypes are shared across experimental groups. Lastly, we show significant biases and a family effect on the usage of V-J gene segments.
104	QtCR : an integrative pipeline for adaptive immune receptor repertoire quality control	Kenz Le Gouge	Next-generation sequencing technologies have revolutionized the Adaptive Immune Receptor Repertoire (AIRR) sequencing analysis. Increasing sample throughput with dedicated kits and increased depth profoundly complexified downstream analysis, often at the cost of a thorough quality control (QC). Indeed, assessing data quality is a crucial step toward subsequent analysis, yet often overlooked. Here, we propose QtCR, an integrative QC pipeline particularly adapted for T-cell receptor sequencing that combines QC on both raw and aligned reads. QtCR also includes correlations between read data and sample metadata with an at-a-glance readout. It provides all the tools to spot inconsistencies between input material and sequenced data, identify outliers and quickly monitor run quality. QtCR pipeline is wrapped in an R shiny interface to efficiently skim through all the metrics with a user-friendly interface. QtCR would benefit to all researchers, neophytes or experienced. Tool will be open source and freely available.
105	Genomic and proteomic antibody repertoire analysis at single-cell and single-molecule resolution	Khang Le Quy	The diversity of the antibody repertoire is crucial to broad pathogen recognition. Therefore, investigating the relationship between the genomic (BCR) and phenotypic (serum antibody) diversity of antibodies is crucial for understanding human adaptive immunity. The capability of accurately predicting and describing the entire antibody repertoire has the potential to dramatically alter vaccine development and disease diagnostics. However, despite advances in high-throughput BCR sequencing, a joint characterization of antibody complexity at the genomic single-cell and proteomic single-molecule level remains elusive due to the fact that mass spectrometry (MS) analysis of the serum antibody repertoires has remained underdeveloped. Specifically, it remains unclear to what extent the blood B-cell receptor repertoire differs from the serum antibody repertoire or, even more fundamentally, what the number of unique antibody clonotypes in the blood is. To address these questions, we isolated blood-borne B cells from a healthy donor and sequenced the BCR repertoire at bulk and single-cell level. Simultaneously, serum antibodies of all major isotypes were isolated, digested, and sequenced with tandem MS, ensuring that the antibody repertoire is covered comprehensively. Systems immunology analysis showed high concordance between bulk- and single-cell sequencing. MS was able to identify CDRH/L3 peptides linked to specific clonotypes in the sequencing libraries, demonstrating how to reconstruct antibody clonotypes from proteomics data at single-molecule resolution. In addition, performance benchmarking of antibody MS revealed that relative CDR3-peptide concentration differences can be measured with high sensitivity. In conclusion, we developed a platform that connects bulk sequencing, single-cell sequencing, and mass spectrometry enabling the quantification of the serum antibody

			repertoire at unprecedented precision thereby removing a blank spot in our understanding of antibody adaptive immunity.
106	Antibody repertoire heavy chain gene usage is explained by common genetic variants in the immunoglobulin heavy chain locus	Oscar Rodriguez	Antibodies (Abs) are critical components of the adaptive immune system. The heavy chains of Abs are encoded by genes at the immunoglobulin heavy chain (IGH) locus. The IGH locus harbors >100 variable (V), diversity (D), and joining (J) gene segments, and is among the most structurally complex regions of the human genome, containing extensive haplotype diversity, with an increased amount of single nucleotide variants (SNVs) and large gene-containing structural variants (SVs). Mounting evidence shows that genetic variants within IGH influences inter-individual variation in the expressed Ab repertoire, with potential downstream consequences on Ab function in human health. However, genetic variants contributing to Ab repertoire differences between healthy adults have not been comprehensively characterized. To address this, we conducted the first genetic association study to identify IGH variants that contribute to expressed Ab repertoire signatures. Utilizing our novel targeted genomic long read sequencing approach, we generated paired IGH haplotype and expressed Ab repertoire sequencing data in 154 healthy adults, the largest dataset of its kind. Our analysis revealed IGH genetic variants strongly influence the composition of the expressed repertoire, principally through impacts on V and D gene usage. Associations revealed effects of both coding and non-coding SNPs, and large SVs. Variants were associated with gene usage for single genes or multiple genes demonstrating signatures of coordinated gene regulation, and were enriched in regulatory regions annotated by the ENCODE project, including co-localization within CTCF binding sites, a transcription factor involved in V(D)J recombination. These data provide the first robust links between IGH polymorphism and the Ab repertoire. Extending this dataset will allow for the discovery of novel genomic factors and molecular mechanisms influencing Ab repertoire development and diversity, and facilitate better characterization of the Ab response in disease and clinical phenotypes.

107	Population genotyping of the germline immunoglobulin repertoire in AIDS-designated rhesus macaque breeding colonies.	Steven Bosinger	"An urgent priority of contemporary HIV vaccine research is the development of strategies to elicit broadly neutralizing antibodies (bNAbs) capable of providing protection against diverse strains. Currently, some of the most promising vaccine strategies are those in which specific classes of germline immunoglobulin (IG) genes are targeted by immunogens and Ab evolution is directed by sequential immunization with a changing set of epitopes. For these pre-clinical studies to progress, a key technical roadblock is a lack of information on the composition, population diversity and extent of human-orthology in the IG loci of Indian-origin rhesus macaques (RMs). The RM IG loci are poorly characterized at the genomic level, in part due to technical challenges assessing these structurally complex loci. To date, detailed resolution of the IG at the genomic level has only been reported for two Indian RMs. As a result, there is very limited knowledge of RM IG diversity, as well as how it may impact the availability of germline sequences within the naive repertoire. Without this information readily available, specific IG genotypes cannot be used as criteria when enrolling RMs into HIV vaccine studies. This contrasts with MHC genotyping, which has become standard practice for vaccine studies and is a core function in the genetic management of RM colonies. The paucity of data on RM IG therefore represents a significant knowledge gap, and is a critical barrier to effectively using RMs in contemporary HIV vaccine research. Our team is well positioned to overcome this barrier, as we have developed novel approaches for re-constructing full-locus IG haplotypes utilizing long-read sequencing and pairing this data with expressed repertoire sequencing (RepSeq) analyses. In work forming the foundation of this proposal, we completed the first long-read assembly of the Indian rhesus macaque genome, from which we compiled a unique dataset of RM germline IG alleles. This dataset was an essential component to differentiating clonotype usage between non-neutralizing immunodominant responses and those leading to autologous neutralization in RMs vaccinated with Env immunogens. The primary objective of this proposal is to build upon these results by generating full-locus reference assemblies and germline variant catalogues for the IG loci across multiple RM centers in the USA. We will accomplish these objectives by pursuing the following specific aims:"

109	Systematic evaluation of B-cell clonal family inference approaches and impact on biological conclusions	Daria Balashova	Adaptive immune receptor repertoire sequencing (AIRR-seq) allows the sequencing of a wide variety of B-cell receptors (BCRs) in a sample. Clustering data into clonal families (CFs) is an important step because many subsequent types of AIRR analysis, such as determining repertoire diversity, identifying dominant and shared CFs, rely on it. However, CF inference remains challenging, and therefore a number of approaches have been developed, which include consideration of different regions of BcRs at the nucleotide or amino acid level and a set of mathematical methods – algorithmic, probabilistic and machine learning. The objective of this project is to systematically evaluate approaches of identifying B-cell CFs from AIRR-seq data and to determine the extent of influence on downstream biological analysis. Datasets with different characteristics were selected, and a series of simulations with parameters taken from experimental data were performed to obtain information on CF inference accuracy by different CF inference methods benchmarked. CF inference results obtained using different approaches was established to vary widely depending on the biological and technical features of the AIRR-seq data. In addition, we found that the choice of CF inference approach can significantly affect the identification of CFs and thus influence the results of subsequent analyses and biological conclusions.
110	AIRRscape: an interactive tool for exploring B-cell receptor repertoires and antibody responses	Eric Waltari	Exploratory analyses of Adaptive Immune Receptor Repertoire sequencing (AIRR-seq) datasets are challenging due to their size and complexity. To aid in data exploration, we have developed AIRRscape, an R Shiny-based interactive web browser application that enables B-cell receptor (BCR) and antibody feature discovery through comparisons among multiple repertoires. Using AIRR-seq data as input, AIRRscape starts by aggregating and sorting repertoires into interactive and explorable bins of germline V-gene, germline J-gene, and CDR3 length, providing a high-level view of the entire repertoire. Interesting subsets of repertoires can be quickly identified and selected, and then network topologies of CDR3 motifs can be generated for further exploration. We use AIRRscape to investigate patterns of immunity to SARS-CoV-2, HIV-1, and DENV among patient BCR repertoires and published antibody sets. AIRRscape reveals convergent antibody sequences among datasets for all three pathogens, although HIV-1 antibody datasets display limited convergence and idiosyncratic responses. We have made AIRRscape available on GitHub to encourage its open development and use by immuno-informaticians, virologists, immunologists, vaccine developers, and other scientists that are interested in exploring and comparing multiple immune receptor repertoires.

111	Development and validation of a novel framework for full length length Adaptive Immune Receptor Repertoire Sequencing (FLAIRR-seq)	Easton Ford	Currently available Adaptive Immune Receptor Repertoire Sequencing (AIRR-seq) methods aim to resolve immunoglobulin (Ig) variable region sequences with very limited inclusion of constant region genes. No method resolves the full-length Ig heavy chain (IGH). We developed a novel full-length AIRR-seq (FLAIRR-Seq) method that utilizes targeted amplification combined with single molecule real time (SMRT) sequencing on the Pacific Biosciences Sequel IIe to generate highly accurate, full length IGH transcripts in peripheral blood mononuclear cells (PBMC), purified B cells, and tissue. To validate the FLAIRR-seq method, Ig transcripts were targeted in RNA purified from healthy donor PBMC, which was converted to cDNA and amplified to full length Ig transcripts with 5' RACE PCR. Sequencing resulted in full-length, accurate (>99.99%) single-molecule resolution of IGH transcripts for both IgG and IgM compartments. These data provided direct association of variable region clonotypes, isotype and subisotype signatures, and Fc polymorphisms. Results were compared to parallel processing of the same samples using standard 5'RACE-based AIRR-seq methods. Analyses were performed using a modified version of the publicly available Immcantation tool suite. FLAIRR-seq and AIRRR-seq datasets showed significant correlations across multiple repertoire metrics. FLAIRR-seq data were able to further define constant region gene usage and haplotypes and examine subisotype-specific repertoires. We also identified several novel constant region alleles, which were confirmed with targeted genomic DNA capture of the IGH locus. These results demonstrate the feasibility and utility of the FLAIRR-seq method, which provides comprehensive characterization of expressed IgG and IgM repertoires.

112	Efficient and sensitive high-throughput human B-cell receptor repertoire profiling using SMART® technology	Bryan Bell	"OBJECTIVE: B-cell receptor (BCR) repertoire profiling is increasingly used in health and pathogenic contexts with the goal of biomarker discovery. However, current sequencing technologies are limited in their ability to generate data accurately and reproducibly for all BCR isotypes. To overcome these limitations, we have developed a new kit to accurately profile all heavy (A, D, E, G, M) and light-chain (K, L) isotypes—an end-to-end solution, from library preparation to streamlined data analysis. Here we present data on an updated approach for efficient and high-throughput BCR repertoire profiling of human samples METHODS: Libraries were prepared from human peripheral blood cells (10 ng–1 μg total RNA) or from B-cells (1 ng–100 ng total RNA) using our new human BCR repertoire profiling kit (~2.5 hours hands-on time). Prepared libraries were then analyzed on the Illumina® Miseq® benchtop sequencer using 300-bp paired-end reads. RESULTS: For each library, >90% of sequencing reads were on[1]target while the most highly represented clonotype was found to remain consistent among technical duplicates across a range of input amounts. In comparison to the previous version of our BCR[1]sequencing kit, the new approach enabled a ~4x increase in total clonotype count observed across various RNA inputs. Furthermore, a sensitivity assay demonstrated that B-cell RNA corresponding to a single clonotype could be detected above background levels when spiked into input total RNA at a relative concentration of 0.001% CONCLUSION: Our new human BCR repertoire profiling kit was found to accurately and reproducibly profile B-cell clones and provide information on the diversity of BCR repertoire in human samples"
113	Evaluating sequence embedding approaches in predicting biological properties of B cell receptors	Mamie Wang	High throughput sequencing of B cell receptors has been increasingly applied to study humoral immune responses, generating insights into the immense diversity of antibodies. Learning biologically meaningful representation of the sequences is an important step in predictive modeling. Word or sentence embedding methods, originally introduced for natural language processing, have been recently applied to adaptive immune receptor sequences and achieve state-of-the-art performance in binding prediction tasks. However, with a lack of benchmarking studies, it is unclear if embeddings that perform well in one task will succeed in a different task. Here we estimate the performances of multiple embedding methods to predict BCR sequence properties, including V, J gene usage, mutation frequency, and CDR3 features (e.g., length). We found that despite the differences in model architectures, most embeddings are good at capturing gene usage and mutation frequency information of BCRs. However, the embeddings different greatly in their ability to predict specificity of BCRs for SARS-CoV2 spike protein. The benchmarking framework we have developed can be applied to gain a better understanding of the properties of BCR embeddings and provide insights into improving downstream prediction applications for antibody discovery.

114	Upstream Region Sequences Support the Analysis of the IGHV Locus	Linnea Thörnqvist	Upstream regions – such as 5'-untranslated regions (UTRs) and leader sequences – of immunoglobulin genes are rarely studied even though they may affect the expression of these genes and can be utilized in the analysis of immunoglobulin gene loci. We are now exploiting a novel pipeline for analysis of 5'UTR-leader sequences in rearranged immunoglobulin heavy chain transcriptome data and identify 166 such upstream region sequences. Analysis of the identified sequences exemplifies how these regions can be used to support antibody genetic studies and aid in the deciphering of the complex IGHV gene locus. For example, the herein studied 5'UTR-leader sequence data indicate a potential duplication of the IGHV4-30-2 gene in one of the examined individuals, provide evidence of evolutionary relationship between certain IGHV genes, sharing insertions/deletions in their upstream regions, and imply that IGHV4-59*12 might in fact reside in the IGHV4-4 location.
115	Genomic characterization of the immunoglobulin light chain lambda locus from individuals of European, Asian and African origin.	William Gibson	The adaptive immune system relies on a diverse set of over one hundred immunoglobulin (IG) genes across three genomic loci that are variably combined to form antibodies (Ab). Studies show that the IG loci are highly variable between individuals and populations. However, the complexity of the IG loci severely limits the effective use of standard short read sequencing, limiting our knowledge of population diversity in these loci. We leveraged existing long read whole-genome sequencing (WGS) data, fosmid technology, and targeted long-fragment DNA capture (IG-cap) combined with single molecule, real time (SMRT) long read sequencing (Pacific Biosciences) to create haplotype-resolved assemblies of the IG Lambda (IGL) locus from 6 ethnically diverse individuals. In addition, we generated 10 diploid assemblies of IGL utilizing IG-cap combined with SMRT sequencing from a diverse cohort of individuals. These data represent highly accurate base-level assemblies of the IGL region. From these 16 individuals, we identified significant allelic diversity, including 37 novel IGLV alleles. In addition, we observed highly elevated single nucleotide variation (SNV) in IGLV genes relative to average IGL intergenic and genomic background SNV density. By comparing SNV calls between our high quality assemblies and existing short read datasets from the same individuals, we show a high propensity for false-positives in the short read datasets. Finally, for the first time we resolved, at nucleotide resolution, common 5-10 Kb duplications in the IGL constant region that contain functional IGLJ and IGLC genes. Together these data represent a significant advancement in our understanding of genetic variation and population diversity in the IGL locus.

116	A structural comparison of human and humanised mouse BCR repertoires	Eve Richardson	Humanised mice continue to be used as the workhorses of antibody discovery and increasingly in vaccine development, but the representativeness of their BCR repertoires of human BCR repertoires has not been fully described. At the sequence level, biases in germline gene usages and shorter CDRH3s indicate potentially important differences between the two but the complex mapping between these sequence properties and the structural properties of the repertoires has not been examined. In this work, we used high-throughput structural homology annotation and state-of-the-art eGNN-based modelling to describe and compare the naïve structural repertoires of humans, mice and the Intelliselect transgenic mouse. We report that recombination biases and reduced junctional diversification leads to shorter and less diverse CDRH3s in humanised mice than humans, but that the structural shapes that these CDRH3s adopt are more commonly observed in human repertoires than murine repertoires. Further, the CDRH3 structural profiles of humanised mice are within the sampled range of naïve human CDRH3 structural profiles. This indicates that the gulf between the naïve human and humanised mouse repertoires is smaller than is indicated by sequence analyses and demonstrates the utility of high-throughput structural modelling of BCR repertoires.
117	Decoding amyloidosis in human antibody light chains	Puneet Rawat	Protein aggregation is one of the major hindrance in the development of antibody-based therapeutics and causes several antibody-related diseases, such as systemic light chain amyloidosis, autoimmune diseases and plasma cell disorders (PCD). The prediction of aggregation capability of protein using sequence/structure information is a challenging task. However, the common architecture of antibodies provides crucial details such as surface exposed binding regions, stabilizing regions in antibodies, which may help better understand the aggregation of antibodies. We have analyzed the sequence/structure features of experimentally known 348 amyloidogenic and 1480 non-amyloidogenic light chains of antibodies. The analysis using state of the art aggregation-prone region (APR) prediction method "ANuPP" revealed that amyloidogenic light chains had marginally higher predicted aggregation propensity and higher percentage of exposed APRs than non-amyloidogenic light chains. Moreover, the percentage of gatekeeper residues in the vicinity of APRs is higher in the non-amyloidogenic light chains of antibodies. The presence of gatekeeper residues (D, E, R, K and P) around the APR regions can greatly hinder the aggregation process. The observations were further translated to the antibody sequence level to develop a high-throughput machine learning method capable for classifying amyloidogenic and non-amyloidogenic light chains. The method has shown prediction accuracy of 80% on the leave-one-out cross-validation and rigorous testing on blind test dataset (71%), clinical stage antibodies (75.6%) and human antibody repertoire (94.1%) has shown robustness of the method. The method will be a useful resource to improve the developability of antibodies and analyzing immune repertoires for potential aggregation-prone light chains.

			The model is freely available as a web server at https://web.iitm.ac.in/bioinfo2/vlamy-pred/index.html.
118	Prime-boost with mRNA-1273 Vaccination Re-elicits a Broad Range of SARS-CoV-2 Infection-Induced Antibody Lineages	Mark Chernyshev	Vaccination of SARS-CoV-2 convalescent individuals induces hybrid vigor immunity, generating a broader and 25 to 100 times more potent antibody response. While it is known that infection-induced B cell responses are highly polyclonal, it is not known how many of these clonal lineages are boosted by vaccination. Our study identifies vaccine re-elicited antibody lineages and their binding characteristics, allowing us to determine the breadth of the boosting effect. We performed a longitudinal study of two individuals who were infected by SARS-CoV-2 and later boosted with mRNA vaccination. We isolated a total of 479 spike-specific monoclonal antibodies (mAbs) across three timepoints and determined their half-maximal inhibitory concentration (IC50) neutralization values against the WT, Beta, Delta, and Omicron variants of SARS-CoV-2. IgG repertoires from all three time points was bulk sequenced to trace the lineages of the mAbs in IgG repertoire data. A total of 205 mAb lineages were traced in our bulk IgG data, including 20 neutralizing Ab lineages. The most frequent Ab lineages traced post-vaccination were specific to the spike S2 subunit, while the vast majority of the neutralizing Ab lineages were specific to the receptor-binding domain (RBD). Examination of somatic hypermutation (SHM) levels indicated ongoing affinity maturation in the time period between infection and vaccination with further increases in SHM in some lineages observed following vaccination. The study demonstrates that mRNA vaccination boosts a broad range of infection-induced antibody lineages, providing a basis for previous observations regarding the potency of hybrid immunity.
119	Longitudinal dynamics of B-cell repertoires in Rhesus Macaques immunized with the transmitted /founder envelop sequence of HIV-1	Rohini Mopuri	Despite overwhelming success with vaccines against myriad infectious diseases, there remains the need for a HIV-1 vaccine. A major part of vaccine development includes understanding the underlying mechanisms by which B-cells respond, diversify and mature. Twenty rhesus macaques (RM) were immunized sequentially with DNA, modified vaccinia Ankara (MVA) and protein derived from one of two transmitted/founder envelopes isolated from HIV-1 infected patients. Here, we use immunological techniques and high throughput B cell receptor and single cell RNA sequencing to characterize expansion, overlap and diversity of B cell responses over the course of vaccination. All animals developed robust antigen specific serum IgG. Additionally, two of the RMs developed autologous neutralizing activity and monoclonal antibodies representing an expanded neutralizing lineage were isolated from one animal. This lineage arose post-MVA and persisted for at least 64 weeks, showing evidence of increasing somatic hypermutation over time. The findings provide new insight into how B cells respond during immunization and why neutralizing antibodies arose in only a small number of animals.

120	VDJbase Updates	William Lees	VDJbase is an open-source database of immune receptor genes and population usage. Here we present recent developments: in particular, the inclusion of large-scale genomic data, adoption of MiAIRR metadata, and a 'turnkey' version you can use in your own lab.
List of Posters – Session 2			
201	Prediction of immunodominant CD4+ SARS-CoV-2 epitopes with TCR repertoire sequencing data	Mikhail Pogorelyy	The worldwide scientific effort to overcome COVID-19 pandemic led to the generation of an extraordinarily large amount of publicly available data describing T cell immune responses to SARS-CoV-2. The T cell receptor (TCR) is a hypervariable molecule defining the specificity of T cells for peptide-MHC complex. Defining exact epitopes is crucial to profile T cell responses and yet very few MHC-II restricted SARS-CoV-2 epitopes are currently known. Here, we propose a reverse epitope discovery technique, which, instead of using large pools of peptides to identify reactive T cells, utilizes TCR repertoire sequencing data as the means to predict immunodominant epitopes. The core idea of the approach is to combine information from large, publicly available TCR repertoire datasets: TCRbeta repertoires from cohorts of COVID-19 patients and healthy controls, and TCRbeta and paired TCR repertoires of single T cells activated by SARS-CoV-2 peptides. Our pipeline allows us to predict the HLA-restriction of public SARS-CoV-2 specific TCR clonotypes, which in turn allows us to predict binding to a specific peptide (instead of the pool, as in existing datasets). We applied this approach to single cell TCR repertoires of CD4+ T cells from COVID-19 patients, and predicted six MHC-II restricted immunodominant epitopes and alpha-beta TCR motifs recognising them and tested our predictions experimentally. We further applied this technique to bulk TCRalpha repertoires from T follicular helper cells from draining lymph nodes of healthy donors after BNT162b2 mRNA vaccination. We found that the DPB1*04 restricted S167-180 epitope is responsible for the largest public CD4+ response, and recognition of this epitope is driven by a semi-invariant TCR alpha chain. This finding led to generation of the DPB1*04 S167-180 tetramer, allowing to track SARS-CoV-2 specific CD4 cells in peripheral blood and lymph node samples with flow cytometry.

202	TCR repertoire as biomarker of immune diseases: applications to autoimmune diseases and COVID-19	Paul Stys	The TCR repertoire constitutes a complex biological component of the immune system that reflects the immune history of each individual. This property is of particular interest for most of the immune pathologies, and in particular for those for which T-cells play a major role. We are particularly interested in autoimmune diseases (AIDs), for which diagnostic and/or prognostic tools remain limited, as well as in COVID-19, as an example of an emerging infectious disease characterized by a high mortality that has recently affected the world population. Moreover, while most of the TCR repertoire studies analyse total blood repertoire, we focused on T-cell subsets, with various functional properties. In particular, we are interested in deciphering the role and repertoire diversity of effector CD4 T-cells (Teff), as major player in the orchestration of the immune response, as well as on regulatory CD4 T-cells (Treg), involved in the maintenance of the immune homeostasis. A defect of the latter is notably highly associated with autoimmune disorders. We therefore aimed at identifying TCR signatures of Teffs and Tregs as marker of immune-related diseases and to characterize their convergence and/or dissimilarities between diseases as well as their putative specificities.
203	Clustering-based prioritization of disease-associated clonotypes	Sebastian Valkiers	The T-cell receptor (TCR) repertoire comprises a comprehensive map of the immune system's past and present antigen exposure. As such, it provides promising opportunities for monitoring immune-related diseases and identification of highly personalized therapeutic targets. However, the complexity of this system imposes drastic challenges with regards to the interpretation. Grouping TCRs according to their reactivity towards the same epitope may remove much of the intrinsic redundancy present in immune repertoires. In addition, it may reveal events of clonal convergence that cannot be captured by analysis of the TCR repertoire at the level of individual sequences. Here, we present ClusTCR, a powerful computational ecosystem for fast and efficient clustering and analysis of TCR repertoires.

204	Highly Reproducible TCR profiling using RNA from rhesus macaque PBMC	John Beckford	Non-human primates (NHP) such as the rhesus macaque (Macaca mulatta) have long been key translational models in biomedical research because of their genetic and physiological similarity to humans1–4. Studies using rhesus macaques have contributed significantly to our understanding of T-cell responses to vaccines, cancer, and infectious diseases. More recently, these NHP have emerged at the forefront of COVID-19 vaccine research. Increasingly complex information can now be gleaned from immune system processes. High-throughput TCR sequencing (TCR-seq) profiles T-cell responses in exquisite detail. A comprehensive understanding of immune responses in such a closely related organism as the rhesus macaque would be a significant advance in science. Numerous tools exist for performing TCR-seq in human samples5–6, but equivalent tools for rhesus samples have been lacking. Because rhesus macaques often serve as surrogates in the lead-up to human studies, there is an industry need for a complete TCR-seq solution for these NHP samples. Due to strong species homology, our SMARTer® Human TCR a/b Profiling Kit v2 (TCRv2) can generate high-quality TCR sequencing libraries using human or rhesus macaque RNA.
205	TCRcloud: a screening tool for biologically and clinically relevant TCR repertoire landscapes.	Eric de Sousa	Deep TCR sequencing is an unbiased, comprehensive way to gauge the TCR repertoire. There are unmet needs to gauge the TCR diversity and anatomical location of the cellular response to checkpoint inhibitor therapies, termed 'clonal repopulation', as well as to trace infused T-cells after active cellular therapy. We present here TCRcloud, a tool to screen the 'TCR data warehouse' for biologically and clinically relevant patterns of the TCR repertoire that can be linked with clinical outcomes. To demonstrate its functionality, robustness, and feasibility, we applied TCRcloud to two different datasets: a cohort of PBMCs samples from patients throughout their course of COVID-19 and a cohort of tumor infiltrating lymphocytes (TIL) samples before and after in vitro expansion from patients with different types of cancer of the gastrointestinal tract. In patients with clinically relevant COVID-19 infection, we observed that the TRA/TRB numbers and TCR diversity steadily increased in association with clinical improvement, suggesting TCR diversity as an indicator of disease recovery. In samples from patients with cancer of the gastrointestinal tract, we observed that our in vitro expansion protocol maintains T-cell diversity establishing its potential use for adoptive cell therapy, along with the possibility to link individual TCR to its nominal target antigen and to follow-up individual TCRs after infusion.

206	Stitchr: a tool for making and modifying full-length T cell receptor sequences	Jamie Heather	Stitchr accurately generates full-length, properly spliced, coding nucleotide sequences for T cell receptors, given only the minimal V gene, J gene, and CDR3 sequences that are typically recorded. It can be operated via a graphical user interface or the command for a few sequences (e.g. for generating expression constructs), or via the command line for high-throughput volumes of TCRs provided in a tab-separated file as part of an analysis pipeline. Modification of the reference genes also allows the introduction of novel sequences, e.g. for rational TCR engineering. Stitchr offers a fast, reliable, and repeatable way to generate TCR sequences, which should be of use to many across the fields of TCR research. Presented by Jamie Heather (Massachusetts General Hospital)
207	pyTCR: a comprehensive and scalable platform for TCR-Seq data analysis to facilitate reproducibility and rigor of immunogenomics research	Serghei Mangul	T cell receptor (TCR) studies have grown exponentially with the advancement in the sequencing techniques of T cell receptor repertoire sequencing (TCR-Seq). The analysis of the TCR-Seq data requires computational skills to run analysis tools. However biomedical researchers with limited computational backgrounds face multiple obstacles to properly and efficiently utilizing bioinformatics tools for analyzing TCR-Seq data. Here we report pyTCR, a computational notebook-based platform for comprehensive and scalable TCR-Seq data analysis. Computational notebooks, which combine code, calculations, and visualization, are able to provide users with a high level of flexibility and transparency for the analysis. Additionally, computational notebooks are demonstrated to be user-friendly and suitable for researchers with limited computational skills. Our platform has a rich set of functionalities including various TCR metrics, statistical analysis, and customizable visualizations. The deployment of pyTCR of large and diverse TCR-Seq datasets will enable the effective analysis of large-scale TCR-Seq data with flexibility.
208	Phylogenetic methods for analyzing B cell migration, differentiation, and evolution over time	Kenneth Hoehn	B cells are an evolutionary system, undergoing rapid somatic hypermutation and antigen-driven selection as part of the adaptive immune response. B cell lineage trees inferred from B cell receptor sequencing data represent the history of mutations in a lineage, and can also link multiple forms of B cell diversity. For example, B cell lineage trees sampled from multiple tissues from the same subject could represent B cell migration between tissues. Trees sampled at different time points in the same subject represent clonal persistence over time. Here, we introduce new phylogenetic methods that use lineage trees to understand patterns of B cell migration, differentiation, and evolution over time. We demonstrate how this framework has been used to understand the role of B cell migration and differentiation in multiple immune conditions. Further, we show how longitudinally sampled data can be used to test for ongoing evolution in B cell lineages. Using large, publicly available AIRR-seq datasets, we demonstrate how different infections, vaccinations, and other conditions differ significantly in their ability to stimulate evolution over time in B cells. Some conditions, such as HIV infection, stimulate high strong signatures of B cell evolution, while others such as influenza vaccination produce a more compartmentalized response

			detectable only with direct germinal center sequencing. These methods are widely applicable and implemented in the R package dowser, available at https://bitbucket.org/kleinstein/dowser.
209	ABlooper: fast accurate antibody CDR loop structure prediction with accuracy estimation	Brennan Abanades Kenyon	"Antibodies are a key component of the immune system and have been extensively used as biotherapeutics. Accurate knowledge of their structure is central to understanding their antigen-binding function. The key area for antigen binding and the main area of structural variation in antibodies are concentrated in the six complementarity determining regions (CDRs), with the most important for binding and most variable being the CDR-H3 loop. The sequence and structural variability of CDR-H3 make it particularly challenging to model. Recently deep learning methods have offered a step change in our ability to predict protein structures. In this work, we present ABlooper, an end-to-end equivariant deep learning-based CDR loop structure prediction tool. ABlooper rapidly predicts the structure of CDR loops with high accuracy and provides a confidence estimate for each of its predictions. On the models of the Rosetta Antibody Benchmark, ABlooper makes predictions with an average CDR-H3 RMSD of 2.49 Å, which drops to 2.05 Å when considering only its 75% most confident predictions."
210	Intra-tumoral Adaptive Immunity Develops Acidic pH-dependent Tumor-specific Antibodies with	Genta Furuya	Important roles of humoral tumor immunity are often pointed out as well as cellular immunity. However, the precise characteristics of tumor-infiltrating B/plasma cells and intra-tumoral antibody biology have not been fully clarified. Our group has discovered that, by combining next-generation repertoire sequencing and biological experiments, the majority of tumor-infiltrating dominant antibodies in the human tumor environment recognize densely sulfated glycosaminoglycan (dsGAG) as their antigens (Cell Reports 2017; and a manuscript in review); however, background molecular mechanisms which produced such antibodies has been elusive. In this study, we explored the origins, developments, and

	Therapeutic Applicability.		therapeutic applicability of the anti-dsGAG antibodies. Examination of histological distributions of the anti-dsGAG B/plasma cells showed characteristic dominance in primary tumor sites, including tumor-associated tertiary lymphoid structures. The affinity of anti-dsGAG antibodies and their positive selection bias were significantly positively correlated, indicating that these antibodies were developed through the adaptive immune system in the tumor microenvironment. Intriguingly, the affinity of anti-dsGAG antibodies was drastically augmented by acidic or tumoral pH, further confirming their intra-tumoral development. We revealed that this dsGAG antigen was selectively expressed in tumor cells of almost all the cases of 7 cancer types examined (lung, breast, pancreas, stomach, large intestine, and prostate cancers), while this antigen was not expressed in normal human vital organs. We developed anti-dsGAG antibody-drug conjugate by conjugating maleimidocaproyl-valine-citrulline-p-aminobenzoyloxycarbonyl-monomethyl auristatin E and demonstrated its tumor-selective efficacy in vivo. Taken together, our study interestingly shows that the intra-tumoral adaptive B-cell immune system develops dominant tumoral-pH selective anti-dsGAG antibodies, and also provides a novel therapeutic modality applicable to various malignancies.
211	Redundancy, sensitivity, and predictability of developability parameters in natural, patent-submitted, and clinical-stage antibodies	Jahn Zhong	Monoclonal antibodies are of paramount therapeutic importance in treating autoimmunity, cancer, and infection. Developability, the set of physicochemical properties of an antibody relevant for manufacturing and success in clinical trials, is one of the key determinants for success during clinical testing. Many antibody developability studies have focused so far on extracting developability guidelines from successful mAb candidates, considering them as the “ground truth” of developability. However, the amount of available therapeutic antibody data might be insufficient to support the rules stated by previous work.In contrast, the natural immune system routinely produces highly effective non-immunogenic antibodies within days to weeks. Thus, we hypothesize that natural antibodies possess a favorable profile of developability parameters and that delineating the rules of natural antibody design could provide critical antibody developability design insights. Although many developability parameters can be computed from the antibody sequence and structure, the sequence and structural distributional landscape of the natural antibody repertoire has not yet been described. Here, we exploit a dataset of ≈2 million antibody sequences (mouse, human, all isotypes) and their corresponding computationally determined structures to build a large scale atlas of 42 sequence- and 48 structure-based computationally determined real-world relevant developability parameters. From this atlas, we aim to elucidate the rules of antibody development by employing mathematical, statistical, and machine-learning methods. Finally, by embedding patent-submitted antibodies in the repertoire-wide developability atlas, we will determine the similarity of patent-submitted antibody sequences to the repertoire of natural antibodies. Our work establishes redundancy, sensitivity, and predictability of antibody

			developability in native and human-derived antibodies. Exploiting the vast amount of available antibody high-throughput data will facilitate the derivation of the underlying rules of developability profiles which will guide discovery of antibody therapeutics.
212	Mining the immune system for protective antibodies	Jake Galson	Alchemab's unique approach identifies groups of resilient individuals, such as long-term cancer survivors, or those with genetic predisposition to develop neurodegenerative disease who remain healthy. The B cell receptor repertoires from these cohorts are deeply sequenced and compared with healthy control and disease progressor data to provide an understanding of how the immune response may help overcome or resist disease. This approach is being used to build a broad pipeline of protective antibodies for hard-to-treat diseases in oncology and neurodegeneration.

213	Multi-omic Investigation of the B cell Abnormality in Lassa Virus Infection	Jonathan Hurtado	Administration of neutralizing antibodies has been shown to protect from Lassa Virus (LASV) infection. However, distinct from most viral infections, the neutralizing antibody response against LASV takes months after convalescence to develop with major defects evident in the B cell maturation pathway. We are interested in uncovering these defects through investigation of the LASV-specific antibody repertoire and snapshots of acute infection of LASV-convalescent and acutely infected individuals by droplet-seq, respectively. For the prior, we developed a high-throughput single-cell genomics antibody discovery pipeline to maximize quantity and quality of antibody to antigen specificity data. With a stabilized trimeric glycoprotein complex (GPC) and monomeric GPC, we probed the B cell repertoires of six Lassa survivors and recovered 206 LASV-specific monoclonal antibodies (mAbs). More than half of the LASV-specific mAbs (50-80%) per subject were trimer-specific with the majority not class-switched. 5 of the 6 subjects had expanded mAb lineages but both expanded and unexpanded lineages lacked significant somatic hypermutation. Binding characterization of recombinantly expressed LASV-specific mAbs correlated with multi-omic data, validating our technical and analytical pipeline. For the latter, we expanded our focus on peripheral blood mononuclear cells and from a representative individual, we observed clonal expansion in multiple T cell lineages, exclusively in the CD8 T cells population. Contrasting the activation of CD4 T cells and B cell populations. Overall, this data confirms previously reported defects in the B cell maturation but narrows the cause to a lack of CD4 T cell help or upstream antigen presentation problems.
214	Potential role of immunoglobulin polymorphisms in celiac disease	Ida Lindeman	Objectives: Plasma cells (PCs) are terminally differentiated lymphocytes of the B-cell lineage secreting large amounts of antibodies. PCs are especially abundant in the lamina propria of the small intestine. Celiac disease (CeD) can be characterized by plasmacytosis in the lamina propria and the presence of high numbers of disease-specific PCs, of which most target the autoantigen transglutaminase 2 (TG2) and a smaller fraction target gluten peptides. The genetic basis for recognition of TG2 and gluten peptides by B-lineage cells in CeD has only been partially explored. Methods: We sorted single TG2-specific, gliadin-specific or unknown specificity PCs from 4 untreated CeD patients, 3 CeD patients on a gluten-free diet and 5 non-CeD controls. We then performed single-cell RNA-sequencing of the sorted cells and reconstructed their B-cell receptors (BCRs) with the computational tool BraCeR. Results: We observed antigen-dependent V-gene selection and stereotypic antibodies towards transglutaminase 2 and deamidated gliadin peptides (DGP), with limited accumulation of somatic mutations. Generation of recombinant DGP-specific antibodies followed by ELISA revealed a key role of a conserved heavy chain residue that displays polymorphism, suggesting that immunoglobulin gene polymorphisms may influence CeD-specific antibody responses.

			Conclusion: Our findings support previous reports showing that germline-encoded residues are important for binding to DGP and TG2 in CeD. Future studies will reveal if genetic polymorphisms within the coding and/or non-coding regions of the BCR loci predispose to CeD by skewing the naïve B cell repertoire.
215	Characterization of Extensive Diversity In Immunoglobulin Light Chain Variable Germline Genes Across Biomedically Important Mouse Strains	Justin Kos	The light chain immunoglobulin genes of biomedically relevant mouse strains are poorly documented in current germline gene databases. We previously showed that IGH loci of wild-derived mouse strains representing the major mouse subspecies contained 247 germline IGHV sequences not curated in the international ImMunoGeneTics (IMGT) information system, which is the most commonly used database that curates the germline repertoires used for sequence alignment in AIRR-seq analysis. Despite containing levels of polymorphism similar to the IGH locus, the germline gene content and diversity of the light chain loci have not been comprehensively cataloged. To explore the extent of germline light chain repertoire diversity across mouse strains commonly used in the biomedical sciences, we performed AIRR-seq analysis and germline gene inference for 18 inbred mouse strains, including the four wild-derived strains with diverse sub-species origins. We inferred 1582 IGKV and 63 IGLV sequences, representing 459 and 22 unique IGKV and IGLV sequences. Of the unique inferred germline IGKV and IGLV sequences, 67.8% and 59%, respectively, were undocumented in IMGT. Across strains we observed germline IGKV sequences shared by three distinct IGK haplotypes and a more conserved IGLV germline repertoire. In addition, J gene inference indicated a novel IGK2 allele shared between PWD/PhJ and MSM/MsJ and a novel IGLJ1 allele for LEWES/EiJ and IGLJ2 allele for MSM/MsJ. Finally, a combined IGHV, IGKV, and IGLV phylogenetic analysis of wild-derived germline repertoires displayed reduced germline diversity for the light chain repertoire compared to the heavy chain repertoire, suggesting potential evolutionary differences between the two chains.
216	The IGX Platform: from repertoire annotation to antibody discovery	Thijs Maas	A large part of the drugs currently under development is antibody-based. Immune repertoire profiling using high-throughput sequencing is applied to profile the candidate pool and to select the best candidate for further development. We present the IGX Platform, a cloud-based analysis environment for immune repertoire data in general and for antibody discovery in particular. We integrate best-in-class tools into the platform so as to enable researchers to simply select better antibodies faster. The platform can be extended to integrate immune repertoire data in other formats or with other tools and apps to perform novel analyses.

217	B cells, T cells and Sepsis: Understanding the Adaptive Immune Response	Lauren Overend	Sepsis is defined as a life-threatening organ dysfunction caused by a dysregulated host immune response to infection. Globally, sepsis is responsible for 27% of adult intensive care unit (ICU) admissions and has high mortality rates (~20-30%), despite this no sepsis-specific therapies currently exist. A greater understanding of disease mechanisms is paramount for reducing mortality and associated socio-economic burden. Mortality abnormally increases 2-3 months after apparent recovery and continues to rise for several years; death is frequently due to healthcare-associated/secondary infection. This implicates the adaptive immune system in the long-term detrimental immune alterations occurring in sepsis. To elucidate the adaptive immune response in sepsis we have generated the first comprehensive longitudinal study of B/TCR cell receptor (B/TCR) repertoires. Using high-throughput multiplex PCR we have amplified and sequenced the repertoire of ~75 patients with sepsis at multiple time points following ICU admission. We then applied weighted correlation network analysis and tensor decomposition and demonstrate that these techniques can identify modules incorporating the effect of several immune repertoire features (e.g. diversity, V gene usage), both across and within receptor chains. Using conventional statistical analysis we have compared modules to clinical variables such as sepsis response signature endotypes, shock and survival status and identified modules differing across patient groups, in a time-dependent manner, and between health and disease. We have validated our findings using BCR/TCR VDJ contigs constructed from matched bulk RNA sequencing data using TRUST4, from within a larger cohort of 903 sepsis samples sequenced by the Genomic Advances in Sepsis Study (GAinS).

218	FAIR Data Curation Promotes Rapid Response to COVID-19: The iReceptor Project promotes data reuse by providing access to ~1 Billion antibody/B-cell and T-cell receptor sequences from COVID-19 patients	Felix Breden	The AIRR community (www.AIRRCommunity.org) has developed protocols and standards for curating, analyzing and sharing AIRR-seq data (antibody/B-cell and T-cell receptor sequences from Adaptive Immune Receptor Repertoires) through the AIRR Data Commons, a set of geographically distributed repositories following the AIRR Community's metadata standards. The iReceptor Gateway (gateway.ireceptor.org) is designed to query the AIRR Data Commons for specific "metadata", e.g. "find all repertoires from studies of ovarian cancer" or for specific sequence annotation features (e.g. CDR sequences). The Gateway then aggregates these repertoires from multiple repositories for further analysis by sophisticated AIRR-seq algorithms. iReceptor Gateway provides access to >5 billion receptor sequences for data reuse, analysis and sharing according to FAIR principles. AIRR-seq data can describe the adaptive immune response to SARS-CoV-2 infection in exquisite detail, and comparison and analysis of these data across studies and institutions can greatly contribute to the development of anti-COVID diagnostics and therapeutics, including vaccines. Many COVID-19 researchers responded following the AIRR Community's call for increased data sharing to help overcome the COVID-19 pandemic, even providing data from studies during pre-print stage. Currently the AIRR Data Commons includes ~1 billion AIRR-seq sequences from 26 studies of COVID-19 patients and 4 studies of healthy individuals vaccinated against COVID, all curated according to the AIRR Community Standards. Initial analyses from these studies indicate shared receptor sequences, restricted V gene usage, and characteristic patterns of TCR and BCR clonal expansion among AIRR-seq repertoires from COVID-19 patients. The ability to confirm such patterns by comparing results across studies and institutions will be greatly facilitated by integrated searches across the AIRR Data Commons through the iReceptor Gateway. For more information on obtaining or sharing COVID-19 data contact support@ireceptor.org.
219	Immunophenotyping of TCR and BCR Clonotypes	Michael Makhanov	TCR and BCR repertoire profiling holds great potential for understanding disease mechanisms and for development of new therapeutics in infectious disease, autoimmunity and immuno-oncology. However, this potential could be greatly improved by combining information about receptor clonotypes with immunophenotypes of T and B cells. To facilitate these studies, we developed a novel technology for combined profiling of all human TCR and BCR variable regions and phenotypic characterization of immune cells. The developed TCR/BCR immunophenotyping method involves multiplex RT-PCR amplification and sequencing of CDR3 regions of TCR and BCR genes and a set of the most informative T- and B-cell phenotyping genes. Bioinformatics analysis of NGS data allows profiling of TCR/BCR clonotypes, and identification of major immune cell subtypes and their activation status. Preliminary studies indicate

			the assay has unparalleled throughput, sensitivity, and improved cost-effectiveness for high-throughput immunity biomarker discovery applications.
220	Phenotypic and phylogenetic characterization of broadly neutralizing antibody development	Collin Joyce	Longitudinal studies of broadly neutralizing antibody (bnAb) evolution from cases of natural HIV infection are highly important for aiding rational vaccine design. Here we characterized the development of a family of bnAbs from Protocol C donor 39 (PC39) that targets the N332 supersite using a combination of memory B cell sorting, bulk antibody repertoire and single-cell RNA sequencing (scRNAseq). Phylogenetic analysis revealed the lineage to contain multiple independent insertions within the heavy chain complementarity-determining region 1 (CDRH1). The most common insertion lengths were 4, 5 and 11 amino acids and each independent insertion was paired to a separate motif in the CDRH3. Fragment antigen-binding (Fab) crystal structures showed evidence of convergent evolution between distinct lineage branches. Natively paired pre-insertion lineage members isolated via scRNAseq were unable to bind to HIV Env or gp120. Gene expression analysis revealed heterogeneity in the memory B cell compartment and defined the phenotype of bnAb lineage members. This work offers actionable information for use in vaccine design by providing a snapshot of the molecular and cellular features involved in bnAb lineage maturation.
221	Characterization of the human immune response to SARS-CoV-2 infection and vaccination using Mapping of B Cell Receptor Sequences to Antigen Specificity	Jan Michler	Using Mapping of B Cell Receptor Sequences to Antigen Specificity through Sequencing, we follow SARS-CoV-2 spike protein specific clones in convalescent infected and vaccinated individuals longitudinally. We describe how BCRs specific for different variants of the spike protein (Wuhan-Hu-1, Beta, Delta, Omicron) emerge and evolve over time, and investigate the heterogeneity in memory B cell phenotypes that these antigen specific B cells adapt.

	Statistical analysis of synthetic AIRR-datasets to guide the development and benchmarking of AIRR-based machine learning	Mariia Chernigovskaia	Machine-learning on adaptive immune receptor repertoire (AIRR) data enables diagnostics and therapeutics design. Experimental AIRR data, labeled with immune events such as immune state or antigen specificity, does not generally represent ground truth data since complete knowledge on immune-event-related as well as unrelated (potential confounding factors) (sub)sequences is unavailable. The lack of large-scale ground truth data renders the development and benchmarking of robust, explainable, and interpretable machine learning approaches unfeasible. To address the lack of large-scale ground truth data, simulation frameworks have been recently developed. However, it remains unclear to what extent a simulation approach impacts a dataset both in terms of its nativeness as well as it classifyability. To address this knowledge gap, we developed the LIgO software suite that enables the modular ("lego"-like) assembly of immune-event-labeled synthetic AIRR-datasets for the development and benchmarking of AIRR-based machine learning. Specifically, LIgO contains (i) different methods for simulating immune events and confounding factors, and we provide a data-driven discussion of the advantages and disadvantages of each simulation approach for user guidance. In three case studies that simulate scenarios that are as of yet not available experimentally, we explore prediction performance and signal recovery using baseline machine learning methods on repertoire-scale paired chain data, repertoire-scale antigen-annotated data, and different sequence-based data leakage scenarios. The LIgO software is integrated into the immuneML ecosystem and outputs AIRRCompliant data for ready use in any AIRRCompliant machine learning software.
	Improving generalization of machine learning-based biomarkers through causal modeling: application to adaptive immune receptor repertoires	Milena Pavlovic	Machine learning is increasingly used to discover diagnostic and prognostic biomarkers from high-dimensional molecular data. However, a variety of factors related to experimental design may affect the ability to learn generalizable and clinically applicable diagnostics. Here, we argue that a causal perspective improves the identification of these challenges, and formalizes their relation to the robustness and generalization of machine learning-based diagnostics. To make for a concrete discussion, we focus on a specific, recently established high-dimensional biomarker – adaptive immune receptor repertoires (AIRRs). We discuss how the main biological and experimental factors of the AIRR domain may influence the learned biomarkers and provide easily adjustable simulations of such effects. In conclusion, we find that causal modeling improves machine learning-based biomarker robustness by identifying stable relations between variables and by guiding the adjustment of the relations and variables that vary between populations.

	What Do We Learn from Millions of HLA Specific Public T cell Receptors	Jabran Zahid	Using the T cell repertoire of ~4,000 subjects with genotyped HLAs, we identify millions of public TCRs associated with specific HLAs. Using these public TCRs, we are able to produce models with high sensitivity and specificity for predicting all commonly occurring HLAs. We show that TCR specificity is typically to the HLA allele (i.e. two-field resolution) and to the alpha+beta heterodimer for class II HLAs. We identify a set of class II HLAs that do not form stable heterodimers. By examining the distribution of the alpha and beta subunits in our sample, we conclude that these heterodimers form via trans-complementary pairing of the alpha and beta chains. The remarkable ability to identify and associate millions of TCRs to HLAs and validate such associations via a predictive model provides a firm foundation for the use of the T cell repertoire as a diagnostic platform and demonstrates the power of using a large samples of T cell repertoires for exploring fundamental issues in human immunology.
online	Dandelion: analyzing single-cell BCR/V(D)J data from 10X	Kelvin Tuong	dandelion is a python package for single-cell BCR-seq analysis visualisation for data from 10X Genomics. It implements a data structure that allows for interoperability with other AIRRCompliant tool kits such as immcantation and scirpy. It performs reannotation of VDJ with igblastn, constant genes reassignment with blastn using a custom curated reference, and performs quality checks to ensure that cells and contigs are matched up correctly. It also implements a network-based approach to quantify and visualise BCR clonal diversity and mutation landscape. Post-processed data from dandelion can be smoothly transferred to the popular single-cell transcriptome analysis package, scanpy/AnnData, to enable simultaneous exploration of BCR-seq data and RNA-seq data. https://www.github.com/zktuong/dandelion https://sc-dandelion.readthedocs.io/ Example notebook: https://colab.research.google.com/github/zktuong/dandelion/blob/master/container/dandelion_singularity.ipynb
online	VDJServer Community Data Portal	Scott Christley	VDJServer is a public analysis and data sharing portal for adaptive immune receptor repertoire sequencing (AIRR-seq) data. The VDJServer Community Data Portal (CDP) is a data repository within the AIRR Data Commons (ADC), which is an internationally distributed set of data repositories for public query and download of AIRR-seq data. The ADC contains post-processed, annotated sequences and study metadata that conforms to the AIRR Community Data Standard, and VDJServer is the primary US-based data repository for NIH-funded studies in the ADC with >3000 repertoires and >2.5B annotated sequences.
online	TCRMatch	Raphael Trevizani	TCRMatch is a tool for identifying epitope candidates for TCR sequences with unknown specificity based on sequence similarity to TCRs with known epitopes.

online	Absolut!: Unconstrained generation of synthetic 3D-antibody-antigen complexes enables machine-learning benchmarking of antibody specificity prediction	Philippe Robert	Machine learning (ML) is a key technology for accurate prediction of antibody-antigen binding. Two orthogonal problems hinder the application of ML to antibody-specificity prediction and the benchmarking thereof: (i) The lack of a unified ML formalization of immunological antibody specificity prediction problems and (ii) the unavailability of large-scale benchmarking datasets. Here, we developed the Absolut! software suite that allows the parameter-based unconstrained generation of synthetic lattice-based 3D-antibody-antigen binding structures with ground-truth access to conformational paratope, epitope, and affinity. Absolut!-generated datasets contain critical sequence and structural levels of complexity to assist development of ML models for antibody-antigen binding prediction. Absolut! helps translating immunological antibody specificity prediction problems into ML tasks and Absolut!-generated datasets are parametrizable to mirror possible experimental conditions or to benchmark ML power. Therefore, the Absolut! framework enables the development and benchmarking of ML strategies for biotherapeutics design.
	IgTreeZ: A toolkit for immunoglobulin gene lineage tree analysis	Hadas Neuman	Somatic hypermutation (SHM) is an important diversification mechanism that plays a part in the creation of immune memory. Immunoglobulin (Ig) variable region gene lineage trees were used over the last four decades to model SHM and the selection mechanisms operating on B cell clones. We present IgTreeZ (Immunoglobulin Tree analyZer), a python-based tool that analyses many aspects of Ig gene lineage trees and their repertoires. Using simulations, we show that IgTreeZ can be reliably used for mutation and selection analyses. We used IgTreeZ on empirical data, found evidence for different mutation patterns in different B cell subpopulations, and gained insights into antigen-driven selection in corona virus disease 19 (COVID-19) patients. Most importantly, we show that including the CDR3 regions in selection analyses – which is only possible if these analyses are lineage tree-based – is crucial for obtaining correct results. Overall, we present a comprehensive lineage tree analysis tool that can reveal new biological insights into B cell repertoire dynamics.

AIRR-C Committee Sub-committees and Working Groups Reports

AIRR-C Committee Sub-committees and Working Groups reports introduced during the AIRR-C Meeting VI are shared on the following order:

2022 AIRR-C Biological Resources WG Report

2022 AIRR-C Common Repository WG Report

2022 AIRR-C Communications SC Report

2022 AIRR-C Diagnostics WG Report

2022 AIRR-C Exec SC Report

2022 AIRR-C GLDB WG Report

2022 AIRR-C IARC Report

2022 AIRR-C Legal _ Ethics WG Report

2022 AIRR-C Meetings SC Report

2022 AIRR-C Software WG Report

2022 AIRR-C Standards WG Report

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

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

Current
Date of this report: April 20, 2022

SC/WG Name: Biological Resources WG

SC/WG Co-leaders: Anne Eugster and Johannes Trück

SC/WG Active Members (list):
Houda Alachkar, Davide Bagnara, John Beckford, Anne Eugster (Co-lead), Nina Luning Prak, Encarnita Mariotti-Ferrandiz, Cinque Soto, Johannes Trück (Co-lead), Nidhi Gupta, Andrew Farmer, Shaveta Goyal, Theam Soon Lim, Wenming Xiao

Purpose: To provide the AIRR Community with biological calibrators and reagents for evaluation of AIRR-seq.
The Biological Resources Working Group is responsible for coordinating the development of reference samples that can be used as controls. The working group aims at reaching out to established organizations such as NIST and Genome in a Bottle, as well as companies to help encourage ease of use and broad adoption.

Goals:
The overarching goal of the working group is to be able to recommend a set of biological standards that can be used for normalization of data sets. Our specific aims are:

- Generate and share in-line and external standards for repertoire analysis
- Generate data using standards and different library prep methods
- Share data for standard method and analysis pipeline comparison and evaluation

Long-term vision and how WG products integrate with the AIRR-C mission:
The long-term aim of the Biological Resources WG is to provide AIRR biological standards and protocol recommendations to the scientific and biomedical community.

Products (if any):
Publication in eLife (Trück et al., 2021).

Chapters in Methods in Molecular Biology (accepted).
Grant application for methods comparison from TAbS (awaiting final funding decision)

Resources (if any):
Controls (RNA templates, spleen sample DNA and RNA and cell lines) in preparation

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

- The WG has led efforts to publish a paper on community opinions regarding needs and challenges for AIRR-seq standards and standardization (Trück et al., 2021).
- In addition, we have participated in the publication of several chapters in Methods in Molecular Biology, Vol. 2453, including 3 wet bench chapters, wet bench methods and a dry bench overview. These chapters provide detailed guidance to the community regarding current state-of-the art methods in AIRR-seq.
- WG members planned and submitted a grant application to compare different BCR sequencing methods using three different types of standards: synthetic RNA templates, spleen DNA/RNA and cell line mixtures.
- WG members are planning a similar effort for TCR sequencing standards comparison, starting from the methods that performed best in the Nature Biotechnology publication that was led by several of the WG members (Barennes et al., 2021).
- Members of the WG participated in the EuroClonality WG and meetings.
- In addition, members of the WG have joined a larger effort spearheaded by Wenming Xiao from the Food and Drug Administration called BCR-SEQC, in which several wet bench methods, standards and data analysis pipelines will be compared and evaluated. The overarching study objectives, quoted from Wenming Xiao, are to elucidate current limitations, address fundamental technical needs, provide standard reference samples and data sets, and establish best practices for constructing BCR repertoires from NGS data.

Proposed plans for the coming interval:

Purpose: To provide the AIRR Community with biological calibrators and reagents for evaluation of AIRR-seq.

The Biological Resources Working Group is responsible for coordinating the development of reference samples that can be used as controls. The working group plans to gain insights from established organizations such as NIST and Genome in a Bottle, as well as companies to help encourage ease of use and broad adoption.

Specific Goals for the Next Cycle:

- Generation of data from different BCR sequencing methods using three different types of standards: synthetic RNA templates, spleen DNA/RNA and cell line mixtures.
- Generation of data from TCR sequencing standards comparison, starting from the methods that performed best in the Nature Biotechnology publication that was led by several of the WG members (Barennes et al., 2021).

- Continued participation in FDA BCR-SEQC initiative
- Continue to participate in EuroClonality WG and meetings
- Generate data and publish a paper about sequencing depth and sample optimization for unique molecular identifiers (UMIs)

Products (if any):
- Data sets for BCR and TCR sequencing methods
- Recommendations for how to use standards in different AIRR-seq methods
- UMI paper

Resources (if any):
- Data sets for BCR and TCR sequencing methods
- Recommendations for how to use standards in different AIRR-seq methods

Long-term vision and how WG or SC products integrate with the AIRR-C mission:
The long-term aim of the Biological Resources WG is to provide AIRR biological standards and protocol recommendations to the scientific and biomedical community.

Proposed SC/WG Co-leaders for 2022-2023 Cycle:

Anne Eugster

Johannes Trück

Nina Luning Prak

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

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

Current
Date of this report: April 8, 2022

SC/WG Name: Common Repository Working Group

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

SC/WG Active Members (list): Attended Meetings
Brian Corrie, Artur Rocha, Scott Christley, Christian Busse, George Blanck, Lindsay Cowell, Eric Waltari, Felix Breden, Katharina Imkeller, Ulrik Stervbo, Veronique Giudicelli, Marco Amaro Oliveira, David Klatzmann, Adrien Six, Enkelejda Miho, Corey Watson

Purpose:

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

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

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

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

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

Products (if any):

AIRR Data Commons

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

Resources (if any):

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

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

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

Feature T Cell Exhaustion and Dedifferentiated Monocytes in Cerebrospinal Fluid," Immunity, vol. 54, no. 1, pp. 164-175.e6, Jan. 2021.
        - R. A. Porritt et al., "HLA class I–associated expansion of TRBV11-2 T cells in multisystem inflammatory syndrome in children," J. Clin. Invest., vol. 131, no. 10, May 2021.
        - A. J. Schmitz et al., "A vaccine-induced public antibody protects against SARS-CoV-2 and emerging variants," Immunity, vol. 54, no. 9, pp. 2159-2166.e6, Sep. 2021.
        - J. Y. Humrich, J. P. Bernardes, R. J. Ludwig, D. Klatzmann, and A. Scheffold, "Phenotyping of Adaptive Immune Responses in Inflammatory Diseases," Front. Immunol., vol. 11, p. 3010, Nov. 2020.
        - F. P. Caruso, G. Scala, L. Cerulo, and M. Ceccarelli, "A review of COVID-19 biomarkers and drug targets: resources and tools," Brief. Bioinform., vol. 22, no. 2, pp. 701–713, Mar. 2021.
- Repository growth
    - 7 repositories in total currently, four "recent" repositories
        - iReceptor Public Archive – iReceptor/SFU
        - iReceptor COVID-19 – iReceptor/SFU
        - VDJServer – VDJServer/UTSW
        - VDJBase (Israel) – Gur Yaari/Bar Ilan
        - sciReptor (Germany) – Christian Busse/DKFZ
        - NICD (South Africa) – Cathrine Scheepers/NICD
        - UKM (Germany) – Nicholas Schwab/Muenster
- ADC API extensions
    - Existing: /repertoire, /rearrangement
    - Version 1.4: /clone, /cell, /expression
    - Longer term: /receptor, /stats, /germline

Proposed plans for the coming interval:
Purpose:

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

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

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

Products (if any):

- AIRR Data Commons
- ADC API
- ADC Registry

Resources (if any):

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

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

Same as above…

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

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

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

Current
Date of this report: April 20, 2022

SC/WG Name: AIRR-C Communications SC

SC/WG Co-leaders: Susanna Marquez and Jean-Philippe "JP" Bürckert

SC/WG Active Members (list): Pam Borghardt, Jean-Philippe "JP" Bürckert (Co-lead), Victor Greiff, Susanna Marquez (Co-lead), Kira Neller and Simon Schafer.

Purpose:
The AIRR-C Communications Sub-committee is responsible for communicating activities of the AIRR-C to both the AIRR Community and the general research community.
https://www.antibodysociety.org/the-airr-community/airr-subcomittees/communications-sub-committee/

Goals:
- Organization of professionally produced webinars, which will be on AIRR biology, AIRR data analysis and AIRR software usage.
- Develop a strategy for widespread adoption of AIRR-C standards (in collaboration with relevant WG's & SC's.
- Support the Meeting SC in the dissemination and communication for AIRR-C Meeting VI in December 2021.
- Create a news showcase post on each WG after significant WG milestones are met.
- Continue to support the AIRR-C in all communication and dissemination activities.

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

The Communications SC:
- Supports the development and maintenance of the AIRR-C corporate identity, updating and maintaining AIRR-C web presence and facilitating external and internal communication.
- Promotes the initiatives and work done by other WG and SC to increase their visibility, strengthen the position of the AIRR-C as a leading organization in the field, and attract new members.

Products (if any):

Resources (if any):

The Communications SC is currently maintaining the following resources:

Resource	Link	Main purpose	Analytics
Website	https://www.antibodysociety.org/the-airr-community/	Dissemination of AIRR-C related information	Page view statistics 19,337 See appendix for more details (AIRR-C website pageviews and AIRR-C Webinar pageviews).
YouTube	https://www.youtube.com/c/AIRRCommunity	Dissemination of AIRR-C videos (meetings, webinars,...)	Subscribers: 323 Total views: 5281 See Appendix for more details (Monthly views by video)
Twitter	https://twitter.com/airr_community	Dissemination of AIRR-C related information	1066 followers. Over the last 28 days, AIRR-C tweets have been seen 1782 times and the profile has been visited ≈ 1,357 times.
LinkedIn	https://www.linkedin.com/company/the-airr-community	Dissemination of AIRR-C related information	101 followers, 8 posts. dedicated hashtags: #airrcommunity and #theairrcommunity See Appendix for more details (visitors and page views).
Slack	http://airrcommunity.slack.com	Internal and external communication, complementary to https://b-t.cr/ forum	Members: 139 (more than 100 new members since last report)

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

1. Goal: Organisation of professionally produced **webinars** (financially supported by TAbS):
   We have published announcements, news items and videos of:
   a. The Adaptive Immune Receptor Repertoires Webinar Series.
      i. 2021-04-06. Steps in data processing and analysis of adaptive immune receptor repertoires: best practices, pitfalls, and future directions. Victor Greiff.
      ii. 2021-05-04. The AIRR Data Commons: 4 billion reasons to store, analyze and share antibody/B-cell and T-cell receptor repertoire data. Felix Breden, Brian Corrie, Kira Neller and Scott Christley.
      iii. 2021-06-03 and 2021-06-15. Fundamentals of the Immune System. Jamie Scott. Links to Part 1 and 2.
      iv. 2021/10/07 Easy, fast, and practical AIRR analysis. Exploration of single-cell and bulk immune repertoire data in R using Immunarch with application to immunotherapy. Vadim I. Nazarov.
      v. 2021/11/09 Reconstruction & analysis of B cell lineage trees from single cell data Immcantation. Kenneth B. Hoehn and Susanna Marquez.
      vi. 2022/02/10 Computational mining of immune receptor germline gene loci variation Martin Corcoran, Ayelet Peres, Oscar Rodriguez
   b. Sessions from the AIRR Community Meeting V: Pre-Meeting "AIRR-seq in the Pandemic. " The meeting happened in 2020-12, but recorded sessions were published in April 2021.
   c. FOCIS 2021.

2. Goal: Develop a **strategy** for widespread adoption of AIRR-C standards.
   Updates:
   a. This is a complex, long-term goal, to be developed in collaboration with AIRR-C Exec and with input from other SC/WG members.

3. Goal: **Support the Meeting SC** in dissemination and communication for AIRR-C Meeting VI.
   Updates:
   a. New Dates. AIRR Community Meeting VI: "Exploring New Frontiers"
   b. AIRR Community Meeting VI: "Exploring New Frontiers"
   c. Follow us on Twitter: #airrc6

4. Create a **news showcase post** on each WG after significant WG milestones are met
   a. Needs and challenges of biological controls for AIRR-sequencing standardization: an AIRR-C review published in eLife
   b. COVID-19 AIRR-seq Vaccine Data Available!

c. The AIRR-C Diagnostics Working Group published on the future of blood testing using AIRR technology
d. A newly certified AIRR-compliant software tool: The Immcantation Framework
e. Three newly certified AIRR-compliant software tools: ImmuneML, CompAIRR, and Dandelion

5. Continue to support the AIRR Community in all **communication and dissemination activities**
   a. The AIRR Community has started a Webinar Series
   b. Interns looking for Integrated Immunology Projects
   c. AIRR Community 2022 Executive Sub-committee Election
   d. Announcing AIRR Community Service Prize 2022 Nominations!
   e. AIRR-C Meeting VI – Registration Is Live!
   f. AIRR Community Meeting VI Event Information – May 17-19, 2022
   g. Glossary of terms (now led C. Busse)
   h. In progress with others: improve internal communications (email, google groups), adoption strategy.
   i. Support the initiation of the AIRR-C Diagnostics WG podcast series (logo, website) onairr.airr-community.org.
   j. In progress: transferring B-T.CR forum to AIRR-C
   k. Merchandising. Launching soon, an AIRR-C online shop where T-shirts and other items with the AIRR-C logo can be ordered.

Proposed plans for the coming interval:

Purpose:

The AIRR-C Communications Sub-committee is responsible for communicating activities of the AIRR-C to both the AIRR Community and the general research community.
https://www.antibodysociety.org/the-airr-community/airr-subcomittees/communications-sub-committee/

Goals:

1. Finalize an AIRR-C **adoption and sustainability** strategy to be developed in collaboration with AIRR-C Exec and with input from other SC/WG members.

2. Improve our communication strategy and networking with **industry** members. We want to build a relationship that will facilitate the organization of sponsored events and activities, will help our students access companies to search for industry opportunities, and the industry members will be able to access a diverse pool of scientists (from junior to senior).
   a. Learn more about our industry members and any specific unmet needs they have, to try to fill the gap.
   b. Do market research (what companies are out there, working in the AIRR field) and identify potential new members. This will be done in collaboration with the Software WG, and will contribute to gaining a better

understanding of the software landscape and promoting the adoption of the AIRR-C standards.

c. Improve our communication strategy to attract new members.
    - Create posters, brochures, cards, and other corporate materials that can be used in conferences to promote AIRR-C.
    - Collect AIRR-C resources particularly useful for industry members and potential members and present them in a way that is useful and attractive for them. Examples are a newsletter with updates relevant to them, and a section on the AIRR-C website that easily helps find tools, experts, and webinars.

d. Organize meetups industry/students, to help AIRR-C mentees explore industry opportunities.

3. Collaborate with Diagnostics WG, Legal and Ethics WG, and Meetings SC in organizing an **event** covering AIRR+AI/ML+legal+regulatory aspects.
4. Continue to:
    a. Organize professionally produced **webinars** (financially supported by TAbS), which will be on AIRR biology, AIRR data analysis and AIRR software usage. We believe these webinars will further the AIRR-C mission, which is to disseminate AIRR-relevant knowledge and best practices.
    b. **Support the Meeting SC** in dissemination and communication for AIRR-C Meeting VII (date and location TBD).
    c. Continue to create an AIRR-C-related **news** showcase post on each WG after significant WG milestones are met
    d. Continue to **support the AIRR Community in all communication** and dissemination activities: news on the website, podcasts, tweets…
5. Bring more members to the Communications SC.

Products (if any):

Resources (if any):

In addition to the resources that the Communications SC maintains (Website, YouTube, Twitter, LinkedIn, Slack), we will create a directory of companies that develop their activity in the field of AIRR analysis. We will create a web page to collect resources of particular interest to industry members, and a newsletter to share relevant updates.

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

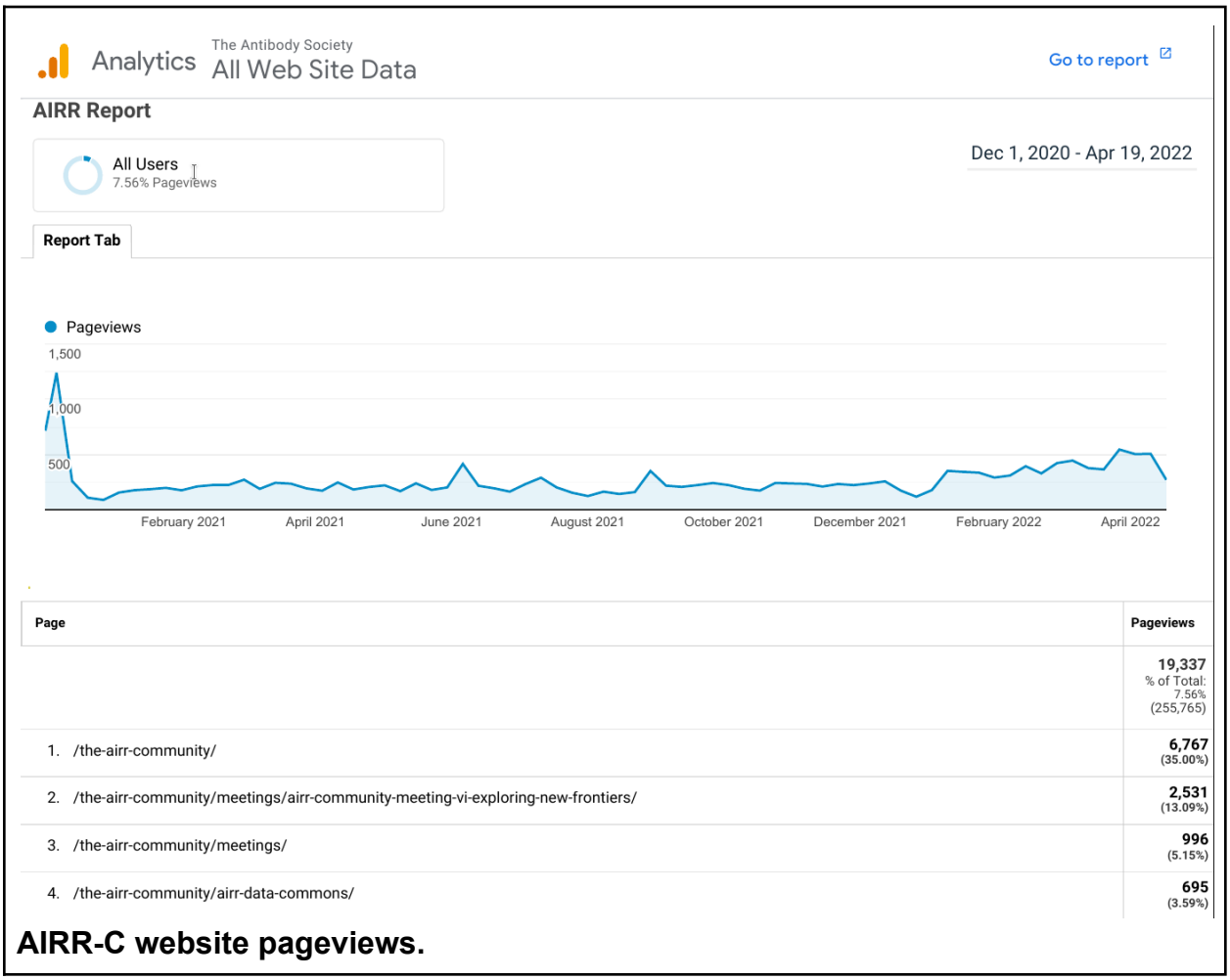
The new goals for the next ~18 months period have been strategically selected to improve the corporate image of the AIRR-C under the industry lens and contribute to the sustainability of the Community. By strengthening the communication with industry, and learning the specific needs of these members, we will be able to improve our efforts targeted to them and attract more members. This will enlarge the Community and the pool of potential sponsors to support our activities. At the same time, we will provide our non-industry members the opportunity to be exposed to new opportunities outside

academia.

Proposed SC/WG Co-leaders:

Susanna Marquez and Jean-Philippe “JP” Bürckert

Appendix



AIRR-C website pageviews.

Analytics The Antibody Society
All Web Site Data

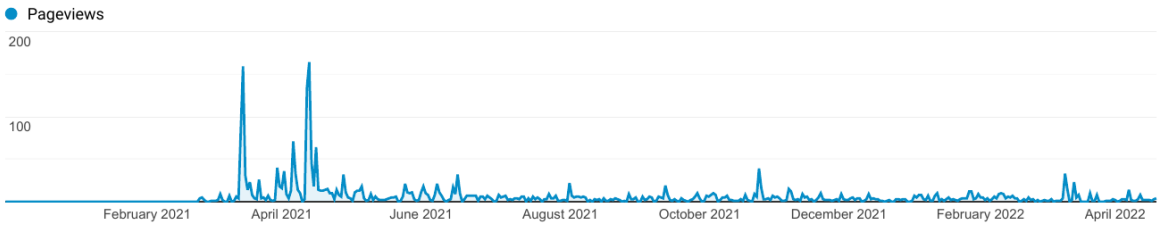
Go to report

AIRR Webinar

All Users
1.10% Pageviews

Dec 1, 2020 - Apr 19, 2022

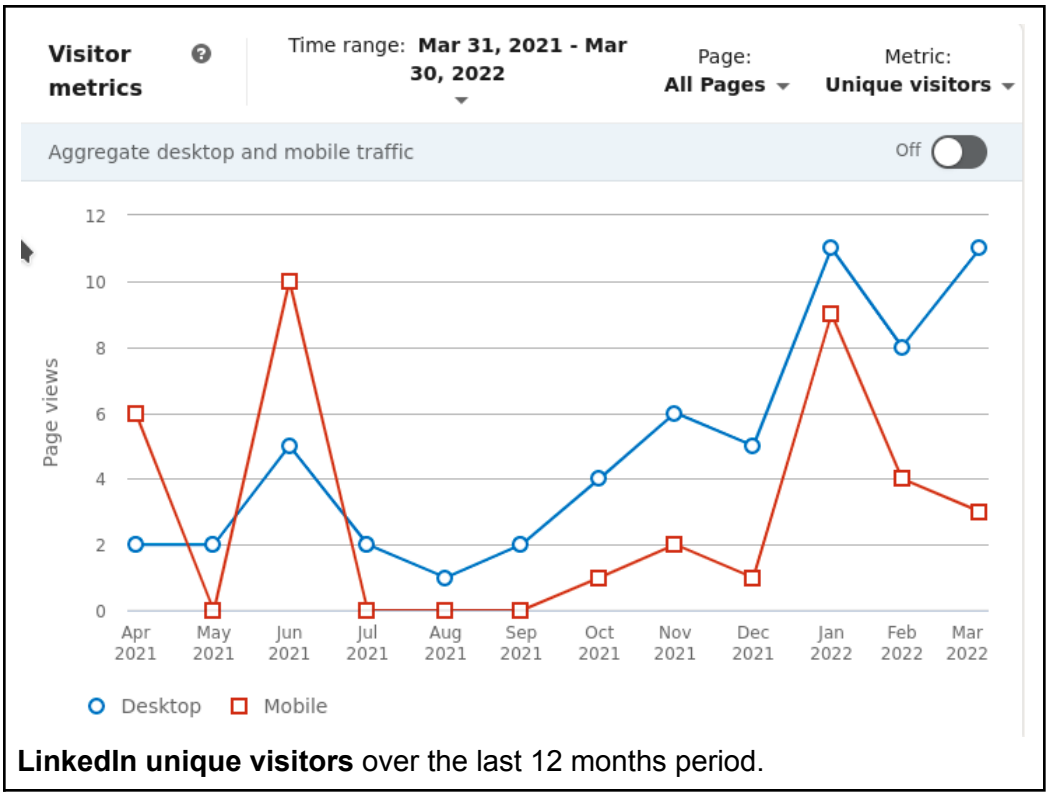
Report Tab



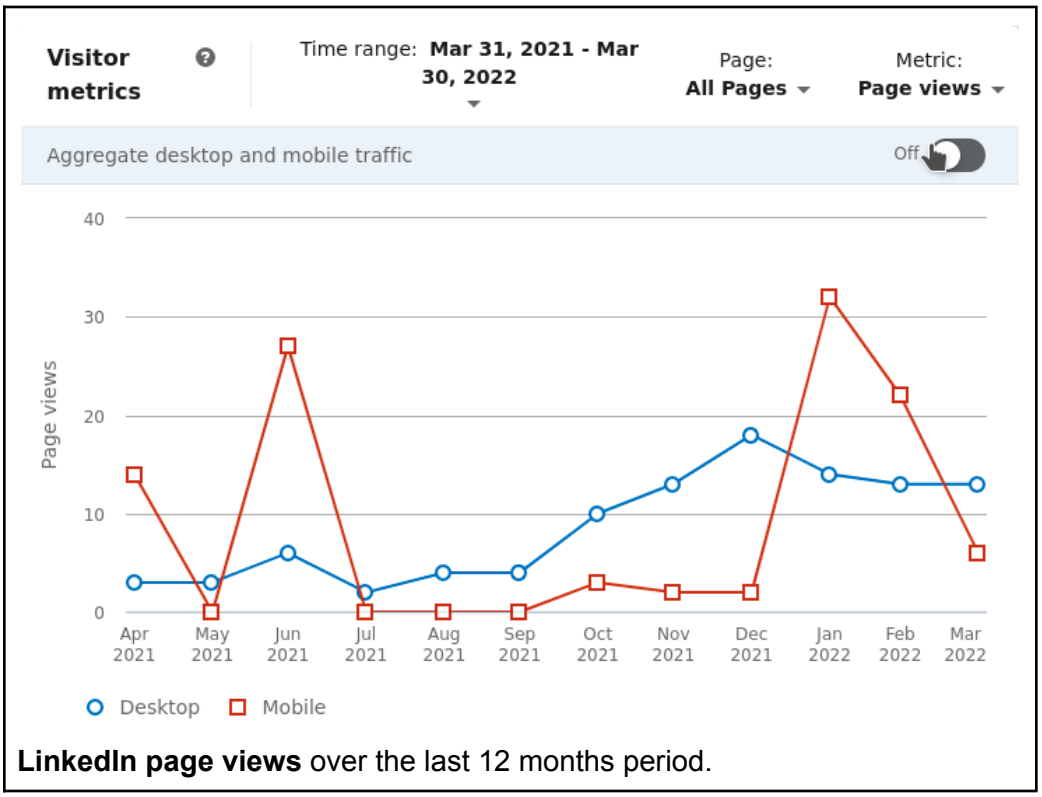
Page	Pageviews
	2,826 % of Total: 1.10% (255,765)
1. /learningcenter/adaptive-immune-receptor-repertoires-webinar-series/	**2,812** (99.50%)
2. /404.html?page=/learningcenter/adaptive-immune-receptor-repertoires-webinar-series/?s=09&from=	**6** (0.21%)
3. /learningcenter/adaptive-immune-receptor-repertoires-webinar-series/?s=	**3** (0.11%)
4. /search?q=cache:GpnJ-h38rbQJ:https://www.antibodysociety.org/learningcenter/adaptive-immune-receptor-repertoires-webinar-series/+&cd=1&hl=es&ct=clnk&gl=mx	**3** (0.11%)
5. /learningcenter/adaptive-immune-receptor-repertoires-webinar-series/?_x_tr_sl=auto&_x_tr_tl=ja&_x_tr_hl=ja&_x_tr_pto=nui	**1** (0.04%)
6. /learningcenter/adaptive-immune-receptor-repertoires-webinar-series/?fbclid=IwAR1Ir09RW8jiik8mY_W0AUMLhXZX0-nSdb4NSngp7buhiRHfGTInw7f09fM	**1** (0.04%)

Rows 1 - 6 of 6

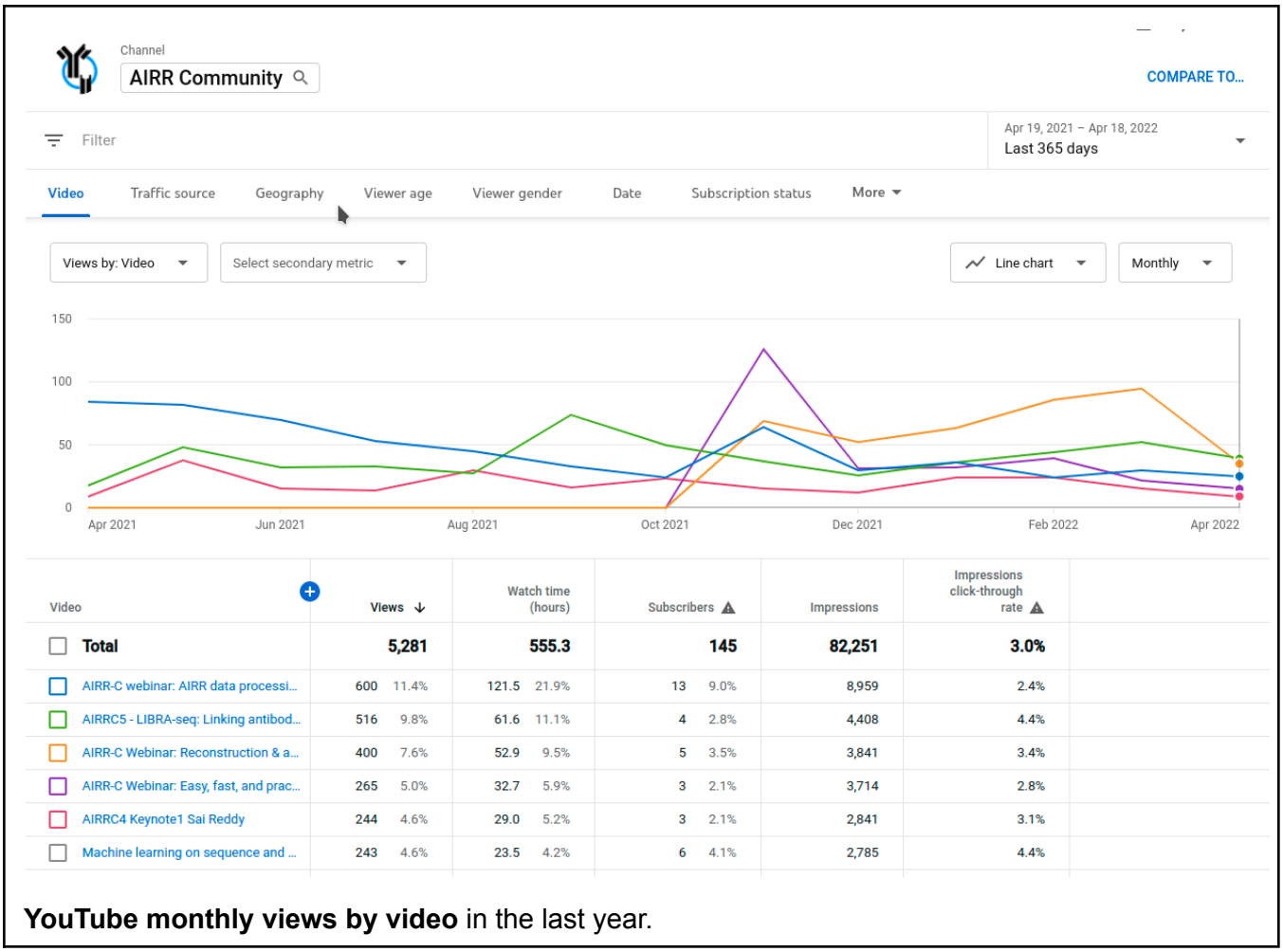
AIRR-C Webinar pageviews



LinkedIn unique visitors over the last 12 months period.



LinkedIn page views over the last 12 months period.



YouTube monthly views by video in the last year.

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

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

Current
Date of this report: April 20th 2022

SC/WG Name: Diagnostics Working Group

SC/WG Co-leaders: Rubelt, Gooley, Schwab

SC/WG Active Members (list):

Rohit Arora, Lmar Babrak, Justin Barton, Rachael Bashford-Rogers, Magnolia Bostick, Felix Breden, Syed Ahmad Chan Bukhari, Brian Corrie, Lindsay Cowell, Zhaoqing Ding, Sol Efroni, Khalil El Mazouari, Christopher Gooley (Co-lead), Victor Greiff, David Klatzmann, Yoshinobu Koguchi, Ton Langerak, Theam Soon Lim, Eline Luning Prak, Susanna Marquez, Pieter Meysman, Enkelejda Miho, Nima Nouri, Milena Pavlović, Florian Rubelt (Co-lead), Geir Kjetil Sandve, Tilman Schneider-Hohendorf, Nicholas Schwab (Co-lead), Erand Smakaj, Cinque Soto, Ulrik Stervbo, Johannes Trück, Henk-Jan van den Ham, Eric Waltari, and Corey Watson

Purpose:
To advance AIRR-seq for clinical use, i.e.in prognosis, diagnostics, and disease monitoring.The working group strives to uncover why there has been little translation of the tremendous progress in the field of AIRR-seq into the clinics. Furthermore, the working group aims to identify and promote ways of expanding AIRR-seq techniques (both sequencing technologies and analysis) into general clinical use. This we will do through:

- Evaluate the usage of AIRR-seq data for diagnostic purposes. These include diagnosis, prognosis, monitoring purposes, inclusion biomarkers for clinical studies, etc.
- Identify bottlenecks or challenges for AIRR-seq based clinical assays. These can be software tools, but also data sets or general approaches (statistical, machine learning, pattern recognition, etc).

Goals for the previous interval:

- Work with other interested WGs on a chapter in Immunogenetics about AIRR sequencing methodology from the standpoint of using them for diagnostics (our part) and help where we can on the other parts
- Establish a podcast-like interview series with experts in the AIRR sequencing diagnostic space to discuss their work and views

    - Work with Diagnostics WG members, Communications and Meetings SC to help source interviewees and for technical support
    - Work with The Antibody Society for further distribution
    - Potential topics:
        - Discuss published and future use cases for AIRR sequencing diagnostic testing starting with eg. MRD, COVID-19
        - Diagnostic methods and study design
        - Commercial and Research Use Only applications of diagnostic techniques
- Support the AIRR-C efforts with regard to the sequencing analysis of the immune response against the novel coronavirus SARS-CoV-2

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

- Act as an umbrella for clinic-oriented members of various AIRR-C working groups, such as the Biological Resources, Legal and Ethics and Software Working Groups, to brainstorm and promote further development and applications of AIRR-seq clinical testing (e.g. potential project design, standard data collection, software tools, extension to new diseases or uses)
- Represent a link between the AIRR-seq community and regulatory agencies when it comes to using AIRR-seq as a diagnostic tool

Products (if any):

Podcast:

- https://onairr.podbean.com/
- https://www.antibodysociety.org/the-airr-community/airr-c-podcast/

Paper:

- The Future of Blood Testing Is the Immunome' (Front. Immunol., 15 March 202)
- https://www.frontiersin.org/articles/10.3389/fimmu.2021.626793/full

Book chapter (contribution):

- Immunogenetics: Methods and Protocols, Methods in Molecular Biology, In press.

Resources (if any):

- Contribution to data sets shared in the AIRR data commons (COVID-19)

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

- Goals have been achieved

Proposed plans for the coming interval:

Purpose: Advance AIRRseq for clinical diagnostics & monitoring

Goals:

- Engage key opinion leaders to uncover barriers in translating progress in AIRR-seq into the clinic
- Broadly disseminate the existing and potential possibilities for AIRR-seq in the clinic
- Engage regulatory bodies (e.g. FDA) to create strategies to overcome obstacles in bringing AIRR-seq into the clinic

Products (if any):

- Produce and extend reach of additional episodes of the podcast “onAIRR” through support and and direct contribution
- Plan and hold a single meeting with diagnostics as a focal point, with support of the AIRR-C Communications SC, Legal and Ethics WG, and Meetings SC.
    - Proposal: e.g. satellite symposium during the IUIS 2023
- Coordinate a white paper response to regulatory bodies (how to include AIRR-seq in regulatory processes)

Resources (if any):

- OnAIRR - Immune receptors in the clinic

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

- Promote the vision described in our perspective paper ‘The Future of Blood Testing Is the Immunome’ (Front. Immunol., 15 March 2021 | https://doi.org/10.3389/fimmu.2021.626793) through community outreach and engagement
- Promote adoption of the standards established in the Standards Working Group for a common framework for clinical classifier discovery
- Act as an umbrella for clinic-oriented members of various AIRR-C working groups, such as the Biological Resources, Legal and Ethics and Software Working Groups, to brainstorm and promote further development and applications of AIRR-seq clinical testing (e.g. potential project design, standard data collection, software tools, extension to new diseases or uses)
- Represent a link between the AIRR-seq community and regulatory agencies regarding clinical application of AIRR-Seq

Proposed SC/WG Co-leaders:

- Ulrik Stervbo
- Susanna Marquez
- The three ex-co-leads as permanent resource and active members of the

working group, not as official co-leaders (only supposed to be 2 co-leads)

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

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

Current

Date of this report: May 17, 2022

SC/WG Name: Executive SC

SC/WG Co-leaders: Lindsay G. Cowell (chair)

SC/WG Active Members (list): Felix Breden, Christian Busse, Lindsay Cowell, Victor Greiff, Eline T. Luning Prak, Encarnita Mariotti-Ferrandiz

Purpose: The Executive SC provides leadership for the AIRR Community (**AIRR-C**)

Goals:

- Monthly meetings with occasional special meetings involving leadership of WGs and SCs
- Represent the AIRR-C on the TAbS Board of Directors and Finance & Audit Committee
- Facilitate and promote the work of the AIRR-C WGs and SCs through:
    - Manuscript endorsements
    - Encouraging adoption of AIRR-C standards and other initiatives with publishers and funding agencies
    - Writing letters of support for grant applications
    - Establishing an infrastructure grant for mission-critical functions of one or more AIRR-C WGs or SCs
- Outreach and mentoring for students and other members of the AIRR-C
    - Questionnaire to determine needs and interests of AIRR-C student members
    - Promoting the career development of AIRR-C members - letters of recommendation, networking

- Contribute to outreach efforts including commercial and academic partnerships that lead to greater sharing, annotation and standardization of AIRR-seq data
- Oversee WG/SC reorganizations, after ratification by the AIRR-C in Meeting V
- Update the AIRR-C mission statement, in part, to align with the revised TAbS mission statement.
- Develop a conflict-of-interest policy.
- Establish one or more AIRR-C prizes

Products (if any):

- Governance document
- Meeting minutes
- Manuscript endorsements

Resources (if any):

The AIRR-C Executive SC has purview over the finances of the AIRR Community, deciding how funds are allocated between different WG and SCs of the AIRR-C including funding for meetings, communications, administrative support, prizes and special projects

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

- Convened monthly meetings including 1-2 special meetings with SC and WG co-leads
- Represented on TAbS Board of Directors and TAbS Finance & Audit (budget and audit) Committee, with financial reporting.
- Reviewed all publications presented for AIRR endorsement (eLife review from the Biological Resources WG, Chapters for the Immunogenetics Volume of Methods in Molecular Biology organized by EuroClonality)
- Facilitated the first round of proposals for funding from TAbS to support mission critical WG/SC activities
- Oversee WG/SC reorganizations, after ratification by the AIRR-C in Meeting V
    - (merging of three WGs into the Standards WG)
- Update the AIRR-C mission statement, in part, to align with the revised TAbS mission statement.
    - TAbS ended up revising theirs, which is now more in line with ours
- Develop a conflict-of-interest policy.
    - Decided to work from the TabS one, which is still under development
- Establish one or more AIRR-C prizes
    - Established the Community Service Prize
- Obtained NIH funding to support travel awards to this meeting.
- Revised governance document:

- Clarifying edits:
    - Clarified the process around WG leadership transitions
    - Resolved a lack of clarity around eligible voters in AIRR-C votes, quorum for termination of merger with TAbS, and entering into force of approved proposals
- Edits to improve organization, readability, and reference:
    - Move WG descriptions to the appendix
    - Remove redundant/duplicated sections
    - Consistent formatting and numbering of paragraphs
    - Add logo licensing form, delegate licensing decisions to Comms SC

Proposed plans for the coming interval:

Purpose: The Executive SC provides leadership for the AIRR Community

Goals:

- Monthly meetings with occasional special meetings involving leadership of WGs and SCs
- Represent the AIRR-C on the TAbS Board of Directors and Finance & Audit Committee
- Facilitate and promote the work of the AIRR-C WGs and SCs through:
    - Manuscript endorsements
    - Encouraging adoption of AIRR-C standards and other initiatives with publishers and funding agencies
    - Writing letters of support for grant applications
    - Establishing an infrastructure grant for mission-critical functions of one or more AIRR-C WGs or SCs
- Outreach and mentoring for students and other members of the AIRR-C
    - Questionnaire to determine needs and interests of AIRR-C student members (comms Pam or Victor)
    - Promoting the career development of AIRR-C members - letters of recommendation, networking
- Contribute to outreach efforts including commercial and academic partnerships that lead to greater sharing, annotation and standardization of AIRR-seq data
- Oversee WG/SC reorganizations, after ratification by the AIRR-C in Meeting VI (establishment of group to develop a strategic plan)
- Develop a conflict-of-interest policy.

Products (if any):

Resources (if any):

Proposed SC/WG Co-leaders:

Victor Greiff is the Chair-elect and is the leader of the Executive SC.

List new members (if any) not in Active Member list above:

Other Executive SC members to be determined (election results pending at the time of writing this report)

AIRR-C Meeting VI - 2022
GLDB WG Report

Current

Date of this report: May 2022
SC/WG Name: Germline Database WG
SC/WG Co-leaders: Andrew Collins, Corey Watson
SC/WG Active Members (list):

Pierre Boudinot, Steve Bosinger, Felix Breden, Christian Busse, Scott Christley, Andrew Collins (Co-lead), Martin Corcoran, Chris Cottrell, Jamie Heather, Gunilla Hedestam, Katherine Jackson, Justin Kos, William Lees, Susana Magadan, Mats Ohlin, Ayelet Peres, Oscar Rodriguez, Cathrine Scheepers, Chaim Schramm, Jamie Scott, Amit Upadhyay, Henk-Jan van den Ham, Corey Watson (Co-lead), Gur Yaari and Jian Ye

Purpose:
This Working Group (WG) was formed to promote the development of complete and accurate sets of reference germline IG and TCR genes, and to promote the accurate analysis and reporting of the germline genes that can be identified in repertoire studies. The WG works to establish processes for documenting novel germline genes and alleles and standards for versioned, inclusive databases. The WG also provides guidance on specific topics relating to data assessment that may be referred to it by the Inferred Allele Review Sub-committee (IARC).

Goals:

- Develop and seek ratification of universal principles for nomenclature systems that meet the needs of IG and TCR researchers
- Work to develop systems to document and report variation in non-coding regions of the IG/TR loci
- Develop appropriate IG Reference Sets for mouse strains and for rhesus macaque
- Establish an IARC focused on the review of TCR gene inferences
- Develop and benchmark tools and systems to improve capacity to fully leverage data types of the future, while maintaining the quality of References Sets
- Continue to work on questions regarding database versioning, programmatic access and licensing of germline reference databases

Products (if any):

- Release of new germline gene sets on OGRDB (see link). Together these represent a significant increase in available V, D, and J germline gene sequences for the mouse:
    - IGH germline genes (based on AIRR-seq inference) for the BALB/cByJ inbred mouse strain. (see publication Jackson et al.)
    - IGH germline genes (based on AIRR-seq inference) for four inbred wild-derived mouse strains representing diverse sub-species origins (CAST/EiJ, LEWES/EiJ, MSM/MsJ and PWD/PhJ), as well as the inbred strain NOD/ShiltJ. (see publication Watson et al.)
    - IGL and IGK germline genes (based on AIRR-seq inference) for 18 inbred mouse strains, including disease models and wild-derived strains representing diverse sub-species origins. (see preprint Kos et al.)
- Contribution of BALB/cByJ IGH germline reference set to IgBLAST (now available: https://www.ncbi.nlm.nih.gov/igblast/)
- Integrated IG germline database for the Rhesus macaque. This database consolidates germline gene/allele sequences from multiple sources, including KIMDB, RhGLDB, and IMGT. Critically, this database includes an extensive set of germlines that more fully represent diversity in this species. Although many germline sequences in this database have not yet formally received formal IUIS names, the database leverages our new temporary label system, and concepts from our germline database schema (see below) to ensure that sequences can receive usable identifiers to improve communication between research groups.
- Tool and registry for the assignment of temporary gene/allele/sequence identifiers for IG and TR sequences that are characterised/published, but do not meet current requirements of review by IARC or the IUIS nomenclature committee. This tool can be found here: https://github.com/williamdlees/IgLabel
- Development of a “new” germline database schema. This schema is being developed to improve the documentation of germline IG and TR sequences through multiple levels of curation. Critically, the schema allows for flexibility/adaptability in the future development of germline databases that draw on more diverse supporting data types (e.g., AIRR-seq, germline DNA sequencing, and other genomic datasets).
- Links between OGRDB and VDJbase to make it easier to track novel alleles discovered in VDJbase and submit them to OGRDB
- Enhancements to VDJbase to support the MiAIRR metadata standard and to hold genomic data alongside AIRR-seq
- Group directed/affiliated manuscripts:
    - A paper describing the future of germline gene databases, outlining the development of a new database schema, and providing short- and

long-term views for leveraging data from multiple sources and building more adaptable systems for IG/TR germline gene/allele curation and nomenclature. This manuscript is currently under consideration for AIRR-C endorsement.
    - A paper describing a BALB/c IGHV Reference Set was published in Frontiers in Immunology (link). Although not an official AIRR-C publication, the paper was the work of the Mouse subgroup of the GLDB WG. The Reference Set is now integrated with IgBLAST.

Resources (if any):

- Expanded features and resources at GLDB-WG affiliated databases:
    - VDJbase: https://vdjbase.org/ (see products list above)
    - OGRDB: https://ogrdb.airr-community.org/ (see products list above)

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

- Develop and seek ratification of universal principles for nomenclature systems that meet the needs of IG and TCR researchers
    - Actions are underway to restructure governance of the the IG/TR/MH Nomenclature Sub-committee of IUIS's Nomenclature Committee. Until these issues are fully resolved, the GLDB-WG is reserving specific recommendations.
- Work to develop systems to document and report variation in non-coding regions of the IG/TR loci
    - There are now multiple tools available in the community to capture non-coding variation using either AIRR-seq data (e.g., 5'UTRs) and genomic data. Efforts to develop systems to catalogue and share such genetic variation have advanced considerably over the past interval. The focus has been on the development and implementation of tools/reporting features made available in VDJbase; this work has been driven by William Lees and the group of Gur Yaari (see vdjbase.org for more details). Specifically, VDJbase will soon offer users the ability to explore and analyse variant data available in both coding and non-coding regions of the IG and TCR gene regions.
- Develop appropriate IG Reference Sets for mouse strains and for Rhesus macaque
    - As outlined in the products section above, IG germline reference sets have been created for 18 mouse strains, and from a compilation of datasets for rhesus macaque, representing population level surveys of allelic variants. Mouse germline sets are currently available on OGRDB. Sources for complete germline sets for Rhesus will be shared in the coming weeks.

- Establish an IARC focused on the review of TCR gene inferences
    - This work has been transferred to the IARC; please see IARC 2022 report for updates on this goal.
- Develop and benchmark tools and systems to improve capacity to fully leverage data types of the future, while maintaining the quality of Reference Sets
    - This goal has been tackled at multiple levels. First, we have laid out a plan for a new germline database schema. This schema is designed to better facilitate the documentation of germline genes and alleles described for any species, including data sources, metadata, and identifiers. This schema is outlined in a manuscript currently undergoing AIRR-C endorsement procedures. As means to address this goal, the schema is structured to provide flexibility at multiple levels: 1) it more easily allows for the documenting of germline gene/alleles coming from a variety of data sources (e.g., AIRR-seq, genomic data, future data types.); 2) it allows for more seamless linking of sequence records and identifiers, promoting transparency and provenance.
    - Second, the reliance on genomic data for the application of existing germline gene/allele nomenclatures has limited our ability to efficiently name and share sequence sets curated from data that don't align with the "gold-standard" data types. For example, there is a growing collection of germline genes and alleles that have been discovered using non-traditional approaches (e.g., AIRR-seq), particularly in non-human species. In these cases, these germlines can currently neither be reviewed by IUIS nor IARC. And while, in the short-term these sets represent high-value data for the community, the effective sharing of these data between research groups is stilted by our inability to apply stable naming systems. Thus we developed an approach that will allow research groups to contribute to growing germline sets, and assign temporary gene/allele labels to newly discovered sequences in any species of interest. These identifiers are being managed by a registry system set up by the GLDB, in which interested participants can employ the registry to get stable unique gene/allele identifiers assigned to their germline sequences. This allows us to circumvent IUIS/IARC for the cataloguing of sequences for short-term utility, if and until such sequences can obtain formal names when genomic evidence is made available. Critically, these names can be utilised in conjunction with formal IUIS names within the database schema mentioned above.
    - Third, several members of the GLDB have been working on method development for curating IG/TR genes from high-throughput genomic data. While complete systems have not been developed yet for the review of such data, we expect these efforts to provide foundational resources for developing such capabilities moving forward (e.g., see progress at VDJbase noted above).

- Continue to work on questions regarding database versioning, programmatic access and licensing of germline reference databases

    - All three topics have seen some progress in the past Interval:
        - Database versioning: The new germline database schema discussed above foresees mechanisms that facilitate fine-grained and transparent updates of germline reference databases.
        - Programmatic access: Germline sets on OGRDB are accessible via a REST API. A defined AIRR Data Commons API will be added in the coming period.
        - Database licensing: In line with previous recommendations of Legal & Ethics WG, all data on ORGDB and VDJbase is licensed under a CC0 licence, allowing for completely unrestricted use and reuse. In addition, GLDB Participants have contributed various use cases to the upcoming detailed analysis of the EU Database Directive (96/9/EG), which has been conducted by Legal & Ethics WG.

Proposed plans for the coming interval:

Purpose:

This Working Group (WG) was formed to promote the development of complete and accurate sets of reference germline IG and TCR genes, and to promote the accurate analysis and reporting of the germline genes that can be identified in repertoire studies. The WG works to establish processes for documenting novel germline genes and alleles and standards for versioned, inclusive databases. The WG also provides guidance on specific topics relating to data assessment that may be referred to it by the Inferred Allele Review Sub-committee (IARC).

Goals:

- Continue to develop appropriate IG Reference Sets for human, various mouse strains, and for rhesus macaque, building upon recent efforts by integrating genomic data currently being generated by multiple groups. This integrated effort should result in robust germline sets for these species that could serve as useful models for other non-human organisms.
- Continue the development and implementation of the germline database schema.
- Collaborate with VDJbase to further develop features for cataloguing and sharing non-coding variation.

- Create an outreach subgroup to identify academic and commercial partners involved in the generation and use of germline data, as a means to encourage broader inclusion of key stakeholders. Specifically, we want to:
    - Improve the provision of datasets from additional species
    - Identify stakeholders and integrate them into community efforts
    - Promote the uptake of germline sets and schema in tools
    - Develop a plan for sustainability through ongoing funding

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

Enhance the accuracy and species coverage of AIRR-seq by providing comprehensive and regularly updated germline reference sets for species of interest. Leverage next-generation techniques (e.g. inference from AIRR-seq and long-read genomic sequencing) to make this possible. Provide this on a sustainable basis with respect to resourcing and funding.

Build understanding and awareness of the importance of comprehensive reference sets in AIRR-seq analysis. Promote the application of 'personalised' genotypes and haplotypes within AIRR-seq analysis, demonstrating the value they can add.

Develop and publish best practice in the application of current and future tools and methods for germline gene discovery.

Proposed SC/WG Co-leaders:

TBD

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

Current
Date of this report: April 20, 2022

SC/WG Name: Inferred Allele Review Committee (IARC)

SC/WG Co-leaders: Mats Ohlin

SC/WG Active Members (list): Ayelet Peres, Gur Yaari, Martin Corcoran, Andrew Collins, Corey Watson, William Lees, Mats Ohlin, James Heather (co-opted)

Purpose: to review human immunoglobulin germline gene inferences from AIRR-Seq datasets, and to communicate affirmed sequences to the IG gene naming authorities.

Goals:

- Continue to evaluate human IGH, IGK, and IGL inferences
- Establish a human TCR IARC
- Establish IARC(s) for non-human species
- Develop tools and guidelines for the evaluation of the 'quality' of submitted datasets
- Streamline procedures for the submission and assessment of human IGHV inferences
- Develop strategies to make the work of the IARC(s) sustainable

Long-term vision and how WG products integrate with the AIRR-C mission:
IARC provides resources and knowhow that enable correct annotation and analysis of IG and TR transcriptomic data, in agreement with the AIRR Community's mission to standardize, analyze and share AIRR-seq data.

Products (if any):
- 9 inferred immunoglobulin alleles submitted to and approved by IUIS
- A published commentary on the use of short read sequencing for AIRR allele discovery
- A submitted manuscript on Balb/c germline alleles identified by inference (co-authored with other WGs)
- Germline gene sets for Balb/c and C57BL/6 mouse strains published on OGRDB

Resources (if any):
- OGRDB with enhanced functionalities
    - Gene set feature being added
    - Better support for submission
    - Improved OGRDBStats plot functionality.
- A Functional Groups Reference Book app to support IARC's assessment of inferred germline genes

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

Since AIRR Community Meeting V, IARC has contributed the following activities/outputs.

- IARC has, as of April 19, 2022, met 36 times since AIRR Community Meeting V.
- Continue to evaluate human IGH IGK, and IGL inferences
    - 9 inferences of IGHV alleles have been affirmed, reported to IUIS and included in the IMGT reference set
    - A Functional Groups Reference Book App is being developed (Ayelet Peres and Gur Yaari) to aid future assessment of novel, inferred alleles by IARC.
    - We have initiated the integration of inferred and genomic data for enhanced understanding of AIRRs.
- Establish a human TCR IARC
    - The process to establish a TCR IARC has been initiated. James Heather has been co-opted to IARC with the intention to bring additional TCR expertise into the team and to transfer expertise of inference, affirmation, and process
    - We have initiated the assessment of novel inferred alleles of the TRBV locus.
- Establish IARC(s) for non-human species
    - Alleles of germline genes of mouse and macaque are currently mostly dealt with within the existing IARC as several members are experts on these germline gene repertoires.
    - Members of IARC co-authored a manuscript on a mouse Balb/c IGHV gene set
    - Balb/c and C57BL/6 expressed germline gene sets have been published on OGRDB.
- Develop tools and guidelines for the evaluation of the ‘quality’ of submitted datasets
    - In process in collaboration with Software WG.
    - New functionalities within OGRDB (enhanced plots, germline sets etc.) and Functional Groups Reference Book will aid submission and assessment of novel alleles.
    - IARC authored a commentary on the use of short read sequence data for AIRR allele discovery (doi: 10.1038/s41435-021-00152-6).
    - Members of IARC co-authored manuscripts on sequencing and annotation.
    - Members of IARC co-developed a temporary allele label system (manuscript in progress).
- Streamline procedures for the submission and assessment of human IGHV inferences
    - OGRDB has been integrated with VDJbase to simplify submission of VDJbase inferred alleles.
    - Tooling to create record sets for NIH/ENA submission has been improved, and documentation explaining the submission process has been rewritten.
    - A submission process with ENA has been established.
- Develop strategies to make the work of the IARC(s) sustainable
    - An application for funding of infrastructure has been submitted (decision pending).

Proposed plans for the coming interval:
Purpose: To generate resources for and advocate for better practice in and understanding of IG and TR annotation and analysis in the community interested in studies of AIRR.

Goals:

- To affirm additional novel human IGHV, IGKV, IGLV, and TRV alleles, and extend partial genes.
- To in collaboration with other WG identify expressed AIRR alleles and publish human IGV germline allele sets
- To, by providing expertise to other WGs/IARCs, consolidate a macaque IG germline set.
- To, in collaboration with other WGs, consolidate the use of temporary labels in instances where formal gene assignment of alleles cannot be made with confidence.
- To further integrate genomic and transcriptomic data of genes and alleles of IG loci.
- To, in collaboration with other WGs, establish processes for allele submission and germline gene sets of other species to OGRDB.
- To submit manuscript(s) on novel alleles and IARC processes.
- To publish guidelines for generation and growth of germline sets to establish formally recognized gene sets.
- To establish funding mechanisms moving forward.
- To consolidate human TCR IARC process.
- To develop IARC sustainability.

Products (if any):

- Affirmed alleles of AIRR germline genes, submitted to IUIS for final approval.
- Published manuscripts of relevance to the IARC's responsibilities.
- Germline gene sets (human and macaque).

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

IARC provides resources and knowhow that enable correct annotation and analysis of IG and TR transcriptomic data, in agreement with the AIRR Community's mission to standardize, analyze and share AIRR-seq data

Proposed SC/WG Co-leaders: Mats Ohlin

AIRR-C Legal & Ethics Working Group Reporting - 5th Interval - 2021-2022

Date of this report: Wed, 2022-04-20

SC/WG Name: Legal & Ethics WG

SC/WG Co-leaders: Tania Bubela, Christian Busse

SC/WG Active Members (list): Alexander Bernier, Tania Bubela, Christian Busse, Dan Emerling, Susanna Marquez, Artur Rocha, Jamie Scott, Jacob Sherkow

Purpose: To address issues pertaining to legal (e.g., IP, GDPR) and/or ethical (e.g., human subjects) standards for the AIRR Community.

Goals:
(after re-scoping in spring 2021)

- Understand the restrictions on data sharing currently imposed by European database protection laws (i.e., Directive 96/9/EC), especially in the context of reference sequence sets (in cooperation with GLDB WG).

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

The Legal & Ethics WG aims to facilitate sharing and reuse of AIRR-seq and associated data types by providing insights into the relevant legal and ethical frameworks. The WG's products are considered individual steps towards this goal.

Products:

- Draft manuscript on Database Directive and its implications for biological data

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

Following the rescoping in May/June 2021, the WG started to meet on a monthly basis, focusing around the topic of the impact of the EU Database Directive. These efforts were led by Alex, who is also the author of the (full-length) manuscript on this topic. Additional topics that were discussed are questions around medical devices. Finally, WG also provided input regarding Database Directive to GLDB WG and to the Comms SC regarding licensing of the AIRR-C logo.

Proposed plans for the coming interval:

Goals:

- Finalize and publish Database Directive manuscript and summary
- Plan, conduct and evaluate a survey regarding obstacles to data sharing
- Provide guidance on GDPR issues.
- Develop an ethical framework for patient engagement and studies involving indigenous people.
- Evaluate legal and ethical aspects of current Open Science practices and recommendations in collaboration with other stakeholders (e.g., RDA, IUIS).
- Support the Diagnostics WG and Meetings SC in organizing an event covering AIRR-seq, AI/machine learning and legal and regulatory aspects.

Products:

- Manuscript on Database Directive and its implications for biological data (audience: legal experts)
- Summary on Database Directive manuscript (audience: researchers, data curators, data stewards)

Proposed SC/WG Co-leaders: NN, NN

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

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

Current
Date of this report: April 22, 2022

SC/WG Name: Meetings Sub-committee

SC/WG Co-leaders: Encarnita Mariotti-Ferrandiz and Pam Borghardt (also Meeting Manager)

SC/WG Active Members (list):
Davide Bagnara, Justin Barton, Jean-Philippe Bürckert (resigned during the term), Ramit Mehr (resigned during the term), Corey Watson. Guest member: Jamie Scott

Purpose:
The AIRR-C Meetings Sub-committee is responsible for the initiation and planning of AIRR-C meetings.

Goals:
Plans for 2021/2022 included:

- re-Planning the AIRR Community Meeting VI to be held at Scripps Research and the La Jolla Hilton Hotel in La Jolla, California
- Meeting title: AIRR Community VI: "Exploring New Frontiers"
- Challenge Session: Engaging/integrating with other types of "Big Immunological Data" (Systems Immunology)
- Challenge Theme: Sustaining AIRR-C initiatives in the future
    - Potential topics include:
        - Improving and promoting MiAIRR standards in the publication
        - "Anchoring" AIRR-C repositories into our institutions
        - Engaging government agencies for new funding mechanisms to build and support "big data infrastructure" (*e.g.*, IMGT)

Products (if any):
NIH grant for travel awards successful application.

8 June 2021: FOCIS 2021 Symposia Application - "**How immune repertoires can inform human immunology"**
https://www.youtube.com/channel/UCtG0yZKjFuiFbYF5tfrAMdA/videos

https://docs.google.com/document/d/1FTTh93RCMhnqFi_JrN07LsUTTt8eROskGwvaqbBCbkE/edit

17-19 May 2022: AiRR Community Meeting VI: “Exploring New Frontiers”
https://www.antibodysociety.org/the-airr-community/meetings/airr-community-meeting-v-zooming-in-to-the-airr-community/

A lot of meetings since January 2021 (>50)
Meeting Minutes

Resources (if any):
Needed resources: Funding for managerial support, support for on-lline services (Zoom, meeting platform), funding from travel awards.
Resources provided: Funds raised through registration fees and corporate sponsorships

Progress report on current purpose, goals, products and resources:
COVID-19 caused a lot of upheaval and uncertainty in our work.
We spent part of the first 9 months of 2021 trying to maintain the initial attempt for the AIRR-C VI meeting to be held in december 2021 at La Jolla. The date motivation was mainly for our annual meeting to be held after the Antibody and Engineering Society meeting.
The progression of the pandemic by waves led the SC to conduct a survey within the AIRR-C members to evaluate the intentions to participate. The results were clear, most of the European people were not allowed to travel to the US (unless a critical obligation) and many US based AIRR-C members were not in favor or allowed to travel either.
In october 2021, we decided to reschedule the meeting in May 2022 after renegotiating once more our contract with the La Jolla Torrey Pines Hilton Hotel. We made the decision to recommend to Executive Sub-Committee that we postpone the in-person meeting to May 2022.

Early 2021, following Jamie’s suggestion, we applied to a FOCIS 2021 symposium session (virtual), as a society member of the FOCIS thanks to our affiliation to the TAbS, and with the help of Janice Reichert. Our successful application led to the planning of a 4h meeting, including 7 AIRR-C members as speakers.

FOCIS 2021 1st AIRR-C Symposium: “How immune repertoires can inform human immunology” , June 8th 2021 (Virtual)

The symposium aimed at providing basic and advanced knowledge to the FOCIS community and meeting attendees, not necessarily experts in the AIRR field but definitely with growing interest in it. We focused on providing learnings on:

- how are AIRR-seq data generated, and how they inform the biology of human immune responses.
- what to do, and not do, when studying AIRRs (wet & dry lab aspects), as well as remaining challenges.
- new tools for AIRR-seq data modeling, sharing, and storage.
- history, contributions, and initiatives of The AIRR Community (AIRR-C).

● ease to reach out the community for AIRR non-expert researchers to find partners and collaborators in this field.

The symposium gathered 89 live viewers, and further attracted 243 on-demand views. The recordings have been recovered and are progressively loaded on the AIRR-C youtube channel. The viewing and spreading continues..

AIRR Community Meeting VI: "Exploring New Frontiers"
Three days in-person sessions including the recurrent AIRR-C Working Group & Sub-committee presentations, scientific basic and biomedical sessions, poster sessions, short talks, software tool demonstrations.
For this 6th meeting, we added several new activities:

- two challenge sessions: systems immunology with invited talks and a session on Encouraging adoption and sustainability of the AIRR-C.
- A mentoring session and activity during the meeting on a voluntary based application for mentors and mentees. 14 mentors and 25 mentees applied.
- An industry meeting networking session organized as round tables to be moderated by AIRR-C WG leaders/active members and industry attendees.

The meeting has been organized with a virtual component: all the talks, tutorial and live demos will be live streamed.

Proposed plans for the coming interval:
Purpose: The AIRR-C Meetings Sub-committee is responsible for the initiation and planning of AIRR-C related meetings.

Goals:
Plans for 2022-23 include:

- Planning AIRR-C Meeting VII, to be held in spring in Europe
    - **Theme:** AIRR Community Meeting VII: TBD
    - **Challenge Theme**: TB
    - Industrial/sponsor networking events
    - Software Tool Demonstrations
    - Poster presentations
    - Workshops

Possible locations : Norway (grant application on going), Genoa (because already done and quite successful), Spain. The final decision will be made with the renewal of the meeting SC.

- Work with the AIRR Community members to host *ad hoc* meeting opportunities as they arise

- Host an AIRR-C satellite symposium at the annual FOCIS meeting in June 2023

- In collaboration with the Executive Sub-committee, potentially host an AIRR-C forum on values regarding conflict of interest, open science, FAIR data principles, diversity, equity

and inclusivity. Values statements on these topics would be developed along with their implementation within the AIRR-C.

Products (if any):
Spring 2023, AIRR-C Meeting VII
June 2023, an AIRR-C satellite symposium at the annual FOCIS meeting (potentially; needs discussion and approval)

Resources (if any):
Funding will be needed for managerial support, on-line meeting sites (e.g., Zoom & Whova) (if necessary), or for holding a face-to-face meeting for AIRR-C Meeting VII.

We expect to raise funding from registration fees and meeting sponsorship.

Proposed SC/WG Co-leaders:

Pam Borghardt and Encarnita Mariotti-Ferrandiz

Members continuing: Corey Watson and Justin Barton

This past year, several members of the Meeting Sub-committee had to resign for personal or professional reasons. We started 6 members in 2021 and finished 3 members. For the next upcoming interval, and given the planning of the meeting in Europe as well as the planning of reasonably sound number of virtual meetings, the Sub-committee calls for new members to join and reach back the necessary size of 6 members, including european based ones.

Meetings are typically bi-weekly, and ratchet up to weekly as an upcoming meeting approaches. Given the uncertainty with the pandemics but also global warming that changed the way people accept to travel, we want to maintain the virtual platforms. That means we need at least two members willing to be in charge of the virtual platform.

Altogether, we are looking for 3 willing and active members interested in helping with the planning of networking events, poster sessions, software tool demonstrations, *etc*.

List new members (if any) who are not in the Active Member list above:

We need 3 new members!! Some of our best AIRR-C members are students, postdocs! Please step up and join us in this important job!

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

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

Current
Date of this report: April 5, 2022

SC/WG Name: Software

SC/WG Co-leaders: William Lees and Chaim Schramm

SC/WG Active Members (list):
William Lees, Chaim Schramm, Bryan Briney, Brian Corrie, Susanna Marquez, Luke Myers, Mats Ohlin, Adrian Shepherd, Wenming Xiao

Purpose: To make it easy to do rigorous analysis of AIRR-seq data.

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

The purpose of our working group is to encourage practices that enable software tools to work, and to work with one another.: this dovetails with the AIRR-C mission to improve methods and standards. We aim to promote awareness of AIRR standards within the software tools community, and to provide a forum through which software tools developers can discuss and progress topics of mutual interest.

Current Period

Goals:

- Complete a benchmarking framework for comparing annotation tools
- Identify datasets for testing annotation tools
- Build a tool for assessing repertoire credibility (RepCred)

A key first priority is to assemble data sets that people can use to test and compare the functionality of various programs. To progress this, we have defined summary statistics that can be used to characterize simulated data sets and compare them to real-world data sets. The result is a software tool, Sumrep, which is described in a recent paper. We will use this tool to assist in the collection of simulated and real-world data sets for testing and benchmarking.

We have also defined a standard for AIRR-Seq software tools. We continue to promote this as a way of encouraging inter-operation and adoption of AIRR standard protocols by providing community support and publicity to compliant tools. We have built the core of a tool to assess the biological credibility of an AIRR-Seq repertoire, and to identify common technical errors

that can occur during its preparation, which can be heard to spot from read quality annotations and other technical measures commonly available today.

Products (if any):

- Benchmarking framework
- RepCred tool

Resources (if any):

- Gold standard datasets

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

- Benchmarking framework was derailed by loss of collaborators
- RepCred tool is nearly complete

Proposed plans for the coming interval:

Purpose and long-term vision as above.

Goals:

- Complete RepCred
- Align benchmarking and gold-standards dataset with PrecisionFDA activities
- Survey our AIRR-C membership for key focus areas (particular emphasis on industry needs)
- Support the Communications WG industry initiative by drawing on WG members' knowledge of the tools marketplace
-

Products (if any):

- **RepCred tool**

Resources (if any):

- **Gold standard datasets**

Proposed SC/WG Co-leaders: William Lees and Chaim Schramm

List new members (if any) not in Active Member list above:

AIRR-C Standards Working Group Reporting for 2021-2022

Date of this report: April 20, 2022

SC/WG Name: Standards

SC/WG Co-leaders: Christian Busse, Jason Vander Heiden, Scott Christley

SC/WG Active Members: Felix Breden, Christian Busse, Brian Corrie, Veronique Giudicelli, Eli Harkins, Kenneth Hoehn, Susanna Marquez, Chaim Schramm, Scott Christley, Jason Vander Heiden, Ulrik Stervbo, William Lees, Kira Neller, Aditi Jain, Jingyun Li, Adrien Six, Artur Roca, Edward Lee, Marco Oliveira, Bjorn Peters, Francisco Arcila, Florian Rubelt, Katharina Imkeller, Lindsay Cowell, Nina Luning Prak, Nicole Knoetze, Enkelejda Miho

Purpose:

Develop a set of metadata standards (MiAIRR) for the submission of adaptive immune receptor repertoire sequencing (AIRR-seq) datasets. Develop standardized file formats, schemas and data field names to represent MiAIRR metadata, annotated antibody and T cell receptor sequences, and any downstream data representations. These standards are defined in formal machine-readable specifications, allowing interoperability between software from different developers.

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

The Standards WG aims to facilitate data sharing and interoperability of analysis tools within the AIRR-seq field through common data and metadata standards and documentation.

Products:

- Machine-readable, open source schema for AIRR-seq data.
    - https://github.com/airr-community/airr-standards
- Reference API libraries in R and python providing read, write and validation operations for finalized schema.
    - https://pypi.org/project/airr
    - https://cran.r-project.org/web/packages/airr
- Detailed schema and software documentation for Standards WG products and those of other WGs, along with documentation resources for public data submission and compliant community tools.
    - https://docs.airr-community.org

Progress in 2021-2022:

- Merged the Minimal Standards WG and Data Representations WG into a single Standards WG.
- Experimental germline database schema finalized, in collaboration with GLDB.

Included provisional support in the R and python libraries.
- Experimental single-cell schemas finalized.
- Experimental clonal lineage schemas finalized.
- Experimental receptor schema development is ongoing.
- Review and harmonization of AIRR terminology documents.
- Rough draft schemas for both a file manifest and aggregation of multiple repertoires.
- Various process improvements on GitHub, concerning unit tests, meeting minutes, and project management.
- Release of AIRR Standards v1.3.1 and associated python and R libraries.

Proposed plans for 2022-2023:

- The next cycle will focus primarily on refinement of experimental schemas for release in production ready versions along with a manuscript.
- Release AIRR Standards v1.4, which is scheduled to include:
    - Experimental release of the germline database schemas.
    - Experimental release of the single-cell schemas.
    - Experimental release of the receptor schema.
    - Updates to abundance fields to account for new technologies.
    - Support for additional schemas in the R and python libraries.
    - Abandonment of Python v2 support.
    - Various minor improvements to field definitions and documentation.
- Release AIRR Standards v2.0, which is scheduled to include:
    - Production release of the germline database schemas.
    - Production release of the single-cell schemas.
    - Production release of receptor schemas.
    - Production release of the lineage schemas.
    - Experimental release of a file manifest schema, repertoire grouping schema, and a persistent identifier definition.
    - Several small, but backwards incompatible changes.
- Draft a manuscript to accompany the v2.0 release describing new standards development since the original Minimal Standards (https://doi.org/10.1038/ni.3873) and Data Representations (https://doi.org/10.3389/fimmu.2018.02206) publications in 2017 and 2018, respectively.

Proposed SC/WG Co-leaders for 2022-2023: Christian Busse, Jason Vander Heiden

AIRR Community Meeting VI Attendees

Attendee	Affiliation	City, Country
Ademar Aguiar	INESC TEC	Porto, Porto, Portugal
Arvin Akoopie	AlivaMab Discovery Services	United States
Ray Alvarez	Icahn School of Medicine at Mt. Sinai Hospital	New York, New York, United States
Lynnette Ang	Omniscope	Barcelona
Edel Aron	Yale University	United States
Hossein Asgharian	Roche	San Jose, California
Davide Bagnara	University of Genoa	Genoa, Liguria, Italy
Daria Balashova	Amsterdam UMC	Amsterdam, North Holland, Netherlands
Rituparna Banerjee	University of British Columbia, Vancouver	Greater Vancouver A, British Columbia, Canada
Justin Barton	Alchemab	Cambridge, UK
Habib Bashour	University of Warwick - School of Life Sciences	Leamington Spa, England, United Kingdom
John Beckford	Takara Bio	San Jose, California
Bryan Bell	Takara Bio	Seattle, WA
Jannick Bendtsen	Pipe Bio	
John Bergen	Amgen	Burnaby, British Columbia, Canada
Jan Berka	Roche	Czech Republic
Jody Berry		
Spela Binter	Kymab Sanofi	Cambridge, England
George Blanck	USF	Tampa, Florida, United States
Owen Bodley	PipeBio	Århus, Middle Jutland, Denmark
Robin Bombardi	Illumina	Nashville, Tennessee, United States
Stefano Bonissone	Abterra Biosciences	San Diego, California
Nicola Bonzanni	ENPICOM BV	Utrecht, Utrecht
Pam Borghardt	Simon Fraser University	Greater Vancouver Metropolitan Area
Steven Bosinger	Emory University	Atlanta, Georgia
Cecilie Boysen	DNAnexus	
Rebecca Brachman	Cornell Tech	New York, New York, United States
Maarten Braspenning	BISC Global	Ghent Metropolitan Area
Felix Breden	iReceptor	Pasadena
Bryan Briney		

Alex Brown	National Jewish Health	Denver, Colorado
Jean-Philippe Burckert	Distributed Bio Inc	San Francisco Bay Area
Dennis Burton	Scripps Research	San Diego County, California, United States
Christian Busse	German Cancer Research Center (DKFZ)	Berlin
Atul Butte	University of California, San Francisco (USA)	San Francisco, California, United States
Natalie Castellana	Abterra Biosciences	San Diego, California, United States
Amit Chaudhuri		Redwood City, California, United States
Pei-Lung Chen	Graduate Institute of Medical Genomics and Proteomics, National Taiwan University	Taipei
Maria Chernigovskaya	University of Oslo	Oslo, Oslo, Norway
Mark Chernyshev	Karolinska Institutet	Sweden
Leila Chihab	La Jolla Institute for Immunology	
Geummi Cho		
Scott Christley	UT Southwestern Medical Center	Dallas, Texas, United States
Sarah Cobey	University of Chicago	Greater Chicago Area
Andrew Collins	University of New South Wales	Greater Sydney Area
Brian Corrie	Simon Fraser University	Summerland, British Columbia
Christopher Cottrell	Scripps Research	
Lindsay Cowell	UT Southwestern Medical Center	Dallas, Texas, United States
Shane Crotty	La Jolla Institute for Immunology (USA)	San Diego County, California, United States
Simon van Dam	Agora/Interteam	Israel
Dalia Daujotyte	Illumina	San Diego, California
Marie-helene Delfau-larue	APHP	Greater Paris Metropolitan Region
Hannah van Dijk	Deutsches Krebsforschungszentrum	Berlin, Berlin, Germany
Zhaoqing Ding	Gossamer Bio	San Diego, California, United States
Artemis Efstratiou	Max Planck Institute for Evolutionary Biology	Plön, Schleswig-Holstein, Germany
Eva Eilers	University Clinic Münster	
Ali Ellebedy	Washington Univ. School of Medicine, St. Louis (USA)	
Daniel Emerling	Atreca, Inc.	El Cerrito, California, United States
Anne Eugster	CRTD TU Dresden	Greater Dresden Area
Andrew Farmer	Takara Bio	
Ingeborg Feil	Janssen R&D	San Diego, California
Lamar Fleming	Fred Hutchinson Cancer Center	Seattle, Washington
Easton Ford	University of Louisville	Louisville, Kentucky, United States

Romain Fournials	Omniscope	
Romain Fournials	Omniscope	
Robert Frank	University of Oslo	Oslo, Oslo, Norway
Genta Furuya	The University of Tokyo	Tokyo
Jake Galson	Alchemab Therapeutics	Greater Cambridge Area
Khadija Ghias	Cellecta	
William Gibson	University of Louisville	Louisville, Kentucky
Sofie Gielis	UAntwerpen	Antwerp Metropolitan Area
Christopher Gooley	Microsoft Research	Redmond, Washington, United States
Kenz Le Gouge	i3 lab, Sorbonne Université	Paris, Île-de-France, France
Victor Greiff	University of Oslo	Oslo, Oslo, Norway
Henk-Jan van den Ham	ENPICOM BV	's-Hertogenbosch, North Brabant, Netherlands
David Hamm	Adaptive Biotechnologies	Seattle
David Hamm	Adaptive Biotechnologies	
Jennifer Hammond	Illumina	San Diego, California, United States
Alena Harley	Tempus	USA
Peter Hawkins	Regeneron Pharmaceuticals	New York City, New York
James Heath	James Heath, Institute for Systems Biology, Seattle (USA)	
Jamie Heather	MGH	Greater Boston
Gunilla Karlsson Hedestam	Karolinska Institutet	
Gurjinder Heer	GlaxoSmithKline	
Jason Vander Heiden	Genentech	Burlingame, California, United States
Jonathan Herman	Jonathan Herman, Massachusetts General Hospital (USA)	Boston, Massachusetts, United States
Holger Heyn	Omniscope	Barcelona, Catalonia, Spain
Laurie Hodge	Adaptive Biotechnologies	Greater Seattle Area
Kenneth Hoehn	Yale School of Medicine	United States
Juliet Neun Hornick	Simon Fraser University	Vancouver, British Columbia, Canada
Xihao Hu	GV20 Therapeutics	Boston, Massachusetts, United States
Yu-Ning Huang	USC School of Pharmacy	Los Angeles, California
Jonathan Hurtado	The Scripps Research Institute - La Jolla Campus	San Diego, California
Katherine Jackson	Garvan Institute of Medical Research	Sydney, New South Wales
Aditi Jain	DKFZ German Cancer Research Center	Saarbrücken, Saarland, Germany
Uddalok Jana	University of Louisville	Louisville, Kentucky
Mahita Jarjapu	La Jolla Institute for Immunology, Division of Vaccine Discovery	San Diego Metropolitan Area
Cole Jensen	Yale University	New Haven, Connecticut

Collin Joyce	Scripps Research	San Diego County, California, United States
Kendall Kearns	La Jolla Institute for Immunology	
Dominic Kelly	University of Oxford	
Brennan Abanades Kenyon	University of Oxford	Oxford, England, United Kingdom
Steven Kleinstein	Yale University (USA)	New Haven, Connecticut, United States
Pieter Martijn Kolijn	Erasmus MC	Rotterdam
Justin Kos	University of Louisville School of Medicine	Louisville Metropolitan Area
Evan Kransdorf		
Konrad Krawczyk	NaturalAntibody	Hamburg, Hamburg, Germany
Angelica LaVallee	MedGenome	San Diego, California, United States
Mansun Law	The Scripps Research Institute	
Donald Lee	Eli Lilly and Company	Frederick, Maryland, United States
Edward Lee	Yale School of Medicine	Palo Alto, California, United States
Jeong Hyun Lee		
Noah Lee	Yale University	New Haven
Jinwoo Leem	Alchemab Therapeutics Ltd	
William Lees	Birkbeck College University of London	London, England
Jingyun Li	DKFZ	
Ida Lindeman	Oslo University Hospital	Oslo
Hong-Hsing Liu	Institute of Molecular and Genomic Medicine, National Health Research Institutes	Taiwan
Diana Lok	Illumina	San Mateo, California, United States
Svitlana Lozova	ENPICOM BV	
Ana Lujan	Twist Bioscience	
Jos Lunenberg	ENPICOM BV	‘s-Hertogenbosch
Jiaqi Luo	0	Hong Kong
Pamela Lussier	Illumina	Oroville, Washington, United States
Thijs Maas	ENPICOM	's-Hertogenbosch, North Brabant
Thomas Maccarthy	Stony Brook University	
Susana Magadan	University of Vigo and University of New Mexico (USA)	Greater Vigo Metropolitan Area
Mikhail Makhanov	Cellecta	
Serghei Mangul	USC	Los Angeles, California
Encarnita Mariotti-Ferrandiz	Sorbonne Université	

Susanna Marquez	Yale University	Valencia, Valencia
Alaine Marsden	Antibody Immunity Research Unit, NICD	
Maria Rodriguez Martinez	IBM Research	Zurich, Switzerland
Ramit Mehr	Bar-Ilan University	Petah Tiqwa, Central, Israel
Jens Meiler	Vanderbilt University (USA) and Leipzig University (Germany)	Nashville, Tennessee, United States
Marcus De Almeida Mendes	La Jolla Institute for Immunology	San Diego, California
Ruth Mercado	Cellecta	Mountain View, California
Shoichi Metsugi	Chugai Pharmacetutical Co., Ltd	Kamakura, Kanagawa
Ahmed Metwally	Illumina	San Francisco Bay Area
Jan Michler	ETH Zurich	Germany
Gael Millot	Institut Pasteur	
Anastasia Minervina	St. Jude Children's Research Hospital (USA)	
Hamid Mirebrahim	Roche	Pleasanton, California, United States
Rohini Mopuri	Emory University	Cumming, Georgia
Rajagopal Ayyanar Murugan	DKFZ German Cancer Research Center	Heidelberg, Baden-Württemberg, Germany
Luke Myers	Vanderbilt Vaccine Center	Nashville, Tennessee, United States
Kira Neller	iReceptor/SFU	Canada
Benjamin Nemoz	Scripps Research	San Diego, California, United States
Hadas Neuman	Bar-Ilan University	Ramat Gan, Tel Aviv
Ngan Nguyen	Atreca, Inc.	San Francisco, California, United States
Mats Ohlin	Lund University	Lund
Lauren Overend	The University of Oxford	Greater Oxford Area
Vagmita Pabuwal	Ligand Pharmaceuticals	Greater Seattle Area
Tim Pan	Ligand Pharmaceuticals	
Ravi Pandya	Microsoft Research	Bellevue, Washington, United States
Rachael Parks	Fred Hutch	Seattle, Washington, United States
Milena Pavlovic	University of Oslo	Oslo, Oslo
Emma Paytas	Takara Bio	San Diego, California, United States
Matt Pennell	University of British Columbia	Vancouver, British Columbia, Canada
Jane Pernes	University of Oxford	Oxford

Bjoern Peters	Lji	San Diego County, California, United States
Mikhail Pogorelyy	St. Jude Children's Research Hospital	Moscow, Moscow, Russia
Nate Polaske	Illumina	Tucson, Arizona, United States
Sabrina Pollastro	Sanquin Research	Amsterdam, North Holland
Nina Luning Prak	Perelman School of Medicine, University of Pennsylvania	Malvern, Pennsylvania, United States
Khang Le Quy	University of Oslo	Oslo, Oslo, Norway
Nagarajan Raju	Emory university	Atlanta, Georgia
Puneet Rawat	University of Oslo	Oslo
Janice Reichert	The Antibody Society	Boston, Massachusetts, United States
Eve Richardson	University of Oxford	Oxford, England, United Kingdom
Philippe Robert	Universitetet i Oslo	Norway
Sarah Robinson	Oxford Protein Informatics Group, University of Oxford	Greater Oxford Area
Artur Rocha	INESC TEC	Porto, Porto, Portugal
Oscar Rodriguez	University of Louisville School of Medicine	New York, New York, United States
Cynthia Ronkowski	University of Southern California	Los Angeles
Florian Rubelt	Roche	San Francisco Bay Area
Magdalena Russell	Fred Hutchinson Cancer Research Center (USA)	Seattle, Washington, United States
Swati Saha	University of Louisville	
Simon Schaefer	Fau Erlangen	Nuremberg, Bavaria, Germany
Cathrine Scheepers	University of the Witwatersrand	City of Johannesburg, Gauteng, South Africa
Lonneke Scheffer	University of Oslo	
Richard Scheuermann	J. Craig Venter Institute	San Diego, California, United States
Tilman Schneider-Hohendorf	University of Muenster, Germany	Münster
Chaim Schramm	NIH/NIAID/VRC	New York, New York, United States
Gary Schroth	Illumina	Danville, California, United States
Nicholas Schwab	University Of Münster	Ruhr Region
Jamie Scott	Simon Fraser University	
Ranjan Sen	National Institute on Aging, NIH (USA)	Baltimore, Maryland, United States
Nina Senna	Omniscope	Oxford
Seung Kak Shin		

Katherina Siewert	German Federal Institute for Risk Assessment	Berlin, Berlin
Troy Sincomb	Scripps Research	
Andrei Slabodkin	Greiff Lab, the University os Oslo	Greater Oslo Region
Melissa Smith	University of Louisville	Louisville, Kentucky, United States
Eva Smorodina	Harvard University	Yerevan, Armenia
Eva Smorodina		
Eric de Sousa	Champalimaud Foundation	Lisbon
Laura Spector	Specifica, Inc.	Santa Fe
Ulrik Stervbo	Ruhr University Bochum	
Paul Stys	Sorbonne University	Paris
Yvonne Suessmuth		
Qian-Quan Sun	University of Wyoming	
Suhas Sureshchandra	University of California	United States of America
Ana Teles	University of Hamburg	Greater Hamburg Area
Riva Tez		
Linnea Thörnqvist	Lund University	Lund, Skåne County, Sweden
Raphael Trevizani	LJI	
Claudia Tsunoda	Takara Bio	San Francisco Bay Area
Kelvin Tuong	University of Cambridge	Cambridge, England
Amit Upadhyay	Emory University	Decatur, Georgia, United States
Eugen Ursu	CellFabrik	Bucharest, Romania
Sebastiaan Valkiers	University of Antwerp	Antwerp
Vijay Vaswani	Omniscope	San Francisco Bay Area
Marcos Vieira		Greater Chicago Area
Jordan (Jo) Villa	KOTAI Biotechnologies, Inc.	Osaka, Ōsaka
Eric Waltari	Chan Zuckerberg Biohub	San Francisco, California, United States
Mamie Wang	Yale University	
Po-hung Wang	KOTAI Biotechnologies, Inc.	Osaka, Japan
Corey Watson	University of Louisville	
Simon Watson	Sanofi	Cambridge
Anna Weber	ETH Zurich	Zürich
Jordan Willis	IAVI Neutralizing Antibody Center at Scripps Research	Kansas City/San Diego
Adam Woolfe	SABER BIO	Paris, Île-de-France, France
Bryan Wu	GSK	Stevenage, England
Zeynep Kosaloglu Yalcin	La Jolla Institute for Immunology	San Diego, California, United States
Kazuo Yamashita	KOTAI Biotechnologies, Inc.	Osaka, Japan

Jabran Zahid	Microsoft	
Maxim Zaslavsky	Stanford University	
Chunzhao Zhang	University of Wyoming	
Yulei Zhang	Ligand Pharmaceuticals	
Weian Zhao	University of California, Irvine	Irvine, California, United States
Jahn Zhong	University of Oslo	Erlangen
Bojan Zimonja	Simon Fraser University	Canada

Thank you to everyone for attending!