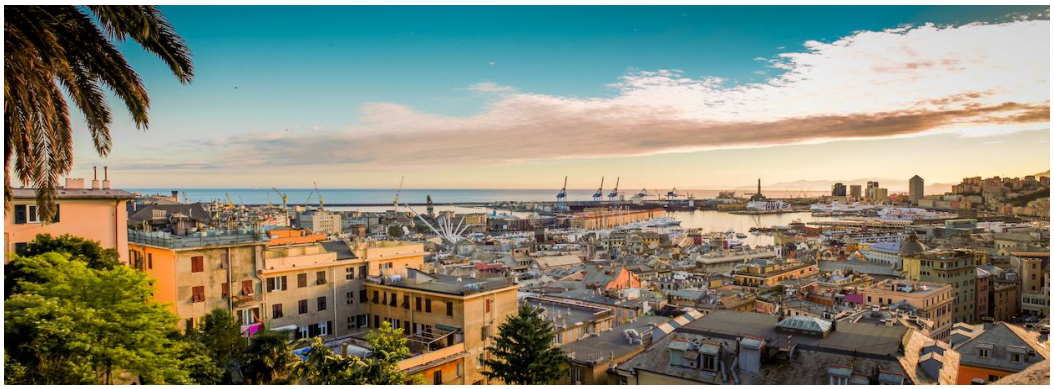
AIRR Community Meeting IV: Bridging the Gaps

May 11-15, 2019, University of Genoa, Italy



Twitter: #AIRRC4

Table of Contents

Welcome from the AIRR-C Meetings Committee and the AIRR Community Chair

Welcome to the fourth meeting of the Adaptive-Immune Receptor Repertoire (**AIRR**) Community of The Antibody Society, hosted by the University of Genoa - AIRR Community Meeting IV: Bridging the Gaps.

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.

The theme of this year's meeting, "***Bridging the Gaps***", addresses two areas of concern:

i. Technological gaps between the huge amounts of AIRR-seq data that are rapidly accumulating and our ability to process and share them; and

ii. Communication gaps reflecting the need to engage and disseminate the shared standards, values and practices envisioned by the AIRR Community to stakeholders, such as industry, medical and patient-advocate communities.

We have planned an exciting meeting, full of activities and opportunities for learning and networking. Over the weekend, AIRR Community activities include two Workshops teaching the basics of the immune system and of AIRR-seq data, progress reports from our Working Groups and Sub-committees, and two panel discussions on germline gene databases and sharing standards and software. The "formal meeting" on Monday and Tuesday comprises basic-science and biomedical Scientific Sessions with an outstanding line up, including keynote lectures by Sai Reddy and Antonio Lanzavecchia, respectively, invited presentations, and short presentations chosen from the poster abstracts. Two Challenge Sessions will address the above-mentioned technical and communication gaps. In addition, attendees can participate in software-tool demos and hands-on pipeline tutorials. Be sure to attend the receptions on Sunday and Monday evenings, and the final group dinner on Tuesday. 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, will be ratified.

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!

Felix Breden (Simon Fraser University), Chair, AIRR Community
Nina Luning Prak (University. of Pennsylvania), Chair-Elect, AIRR Community

And Members of the Meeting Subcommittee:
Davide Bagnara (University of Genoa, Local Host)
Pam Borghardt (Simon Fraser University, Co-leader & Meeting Manager)
Jean-Philippe Bürckert (BISC Global)
Ramit Mehr (Bar Ilan University.)
Jamie Scott (Simon Fraser University, Co-leader)

Thanks to our amazing volunteers!
Lorissa Corrie (University of Victoria, Canada)
Monica Colombo (Ospedale Policlinico San Martino, Italy)
Filippo Vit (Centro di Riferimento Oncologico, Italy)
Martina Cardillo (University of Genoa, Itlay)

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 in our sponsor reception on Tuesday, May 14th.

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

Sponsor	**Website**	**Level**
Canadian Institutes of Health Research	www.cihr-irsc.gc.ca	Funding Agency
Simon Fraser University	www.sfu.ca	Funding Agency
10X Genomics	www.10xgenomics.com	Silver
BISC Global, Inc.	www.biscglobal.com	Silver
Distributed Bio	www.distributedbio.com	Silver
Geneious Biologics	www.geneious.com/biopharma	Silver
Roche	www.roche.com	Silver
The Antibody Society	www.antibodysociety.org	Silver
Dimes	www.dimes.unige.it	Bronze
Chan Zuckerberg Biohub	www.czbiohub.org	Bronze
Grifols	www.grifols.com	Bronze
Hamilton	www.hamiltoncompany.com	Bronze
SFU Faculty of Health Sciences	www.sfu.ca/fhs.html	Bronze
Takara	www.takarabio.com	Bronze
Universita Degli Studi Di Genova	www.unige.it	Bronze
Abnomx	www.abnomx.com	Bronze

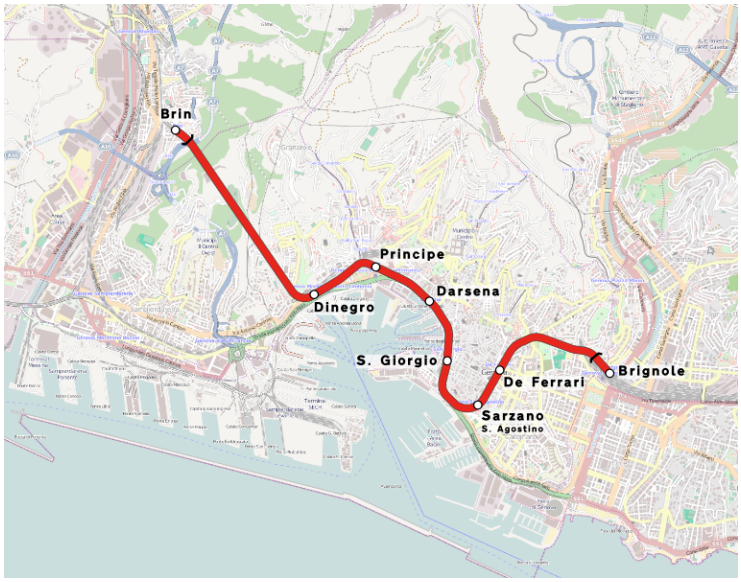
Genoa, Italy (Transit, University Map, Restaurants, Lodging, Tourism)

We have created an interactive map that includes links to restaurants, lodging, transit and directions to the Meeting location. Further details below including a university map with names of rooms where events will take place.

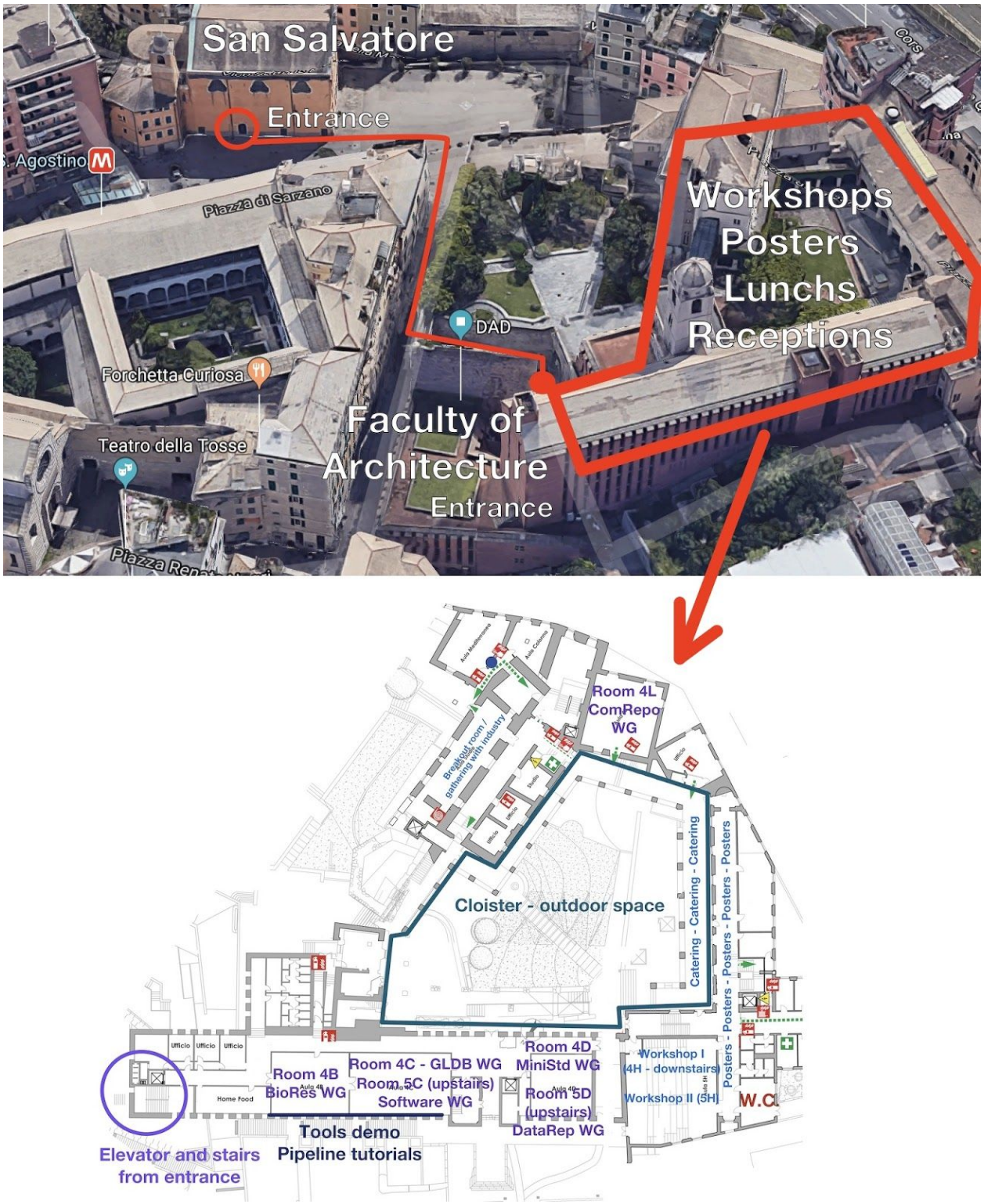
Transit Options

From the Airport:
From the airport the 2 most convenient connection with the city center are taxi and a bus called “Volabus”. Suggest ignoring other buses and the train.
The airport website (https://www.airport.genova.it/en/transportation-airport/) has all the detailed the information in English.
The only extra note is about the Volabus. Participants to the conference should drop-off at the stop “DE FERRARI/METRO”, the closest to the conference and Hotels (https://www.amt.genova.it/amt/trasporto-multimodale/partenze/?lineatiVLB), which is indeed located in Piazza De Ferrari.

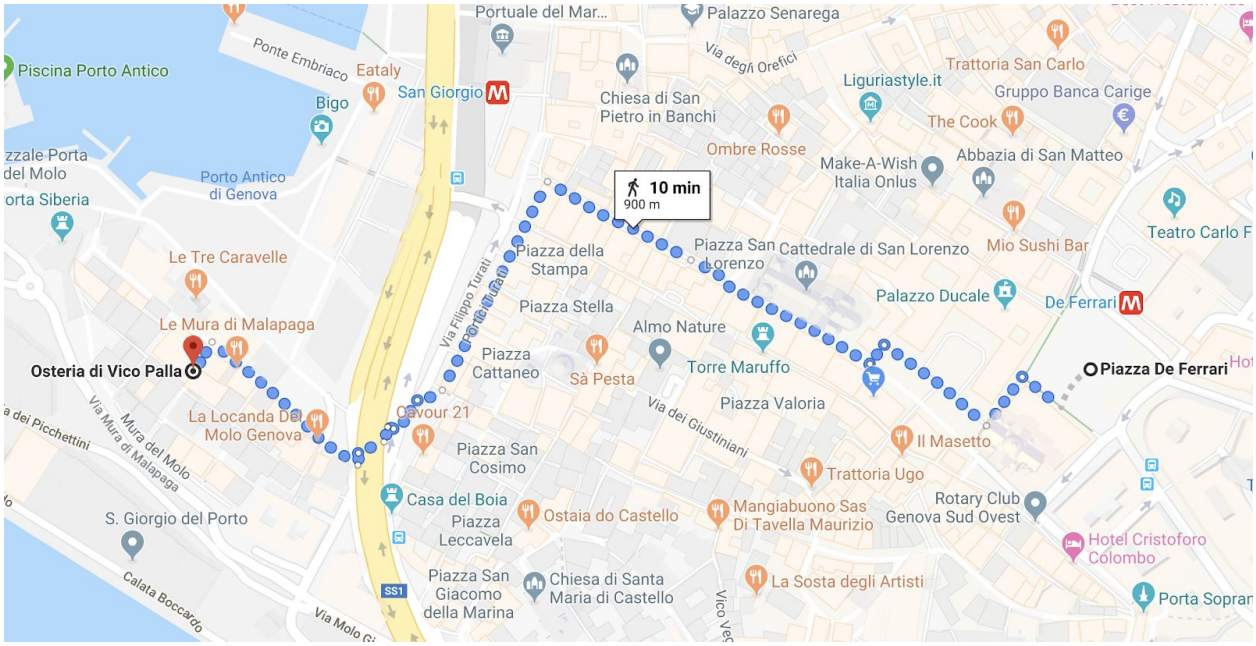
From the Train Station:
From train station Piazza Principe take the subway from “Principe” direction “Brignole”.
The meeting will be next “Sarzano”, while most hotels will be close to stop “De Ferrari” or “San Giorgio”.



Meeting Location, University of Genoa:



Group Dinner Tuesday May 14, 7:30 pm
Pre-registration required
Osteria di Vico Palla,
Vico Palla, 15, 16128 Genova GE



Tourism Genoa
http://www.visitgenoa.it/en/homepage
http://www.portoantico.it/en/information/information-and-welcoming-service/

Phone SIM cards
https://www.tim.it/tim-tourist-en

AIRR-Community Meeting IV: Bridging the Gaps
University of Genoa, Genoa, Italy
11-15 May 2019

Meeting at a glance

<table>
<tr><th>Time</th><th colspan="3">Saturday
May 11</th><th>Sunday
May 12</th><th>Monday
May 13</th><th>Tuesday
May 14</th><th>Wednesday
May 15</th><th>Time</th></tr>
<tr><td>08:00 - 08:30</td><td colspan="3" rowspan="2"></td><td rowspan="2"></td><td colspan="3">Registration, Coffee</td><td></td><td>08:00 - 08:30</td></tr>
<tr><td>08:30 - 09:00</td><td rowspan="4">Keynote I and Scientific Session I</td><td rowspan="4">Keynote II and Scientific Session II</td><td>Coffee</td><td>08:30 - 09:00</td></tr>
<tr><td>09:00 - 09:30</td><td colspan="3" rowspan="2">Registration, Coffee & Snack</td><td>Registration, Coffee & Snack</td><td rowspan="5">AIRR-C business</td><td>09:00 - 09:30</td></tr>
<tr><td>09:30 - 10:00</td><td rowspan="3">Introduction
WG Reports: Software
Data Representation (DataRep)
Minimal Standards (MiniStd)
Common Repository (Common Repo)</td><td>09:30 - 10:00</td></tr>
<tr><td>10:00 - 10:30</td><td rowspan="6">WGs meet separately (Lunch provided)</td><td rowspan="3">Workshop I - Fundamentals of the Immune System</td><td rowspan="3">Workshop II - AIRR-Seq Data Processing and Analysis</td><td>10:00 - 10:30</td></tr>
<tr><td>10:30 - 11:00</td><td>10:30 - 11:00</td></tr>
<tr><td>11:00 - 11:30</td><td>11:00 - 11:30</td></tr>
<tr><td>11:30 - 12:00</td><td colspan="3">Light refreshment break</td><td>11:30 - 12:00</td></tr>
<tr><td>12:00 - 12:30</td><td colspan="2" rowspan="2">Registration open</td><td rowspan="2">Panel Discussion: Common Repo, DataRep and MiniStd WGs</td><td rowspan="3">Lunch + Poster Session</td><td rowspan="3">Lunch + Poster Session</td><td rowspan="3">Meeting Adjourns. (Lunch on your own)</td><td>12:00 - 12:30</td></tr>
<tr><td>12:30 - 13:00</td><td>12:30 - 13:00</td></tr>
<tr><td>13:00 - 13:30</td><td>13:00 - 13:30</td></tr>
<tr><td>13:30 - 14:00</td><td colspan="2" rowspan="2">Bio break</td><td rowspan="2">Intro to Communications SC, report & discussion</td><td>13:30 - 14:00</td></tr>
<tr><td>14:00 - 14:30</td><td rowspan="3">Challenge Session I</td><td rowspan="3">Challenge Session II</td><td rowspan="4">Committee & WG wrap-up meetings</td><td>14:00 - 14:30</td></tr>
<tr><td>14:30 - 15:00</td><td colspan="3">Registration open</td><td rowspan="3">AIRR-C Meeting Welcome Reception</td><td rowspan="2"></td><td>14:30 - 15:00</td></tr>
<tr><td>15:00 - 15:30</td><td colspan="3" rowspan="2">Welcome & Introduction
Executive SC Report</td><td>15:00 - 15:30</td></tr>
<tr><td>15:30 - 16:00</td><td rowspan="2">AIRR-Seq Tool Demos</td><td>15:30 - 16:00</td></tr>
<tr><td>16:00 - 16:30</td><td colspan="3">Biological Resources WG Report</td><td rowspan="2">Pipeline tutorials - sign-up required</td><td rowspan="2">Reception w Industry</td><td rowspan="6"></td><td>16:00 - 16:30</td></tr>
<tr><td>16:30 - 17:00</td><td colspan="3">Germline Database (GLDB) WG Report</td><td>16:30 - 17:00</td></tr>
<tr><td>17:00 - 17:30</td><td colspan="3" rowspan="2">Panel Discussion: Germline Databases WG</td><td rowspan="3">Dinner on your own</td><td>17:00 - 17:30</td></tr>
<tr><td>17:30 - 18:00</td><td rowspan="3">Dinner on your own</td><td>17:30 - 18:00</td></tr>
<tr><td>18:00 - 18:30</td><td colspan="3" rowspan="4">Dinner on your own</td><td rowspan="4">Group Dinner (requires separate reservation)</td><td>18:00 - 18:30</td></tr>
<tr><td>18:30 - 19:00</td><td>18:30 - 19:00</td></tr>
<tr><td>19:00 - 19:30</td><td>19:00 - 19:30</td></tr>
<tr><td>19:30 - 20:00</td><td>19:30 - 20:00</td></tr>
<tr><td>20:00 - 20:30</td><td colspan="3"></td><td colspan="3"></td><td>20:00 - 20:30</td></tr>
<tr><td>20:30 - 21:00</td><td colspan="3"></td><td colspan="3"></td><td>20:30 - 21:00</td></tr>
<tr><td>21:00 - 21:30</td><td colspan="3"></td><td colspan="3"></td><td>21:00 - 21:30</td></tr>
<tr><td>21:30 - 22:00</td><td colspan="3"></td><td colspan="3"></td><td>21:30 - 22:00</td></tr>
</table>

AIRR-Community Meeting IV: Bridging the Gaps
University of Genoa, Genoa, Italy
11-15 May 2019

Full Agenda

See Virtual Attendance details at the bottom of full agenda

Saturday, 11 May: WG Meetings, Workshops and WG Progress Reports			
Start	**End**	**Activities**	**Location**
9:00	10:00	*Registration and Coffee*	Faculty of Architecture Cloister 4th floor
10:00	15:00	**Working Groups (WG) meeting** Biological Resources (BioRes): Room 4B Germline DB (GLDB): Room 4C Software: Room 5C Minimal Standards (MiniStd): Room 4D Data Representation (DataRep): Room 5D Common Repository (ComRepo): Room 4L *(Each WG will provide a separate agenda to its members)*	Faculty of Architecture 6 rooms
10:00	13:00	**Workshop I: Jamie Scott**, Simon Fraser University (CA): "*Fundamentals of the immune system"*	Faculty of Architecture Room 4H
10:00	13:00	**Workshop II: Victor Greiff,** University of Oslo (NO): "*Steps in AIRR data processing and analysis: best practices, pitfalls, and future directions"*	Faculty of Architecture Room 5H
13:00	14:00	*Registration open*	
13:00	14:00	*Lunch*	Faculty of Architecture Cloister
15:00	15:30	*Break*	
15:00	15:30	*Registration open*	Chiesa di San Salvatore
15:30	15:40	**Welcome and voting reminder by the Chair of the AIRR-C Executive Sub-committee:** Felix Breden, Simon Fraser University (CA)	
15:40	16:40	**Report from Executive Sub-committee**	
		Introduction to Working Groups, and reports on progress, future plans & leadership (20-min Reports + 10-min Discussion)	
16:40	17:10	**Biological Resources WG Report**	
17:10	17:40	**Germline DB WG Report**	
17:40	18:30	**Panel Discussion: Germline Databases (GLDB) WG** *"Beyond inferred alleles: Improving the quality, scope and openness of GLDBs through software inference, open publication and population-level data"* **Chair:** Christian Busse, German Cancer Research Center (DE) **Panel Members:** **Martin Corcoran**, Karolinska Institute (SE) **Henk-Jan van den Ham**, Enpicom BV (NL) **William Lees**, Birkbeck College, University of London (GB) **Mike Stubbington**, 10X Genomics (GB) **Corey Watson**, University of Louisville School of Medicine (US)	
		Dinner on your own	

Sunday, 12 May: WG and Communication Subcommittee Progress Reports			
Star	**End**	**Activities**	**Location**
9:00	9:30	*Registration open; Coffee & snack*	Chiesa di San Salvatore
9:30	9:40	**Introduction to Day 2 of Progress Reports by the Chair of the AIRR-C Executive Sub-committee: Felix Breden,** Simon Fraser University (CA) (20-min Report + 10-min Discussion)	
9:40	10:10	**Software WG Report**	
10:10	10:40	**Data Representation WG Report**	
10:40	11:10	**Minimal Standards WG Report**	
11:10	11:40	**Common Repository WG Report**	
11:40	12:30	*Light refreshment*	
12:30	13:50	**Panel Discussion: Common Repositories, Data Representation and Minimal Standards WGs** *"Towards Solutions: The Benefits and Challenges of Sharing AIRR-seq Data"* **Chairs: Christian Busse**, German Cancer Research Center (DE) & **Lindsay Cowell**, University of Texas, Southwestern Medical Center (US) **Panel Members:** **Bryan Briney**, Scripps Research (US) **Konrad Krawczyk**, Natural Antibody (DE) **Anton Langerak**, Erasmus MC (NL) **Xiao Liu**, BGI Shenzhen (CN)	
13:50	14:10	**Introduction to the Communications Sub-committee & Report**	Chiesa di San Salvatore
14:10	14:30	**Meeting Sub-committe report & recommendations for AIRR C Meeting V**	
15:00	18:00	**AIRR-C Meeting Welcome Reception & AIRR-seq Tool Demonstrations**	Faculty of Architecture Cloister 4B-4C
16:00	18:00	**AIRR-seq Tool Demonstrations** (10-min/presenter) **Chair: Lindsay Cowell**, University of Texas, Southwestern Medical Center (US) **Presenters:** **Lorenzo Fanchi**, Enpicom B.V. (NL): **"***Intuitively manage, store, analyze, and visualize repertoire sequencing data using the ImmunoGenomiX ('IGX') platform"* **Bryan Briney**, Scripps Research (US): "*Massively scalable genetic analysis of antibody repertoires"* **Brian Corrie**, Simon Fraser University (CA): "*iReceptor – A platform for querying and analyzing antibody/B-cell and T-cell receptor repertoire data across federated repositories"* **Kenneth Hoehn**, Yale University (US): "*Phylogenetic analysis of B cell repertoires with IgPhyML"* **Steven Kleinstein**, Yale School of Medicine (US): "*The Immcantation Framework for analysis of AIRR-seq data"* **William Lees**, Birkbeck College, University of London (GB): "*OGRDB – the Open Germline Receptor Database"* **Vadim Nazarov**, National Research University (RU): "*Bioinformatics software problems and how to make immune data analysis effortless with immunarch"* **Duncan Ralph**, Fred Hutchinson Cancer Research Centre (US): "*Partis: A tool for accurate and efficient annotation of germline and clonal family inference"* **Mikael Salson**, Universite de Lille - VidjlNet (FR): *"Vidjil: an open-source platform for interactive repertoire analysis"*	Faculty of Architecture Room 4B-4C
		Dinner on your own	

Monday, 13 May: Scientific Progress & Challenges (Basic Science)			
Start	**End**	**Activities**	**Location**
8:00	8:30	*Registration open*	Chiesa di San Salvatore
8:00	8:30	*Coffee available at the meeting site*	
8:30	8:40	**Introduction to the Meeting and the Keynote by the AIRR-C Executive Sub-committee Chair, Felix Breden,** Simon Fraser University (CA)	
8:40	9:30	**Keynote: Sai Reddy,** ETH Zurich (CH): *"Integration of systems immunology and immune repertoires"* (40-min talk + 10-min Q&A)	
9:30	9:40	**Advances in Understanding Biological Processes of the AIRR** (25-min talks + 10-min Q&A) **Chair: Andrew Collins,** University of New South Wales (AU): *Session introduction*	
9:40	10:10	**Talk 1: Nicholas Schwab,** University of Muenster (DE): "*Sex bias in MHC I-associated shaping of the adaptive immune system"*	
10:10	10:50	**Talk 2: Lindsay Cowell**, University of Texas, Southwestern Medical Center (US): "*Statistical classifiers for detecting phenotype-associated biophysicochemical motifs in AIRR"*	
10:50	11:20	*Coffee Break and Group Photo*	
11:20	11:55	**Talk 3: David Klatzmann**, Sorbonne Universite (FR): "*How specific is specific (in the T-cell response)?"*	
11:50	12:30	**Talk 4: Tom Parks**, University of Oxford (UK): "*Germline immunoglobulin heavy-chain locus variation and susceptibility to group A streptococcal disease"*	
12:30	13:00	**Short Oral Presentations selected from Poster Abstracts** (3x10 minutes, no discussion) **Chair: Ramit Mehr,** Bar-Ilan University (IL) **Presenters:** **Kenneth Hoehn**, Yale University (US): *"Repertoire-wide phylogenetic models of B-cell molecular evolution reveal evolutionary signatures of aging and vaccination"* **Ivana Mikocziova**, University of Oslo (NO): "*AIRR-seq reveals a large number of germline polymorphisms in the variable immunoglobulin genes"* **Pieter Meysman**, University of Antwerp (BE): "*On the viability of unsupervised T-cell receptor sequence clustering for epitope preference"*	
13:00	15:00	*Lunch Break* + **Poster Session**	Faculty of Architecture Cloister + Hallway
15:00	15:10	**Challenges in Processing and Storing Massive AIRR-seq Data (>1bn sequences)** (20-min talk + 15-min discussion) **Chair: Uri Laserson**, Icahn School of Medicine at Mount Sinai (US): *Session introduction*	Chiesa di San Salvatore
15:10	15:40	**Talk 1 + Discussion - Processing & Annotation: Mikhail Shugay**, Skolkovo Institute of Science and Technology (CZ): **"***Challenges in motif discovery and comparative analysis of large T-cell receptor repertoire datasets***"**	
15:40	16:20	**Talk 2 + Discussion - Analysis: Bryan Briney**, Scripps Research (US): "*Commonality despite exceptional diversity in the baseline human antibody repertoire*"	
16:20	16:50	**Talk 3 + Discussion - Data Storage & Access: Xiao Liu**, BGI Shenzhen (CN): "*Pan Immunome Initiative and applications in health and disease"*	
17:00	19:00	**Pipeline + Data Repository Tutorials** + light reception (sign-up required through registration site, attendees may register for up to two 45-min sessions) **Presenters:** **Henk-Jan Van den Ham**, Enpicom B.V.: "*IGX"* **Cecilie Boysen**, Biomatters Ltd.: *"Geneious Biologics"* **Bryan Briney**, Scripps Research: "*AbCloud"* **Scott Christley**, UT Southwestern Medical Center: "*VDJServer"* **Brian Corrie**, Simon Fraser University: "*iReceptor"* **Khalil EL Mazouari**, AbnomX: *"Antibody-Extractor"* **Jacob Glanville**, Distributed Bio: "*AbGenesis"* **Kenneth Hoehn**, Yale University: "*Immcantation and IgPhyML"* **William Lees**, Birkbeck College, University of London: **"***OGRDB"* **Vadim Nazarov**, National Research University: "*PyTorch"* **Ayelet Peres**, Bar Ilan University: *"RAbHIT"*	Faculty of Architecture Cloister 4B-4C

		Mikhail Shugay, Skolkovo Institute of Science and Technology: "*VDJtools"*	
		Dinner on your own	

Tuesday, **14 May:** Scientific Progress & Challenges (Clinical and Translational Science)			
Start	**End**	**Activities**	**Location**
8:00	8:30	*Registration open*	Chiesa di San Salvatore
8:00	8:30	*Coffee available at the meeting site*	
8:30	8:40	**Introduction to the Keynote by AIRR-C Chair-Elect Nina Luning Prak**, University of Pennsylvania (US)	
8:40	9:30	**Keynote: Antonio Lanzavecchia,** Institute for Research in Biomedicine (CH): "*Lessons from the analysis of the immune response to P. falciparum: The power of clonal selection"* (40-min talk + 10-min Q&A)	
9:30	9:40	**Advances in AIRR-seq Data Addressing Human Health and Disease** (25-min talks + 10-min Q&A) **Chair: Steven Kleinstein,** Yale School of Medicine (US): *Session introduction*	
9:40	10:10	**Talk 1: Rachael Bashford-Rogers,** University of Oxford (UK): "*Mapping B-cell receptor repertoires in immune-mediated diseases"*	
10:10	10:50	**Talk 2: Rob Holt**, BC Cancer Research Centre, UBC/SFU (CA): *"Mining the tumor-associated T cell repertoire for TCR therapeutics"*	
10:50	11:20	*Coffee Break*	
11:20	11:50	**Talk 3: Olga Britanova,** Shemyakin and Ovchinnikov Institute of Bioorganic Chemistry (RU): *"Potential use of monoclonal anti-TCR antibody for treatment of autoimmune diseases."*	
11:55	12:30	**Talk 4: Ziv Shulman**, The Weizmann Institute of Science (IL): *"Dynamics of B cell immune responses in intestinal tissues"*	
12:30	13:00	**Short Oral Presentations selected from Poster Abstracts** (3x10 minutes, no discussion) **Chair: Ramit Mehr,** Bar-Ilan University(IL) **Presenters:** **Felix Breden**, Simon Fraser University (CA): *"iReceptor Plus facilitates sharing and analysis of AIRR-seq data"* **Xihao Hu**, Dana Farber Cancer Institute (US): "*Landscape of B-cell immunity and related immune evasion in human cancers"* **Keshav Motwani**, University of Florida (US): "*T-cell receptor repertoires in peripheral blood encode type-1 diabetes status"*	
13:00	14:30	*Lunch Break* + **Poster Session**	Faculty of Architecture Cloister + Hallway
14:30	14:40	**Patient & Community Engagement** (15-min talk + 5-min Q&A) **Chair: Tania Bubela**, Simon Fraser University (CA): *Session introduction*	Chiesa di San Salvatore
14:40	15:00	**Short Talk 1: Tania Bubela**: "*A primer on patient-oriented research*"	
15:00	15:20	**Short Talk 2: Harriet Teare**, University of Oxford (UK): "*Dynamic consent: Individual oversight and control of data and samples"*	
15:20	15:40	**Short Talk 3: Namaste Marsden**, BC First Nations Health Authority (CA): "*We are relational: The challenge of Indigenous collective rights in bio-ethics"*	
15:40	16:10	**Panel Discussion: Patient & Community Engagement**	
16:10	16:30	*Coffee*	
16:30	16:40	**Engaging Industry** (15-min talk + 5-min Q&A) **Chair: Jean-Philippe Bürckert,** BISC Global Inc. (US): *Session introduction*	
16:40	17:00	**Short Talk 1: Jacob Glanville**: Distributed Bio (US): "*Applied repertoire analysis - Computational immunoengineering of therapeutic antibodies and TCRs"*	
17:00	17:20	**Short Talk 2: Jonathan Carlson**: Microsoft Research (US): *"The Adaptive-Microsoft TCR-Antigen Map Project"*	
17:20	17:40	**Short Talk 3: Nicola Bonzanni**: Enpicom (NE): "*Bridging the gap: Manage, analyze, visualize and*	

		interpret repertoire sequencing data using the IGX platform"	
17:40	17:50	**Poster Sessions - Winner Announcements**	
17:50	19:00	**Industry Reception** Meet with and learn more about the innovations of our meeting sponsors: 10x Genomics, AbnomX, BISC Global, CZ Biohub, Distributed Bio, Geneious Biologics, Grifols, Hamilton, Roche, Takara Bio	Faculty of Architecture Glass Room 4th floor
19:30	22:00	*Group Dinner at Osteria di Vico Palla (pre-registration required)*	Osteria di vico Palla

Wednesday, 15 May: AIRR-C Organisational Session			
Star	**End**	**Activities**	**Location**
8:30	9:00	*Coffee available at Meeting Site*	Chiesa di San Salvatore
9:00	9:10	**Introduction to the General Meeting of the AIRR Community Members by Chair and Chair-Elect: Felix Breden,** Simon Fraser University (CA) **& Nina Luning Prak,** University of Pennsylvania (US)	
9:10	11:10	**AIRR-C Working Groups work plans and leadership** Each of the 6 WGs ● Recap proposed work plan as a motion (vote) ● Leadership as a motion (vote) (20 min each including voting!)	
11:10	11:30	*Coffee*	
11:03	12:30	**AIRR-C Sub-committee(s) plans and voting** ● Recap proposed work plans as a motion (vote) ● Leadership as a motion (vote)	
11:30	12:00	Executive Sub-committee (ratify governance revisions)	
12:00	12:10	Meeting Sub-committee (ratify where & when will the next meeting be held)	
12:10	12:30	Communications Sub-committee (ratify as a new standing sub-committee)	
12:30	14:30	*Meeting adjourned & lunch on your own*	
14:30	17:30	Sub-committees and Working Groups - wrap-up meetings Biological Resources (BioRes): Room 5A Germline DB (GLDB): Room 5B Software: Room 5C Minimal Standards(MiniStd): Room 6C Data Representation (DataRep): Room 6E Common Repository (ComRepo): Breakout Room	Faculty of Architecture 6 rooms

Attending Virtually

The AIRR-C Meetings Sub-committee is pleased to provide a Virtual Attendance option for AIRR-Community Meeting IV. For meeting statistics purposes please register as a Virtual Only attendee at https://www.regonline.com/builder/site/Default.aspx?EventID=2549580.

AIRR-Community Meeting IV will be streamed live through GoToMeeting. Virtual attendees will be able to follow all of the scientific sessions seeing the slides and listening to the speakers. They will also have the chance to ask questions via live chat during the Q&A portion of each session. Login to our video conference system using the link or phone connection and access code below.

First-time users: Be sure to give yourself time to upload and test the GoToMeeting app well in advance of the first session you plan to attend. Please don't hesitate to contact us at meetings@iairrc.antibodysociety.org if you experience any issues during our live sessions or if you need additional phone numbers to dial in. Sessions will also be recorded, and will be uploaded to the AIRRC website after the meeting's end.

GoToMeeting Event Title: AIRR-C Meeting IV - Bridging the Gaps

Please join my meeting from your computer, tablet or smartphone.
https://global.gotomeeting.com/join/965748445

Access Code: 965-748-445

You can also dial in using your phone.
Australia: +61 2 8355 1050
Austria: +43 7 2081 5427
Belgium: +32 28 93 7018
Canada: +1 (647) 497-9391
Denmark: +45 32 72 03 82
Finland: +358 942 72 1060
France: +33 170 950 594
Germany: +49 692 5736 7317
Ireland: +353 16 572 651
Italy: +39 0 230 57 81 42
Netherlands: +31 207 941 377
New Zealand: +64 9 280 6302
Norway: +47 21 93 37 51
Spain: +34 932 75 2004
Sweden: +46 853 527 836
Switzerland: +41 225 4599 78
United Kingdom: +44 20 3713 5028
United States: +1 (646) 749-3129

New to GoToMeeting? Get the app now and be ready when your first meeting starts:
https://global.gotomeeting.com/install/965748445

Workshops

Title: Steps in AIRR data processing and analysis: best practices, pitfalls, and future directions

Presenter: Victor Greiff
Department of Immunology
University of Oslo, Oslo, Norway

Abstract: High-throughput sequencing (HTS) has enabled the capture of adaptive immune receptor repertoire (AIRR) data at unprecedented depth and precision. This workshop will give an in-depth walk-through on best practices to conceive, analyze and perform AIRR studies for answering fundamental immunological questions as well as discovering novel immunodiagnostic biomarkers and design (therapeutic) immune receptors. Specifically, I will address current approaches to perform AIRR-compliant AIRR data processing encompassing bulk and single-cell approaches and experimental and bioinformatics quality control. Furthermore, I will summarize the computational methods that have been recently developed to deconstruct the high-dimensional complexity of immune receptor repertoires. For example, (i) diversity-, (ii) phylogenetic-, (iii) networks- and (iv) machine learning-based methods have been applied to dissect and understand the diversity, architecture, evolution and antigen specificity of immune repertoires. Finally, I will discuss experimental and computational methods in light of their underlying assumptions, limitations and pitfalls and highlight promising avenues of future research in basic and applied AIRR systems immunology.

Funding: University of Oslo [UiO] World-Leading Research Community, UiO Convergence Environment ImmunoLingo, JDRF/Helmsley Charitable Trust, EU-H2020/iReceptor+

Title: Fundamentals of the Immune System

Presenter: Jamie Scott
Simon Fraser University,
Burnaby, Canada

Abstract: In this workshop, I will first provide an overview of humoral and cellular immunity, and the basic structure of the immune system, including its cells, tissues and compartments, along with the “superhighway” of the immune system: the circulatory and lymphatic systems. In that context, innate and adaptive immune systems and their interaction, and the general timing and dynamics of immune responses will be discussed.

The processes of lymphocyte development, including the various B- and T-cell subsets, positive and negative selection, and the genetic basis of B-cell and T-cell receptor diversification, will be presented to provide a clear idea of what *adaptive-immune receptor repertoires (AIRRs)* are, and in general terms, how they are currently assessed via high-throughput sequencing.

We will then cover the signaling, activation, proliferation and differentiation of T-cell and B-cell clones in the context of lymphoid compartments where antigen is concentrated and presented to naïve and memory B and T cells. The role of co-stimulation in determining the type immune response generated will be emphasized.

The orchestration of systemic and mucosal immune responses will be reviewed, including the roles of tolerance and inflammation in these processes. Finally, a few examples of immune responses to vaccines, chronic viral infection, and/or cancer, as well as autoimmunity, will be presented as variations on a common theme, reiterating the dynamics of the immune response. Some engineered immunotherapies, such as therapeutic antibodies, CAR-T cells and dendritic-cell vaccines, will be introduced as well.

The importance of AIRR-sequencing (AIRR-seq) data to our understanding of immune responses will be emphasized throughout the latter half of the workshop.

Funding:

Invited Speakers

Title: Encoding and decoding specificity in antibody repertoires

Presenter: Sai Reddy
Associate Professor, ETH Zurich, Department of Biosystems Science and Engineering
4058 Basel, Switzerland

Abstract: The ability to predict and correspondingly manipulate antibody responses is highly valuable for biotechnology and medicine. To achieve this requires a greater molecular understanding of antigen selection and specificity in antibody repertoires. In this presentation I will describe how we are decrypting antibody repertoire sequence space by identifying antigen-specific molecular patterns. We combine deep sequencing with deep learning to accurately predict antigen exposure and antigen specificity based on antibody sequencing.

Brief Biography
Sai Reddy is an Associate Professor in the Department of Biosystems Science & Engineering, ETH Zurich, Switzerland. His research group uses methods in systems and synthetic biology to study and manipulate immune responses for applications in biotechnology, vaccination, and immunotherapy. Sai Reddy holds B.S. (2003) and M.S. (2004) in Biomedical Engineering from Northwestern University (Evanston, IL, USA). He completed his Ph.D. thesis at Ecolé Polytechnique Féderale de Lausanne (EPFL, Switzerland) in Bioengineering and Biotechnology (2008). Sai Reddy did post-doctoral research at the University of Texas, Austin (2008-2011).

Funding:

Title: Sex bias in MHC I-associated shaping of the adaptive immune system

Presenter: Nicholas Schwab
Tilman Schneider-Hohendorf[a], Dennis Görlich[b], Paula Savola[c], Tiina Kelkka[c], Satu Mustjoki[c], Catharina C. Gross[a], Geoffrey C. Owens[d], Luisa Klotz[a], Klaus Dornmair[e], Heinz Wiendl[a], and Nicholas Schwab[a,1]
A - Department of Neurology, University of Muenster, 48149 Muenster, Germany;
B - Institute of Biostatistics and Clinical Research, University of Muenster, 48149 Muenster, Germany;
C - Hematology Research Unit Helsinki, Department of Clinical Chemistry and Hematology, University of Helsinki and Helsinki University Hospital Comprehensive Cancer Center, 00029 Helsinki, Finland;
D - Department of Neurosurgery, David Geffen School of Medicine at the University of California Los Angeles, Los Angeles, CA 90095; and
E - Institute of Clinical Neuroimmunology, University Hospital and Biomedical Center, Ludwig-Maximilians University Munich, 80539 Munich, Germany

Abstract: HLA associations, T cell receptor (TCR) repertoire bias, and sex bias have independently been shown for many diseases. While some immunological differences between the sexes have been described, they do not fully explain bias in men toward many infections/cancers, and toward women in autoimmunity. Next-generation TCR variable beta chain (TCRBV) immunosequencing of 824 individuals was evaluated in a multiparametric analysis including HLA-A-B/MHC class I background, TCRBV usage, sex, age, ethnicity, and TCRBV selection/expansion dynamics. We found that HLA-associated shaping of TCRBV usage differed between the sexes. Furthermore, certain TCRBVs were selected and expanded in unison. Correlations between these TCRBV relationships and biochemical similarities in HLA-binding positions were different in CD8 T cells of patients with autoimmune diseases (multiple sclerosis and rheumatoid arthritis) compared with healthy controls. Within patients, men showed higher TCRBV relationship Spearman's rhos in relation to HLA-binding position similarities compared with women. In line with this, CD8 T cells of men with autoimmune diseases also showed higher degrees of TCRBV perturbation compared with women. Concerted selection and expansion of CD8 T cells in patients with autoimmune diseases, but especially in men, appears to be less dependent on high HLA-binding similarity than in CD4 T cells. These findings are consistent with studies attributing autoimmunity to processes of epitope spreading and expansion of low-avidity T cell clones and may have further implications for the interpretation of pathogenic mechanisms of infectious and autoimmune diseases with known HLA associations.

Our findings add to the understanding of sex bias in diseases with immune system involvement: autoimmunity, infection, and cancer. These results also reveal pathology-associated TCRBVs of interest for future studies and support the argument for sex-separated analysis of HLA disease associations in general.

The presentation will conclude with an outlook on current projects to better understand autoimmunity (e.g. MS) and infectious diseases (e.g. John Cunningham infection and its associated pathology) by using bioinformatical tools.

Funding: This study was funded by DFG Grant SFB/CRC128 Project B1 (to N.S.), A5 (to K.D.), and Z2 (to H.W.); single Grant GR3946_3/1 (to C.C.G.); the KKNMS funded by the Federal Ministry of Education and Research (to H.W.); the IZKF Münster (Wie3/009/16) (to N.S. and

H.W.); the EU, M-IMM project; Academy of Finland; Finnish special governmental subsidy for health sciences, research and training; the Sigrid Juselius Foundation; and the Finnish Cancer Institute (S.M.).

Title: Statistical classifiers for detecting phenotype-associated biophysicochemical motifs in AIRR

Presenter: Lindsay Cowell
Jared Ostmeyer[1], Scott Christley[1], Lindsay G. Cowell[1,2]
[1] Department of Population and Data Sciences, Division of Biomedical Informatics, UT Southwestern Medical Center, Dallas, Texas, USA
[2] Presenter

Abstract: We have developed a novel approach for discovering immune repertoire sequence patterns associated with clinical phenotypes. Our approach has two important features: (1) By utilizing Multiple Instance Learning as the machine learning framework, our approach considers all sequences within a patient's repertoire, rather than relying on repertoire-level summary statistics, such as diversity. (2) Our approach searches for signatures within the space of biophysicochemical representations of AIR amino acid sequences, because receptors binding the same antigen are not likely to have the same amino acid sequence but are expected to have similar biophysicochemical properties. We have applied our approach to both B cell and T cell receptor repertoires. In the former case, we developed an IgH statistical classifier that distinguishes patients with relapsing remitting multiple sclerosis from patients with other neurological diseases with 87% accuracy by leave-one-out cross-validation on training data and 72% accuracy on validation data from a separate study. For TCR, we developed TCRβ statistical classifiers that distinguish tumor-associated AIRR from those derived from healthy tissue of the same organ. When comparing tumor with patient-matched healthy tissue, our method achieved classification accuracy of 93% and 94% by patient-holdout cross-validation for colorectal and breast cancer, respectively. When comparing tumor from ovarian cancer patients with healthy ovaries removed during hysterectomy, our method achieved 95% accuracy by leave-one-out cross-validation. For all data sets, permutation analysis gave accuracies of ~49%. The parameter values for each classifier revealed distinct biophysicochemical properties for each phenotype. We hypothesize that the biophysicochemical motifs could be useful for identifying disease-relevant antigens, monitoring response to therapy, predicting relapse, or developing diagnostics.

Funding:

Title: How specific is specific (in the T cell response)?

Presenter: David Klatzmann[1,2]
[1] Sorbonne Université, INSERM, Immunology-Immunopathology-Immunotherapy (i3), F-75005 Paris, France
[2] AP-HP, Hôpital Pitié-Salpêtrière, Biotherapy (CIC-BTi) and Inflammation-ImmunopathologyBiotherapy Department (i2B), F-75651, Paris, France

Abstract: Specificity is an unchallengeable attribute of the adaptive immune response. Vaccination against measles does not protect from the flu. However, findings are accumulating that should challenge the simplistic paradigm that specificity is conveyed by the triggering and expansion of cells that express TCRs specific for immunizing antigens. Our recent study of follicular T cells TCR repertoires led us to revisit the notion of specificity1 .

Follicular T cells are characterized by markers allowing their stringent characterization. They are found in germinal centers of lymph nodes, niches for the production of high-affinity antibodies by B cells. As the antibody produced are specific for the immunizing antigens, the TCR repertoire of follicular T cells responding to immunization was anticipated to be quite specific and thus restricted. This is not the case. A stimulated (expanded) quite polyclonal TCR repertoire is detected, revealing a bystander activation of “non-specific” T cells. We will speculate that this bystander activation is the 2nd “immunologists’ dirty little secret”2 and that the ultimate specificity of the T cell response is supported by fuzziness.

1 Ritvo PG, Saadawi A, Barennes P, Quiniou V, Chaara W, El Soufi K, Bonnet B, Six A, Shugay M, Mariotti-Ferrandiz E and Klatzmann D High-resolution repertoire analysis of Tfr and Tfh cells reveals unexpectedly high diversities indicating a bystander activation of follicular T cells. PNAS. 2018 Sep 18;115(38):9604-9609 2 Janeway CA, Approaching the asymptote? Evolution and revolution in immunology. Cold SpringHarbor Symp Quantit Biol (1989) 54: 1-13].

Funding:

Title: Germline immunoglobulin heavy chain locus variation and susceptibility to group A streptococcal disease

Presenter: Tom Parks
London School of Hygiene & Tropical Medicine
University of Oxford

Abstract: Globally the diseases associated with the bacterial pathogen group A streptococcus are an important cause of death and disability. There is currently no licence vaccine against the causative bacteria and disease control is therefore based on public health measures that are difficult to implement in low resource settings. Foremost among these diseases is rheumatic heart disease a post-infective autoimmune process that leads to scarring of the heart valves.

Our work in the Pacific region identified a novel susceptibility signal for rheumatic heart disease located in the immunoglobulin heavy chain (IGH). Initial fine-mapping attributed the signal to the *IGHV4-61**02 allele, each copy of the allele associated with a 1.4-fold increased risk of disease. In the talk I will describe follow-up analyses of the IGH locus in a new case-control study set in Northern India as well as case-control dataset derived from the UK Biobank study. In particular, I will illustrate the challenges of delineating causal variation in the IGH locus, highlighting the need for IGH-focused genomic resources from diverse populations.

Finally, I will show data indicating that IGH variation may also alter susceptibility to other group A streptococcal diseases potentially including both shared and opposing effects.

Funding:
National Institute for Health Research
British Heart Foundation
British Medical Association
Medical Research Council UK

Title: Challenges in motif discovery and comparative analysis of large T-cell receptor repertoire datasets

Presenter: Mikhail Shugay [1,2,3]
[1]Center of Life Sciences, Skolkovo Institute of Science and Technology, Moscow, Russia
[2]Immunosequencing Algorithms Group, Institute of Bioorganic Chemistry RAS, Moscow, Russia
[3]Department of Molecular Technologies, Pirogov Russian National Research Medical University, Moscow, Russia

Abstract: The V(D)J rearrangement process generates an extremely diverse repertoire of T-cell receptor (TCR) sequences that is designed to protect the organism against various pathogen encounters throughout its lifetime. The structure of the TCR repertoire is further shaped by antigen-driven selection and requires ultra-deep sequencing for proper quantification and analysis. In this talk I will outline currently available bioinformatic methods for background subtraction and motif discovery that can run on large datasets. Such methods can infer sets of TCR sequences of interest that are usually hidden by the noise and complexity of TCR repertoire structure. I will next discuss technical problems in comparative analysis of TCR repertoires that arise from differences in structure of TCR repertoires specific to distinct epitopes, MHC restriction of antigen-specific response and a large imprint of common pathogen infection history in individuals. Finally, I will share my thoughts on data reduction and optimal representation of AIRR-seq data that can both speed up and increase the robustness of sequence search and comparative analysis.

Funding: This work was supported by Russian Science Foundation grant No 17-15-01495

Title: Commonality despite exceptional diversity in the baseline human antibody repertoire

Presenter: Bryan Briney
Bryan Briney[1,2,3,4,5], Anne Inderbitzin[6], Collin Joyce[1,2,4], Dennis R. Burton[1,2,4,5,7]
[1]Department of Immunology and Microbiology, The Scripps Research Institute, La Jolla, CA, USA.
[2]Center for HIV/AIDS Vaccine Immunology and Immunogen Discovery, The Scripps Research Institute, La Jolla, CA, USA.
[3]Center for Viral Systems Biology, The Scripps Research Institute, La Jolla, CA, USA.
[4]IAVI Neutralizing Antibody Center, The Scripps Research Institute, La Jolla, CA, USA.
[5]Human Vaccines Project, New York, NY, USA.
[6]Division of Infectious Diseases and Hospital Epidemiology, University Hospital Zurich, University of Zurich, Zurich, Switzerland.
[7]Ragon Institute of MGH, MIT and Harvard, Cambridge, MA, USA.

Abstract: In principle, humans can produce an antibody response to any non- self-antigen molecule in the appropriate context. This flexibility is achieved by the presence of a large repertoire of naive antibodies, the diversity of which is expanded by somatic hypermutation following antigen exposure. The diversity of the naive antibody repertoire in humans is estimated to be at least 10^{12} unique antibodies. Because the number of peripheral blood B cells in a healthy adult human is on the order of $5x10^9$, the circulating B cell population samples only a small fraction of this diversity. Full-scale analyses of human antibody repertoires have been prohibitively difficult, primarily owing to their massive size. The amount of information encoded by all of the rearranged antibody and T cell receptor genes in one person—the 'genome' of the adaptive immune system—exceeds the size of the human genome by more than four orders of magnitude. Furthermore, because much of the B lymphocyte population is localized in organs or tissues that cannot be comprehensively sampled from living subjects, human repertoire studies have focused on circulating B cells3. We examined the circulating B cell populations of ten human subjects and present what is, at the time of publication, the largest single collection of adaptive immune receptor sequences described to date, comprising almost 3 billion antibody heavy-chain sequences. This dataset enables genetic study of the baseline human antibody repertoire at an unprecedented depth and granularity, which reveals largely unique repertoires for each individual studied, a subpopulation of universally shared antibody clonotypes, and an exceptional overall diversity of the antibody repertoire.

Funding: This work was supported by the National Institute of Allergy and Infectious Diseases (Center for HIV/AIDS Vaccine Immunology and Immunogen Discovery, UM1AI100663 (D.R.B.); Center for Viral Systems Biology, U19AI135995 (B.B.)), the International AIDS Vaccine Initiative (IAVI) through the Neutralizing Antibody Consortium SFP1849 (D.R.B.), and the Ragon Institute of MGH, MIT and Harvard (D.R.B.).

Title: Pan Immunome Initiative and applications in health and disease

Presenter: Xiao Liu
PhD
BGI-Shenzhen

Abstract: We have launched Pan Immunome Initiative (PII) to create an integrated map of T and B cell receptors in health and diseases. To accomplish this mission, we used synthetic receptor sequences to optimize an amplification system, and created a bioinformatic software to analyze the data (IMonitor). As the receptor germline gene references are pivotal to align the rearranged sequences, and far less than complete currently, we have developed a germline gene inference pipeline to predict the germline sequences from the rearranged data (IMpre), and discovered substantial novel germline genes and alleles in rhesus macaque as the important pre-clinical animal model. To facilitate data storage and sharing, we have designed and built PIRD (Pan Immune Repertoire Database), to store and share the latest data from PII, and help users to upload, analyze, compare and visualize their own data with the data in PIRD. In this presentation, I will also introduce some early results from PII and applications in various immunological settings, including leukemia and solid cancer, as well as autoimmune diseases, with the emphasis on accurate diagnosis, molecular classification and accurate prediction of prognosis.

Funding:

Title: Lessons from the analysis of the immune response to *P. falciparum:* the power of clonal selection

Presenter: Antonio Lanzavecchia
Institute for Research in Biomedicine, Università della Svizzera italiana, Bellinzona, Switzerland

Abstract: We use cell culture-based high-throughput methods to interrogate human memory B cell and plasma cell repertoires and isolate antibodies selected on the basis of their neutralizing potency and breadth. Recently, we focused our analysis on the antibody response to the blood stage and pre-erythrocytic stage of *P. falciparum* parasites. We discovered that up to 10% of malaria infected individuals produce a new type of antibodies that contain a large templated insertion that comprises the entire extracellular domain of LAIR1, a collagen binding inhibitory receptor encoded on chromosome 19. In the clones producing LAIR1-containing antibodies the insertions are found either between V and DJ segments or in the switch region, leading to the positioning of the LAIR1 domain on the tip of HCDR3 or in the VH-CH1 elbow. The inserted LAIR1 domain is both necessary and sufficient for binding to infected erythrocytes and somatic mutations abolish collagen binding and modulate binding activity to the parasite antigens, which we identified as distinct RIFINs. Templated insertions derived from transcribed genes are frequently found in memory B cells of healthy individuals, suggesting that this mechanism represents a new and general mode of antibody diversification.

In another study, we analysed the antibody response of African individuals that were immunized by repeated injection of irradiated sporozoites and found to be protected from a challenge with infectious sporozoites. All isolated IgG antibodies bound to the circumsporozoite protein (CSP) and recognized distinct epitopes in its N terminus, NANP-repeat region, and C terminus. Strikingly, the most effective antibodies, as determined in a humanized mouse model, bound not only to the repeat region, but also to a minimal peptide at the PfCSP N-terminal junction that is not present in the RTS,S vaccine. These dual-specific antibodies were isolated from different donors and were encoded by VH3-30 alleles that encode tryptophan or arginine at position 52. Using structural and mutational data, we identified the elements required for germline recognition and affinity maturation. These potent neutralizing antibodies provide relevant information for lineage-targeted vaccine design and immunization strategies.

References: Tan et al, *Nature* 529:105 (2016). Pieper et al, *Nature* 549:597 (2017).

Tan et al, *Nat. Med* 24:401 (2018)

Funding:

Title: Mapping B cell receptor repertoires in immune-mediated diseases

Presenter: Rachael Bashford Rogers
RJ Bashford-Rogers[1,2], L Bergamaschi[2], DC Pombal[2], F Mescia[2], EF McKinney[2], JC Lee[2], TDC Thomas[2], SM Flint[1], P Kellam[3], DRW Jayne[2], PA Lyons[2], KGC Smith[2]
[1]Wellcome Trust Centre for Human Genetics, University of Oxford
[2]Department of Medicine, University of Cambridge
[3]Department of Medicine, Division of Infectious Diseases, Imperial College London.

Abstract: B cells are important in the pathogenesis of many, and perhaps all, immune-mediated diseases. Here we describe novel methods that characterise the human adaptive immune response by high-throughput sequencing of B-cell receptors.
Using these methods, we next examined the isotype-specific B cell receptor repertoire (BCR) in a range of immune-mediated diseases; systemic lupus erythematosus, ANCA-associated vasculitis, Crohn's disease, Behçet's Disease, eosinophilic granulomatosis with polyangiitis EGPA and IgA vasculitis and B cell lymphoid malignancies. This revealed unexpected differences in isotype, clonality and IGHV gene usage between diseases. An IgA-dominated increased clonality in SLE and Crohn's disease, together with skewed IGHV gene usage in other diseases, suggested a microbial contribution to disease pathogenesis. Different immunosuppressive treatment had specific and very different impacts on the repertoire – for example the B cells persisting after rituximab therapy were predominately isotype-switched and clonally expanded, the inverse of those persisting after mycophenolate mofetil. A comparative analysis of the BCR repertoire thus reveals an unexpectedly complex B cell architecture, providing a platform for a better understanding of pathological mechanisms and treatment responses in immune-mediated disease.

Funding: This work was supported by the Wellcome Trust (grant WT106068AIA and 083650/Z/07/Z), UK National Institute of Health Research, Cambridge Biomedical Research Centre and UK Medical Research Council.

Title: Mining the tumor-associated T cell repertoire for TCR therapeutics

Presenter: Robert A. Holt
British Columbia Cancer Agency
University of British Columbia
Simon Fraser University

Abstract: Effective anti-cancer immunity depends on TCR-mediated molecular recognition of tumor antigens and it is well recognized T cell infiltration of solid tumors is associated with better patient outcomes. In an integrated multi-region analysis of metastatic high grade serous ovarian cancer we have found that samples with high epithelial T cell density have low tumor clone diversity and that T cell repertoire similarity is associated with malignant clone composition, consistent with T cell clonotypes spatially tracking tumor clones in these patients. In some cases tumor clones have escaped immune clearance by somatic genomic loss of HLA haplotypes. Although immunogenomic profiling studies of this type can provide important insights into the nature of the anti-cancer immune response they do not reveal the relationship between specific T cell clonotypes and their cognate neoantigens directly. For this, in a separate study, we identified the neoantigen target of the T cell response in a high grade serous ovarian cancer patient and we then followed the tumor and T cell dynamics in this patient over time. Reactive T cells were present at very low abundance at the time of initial diagnosis, became elevated upon progression, and were then lost upon final cancer recurrence of the cancer, without loss of the neoantigen or HLA.

In principle, tumor-reactive TCRs reconstituted in effector T cells could offer a new therapeutic intervention for patients who have not initiated or who have failed to sustain effective anti-cancer immunity. Sporadic mutations will be difficult to approach in this manner for practical reasons, but hotspot mutations that yield neoantigens presented by common HLAs are attractive targets. We are currently screening immune repertoires to generate panels of recombinant TCRs for this purpose and have a series of recombinant TCRs against HLA-A*02:01 restricted KRAS codon 12 hotspot mutations in pre-clinical development.

Funding:

Title: TCR repertoire profiling in non-model species.

Presenter: Olga Britanova
Izraelson M.[1], Metsger M.[4], Davydov A.N.[4], Dronina M.A.[5], Miskevich D.A.[5], Mamedov I.Z.[1,2,3], Shugay M.[1,6,2,3], Staroverov D.B.[1], Kondratyuk E.Y.[7], Shams I.[5*],.[1,**] and Chudakov D.M.[1,6,2,3,4*,**], Britanova O.V[1]
[1]Shemyakin and Ovchinnikov Institute of Bioorganic Chemistry, Moscow 117997, Russia.
[2]Privolzhsky Research Medical University, Nizhny Novgorod, Russia.
[3]Pirogov Russian National Research Medical University, Moscow 117997, Russia.
[4]Central European Institute of Technology, Brno 60177, Czech Republic
[5]Institute of Evolution & Department of Evolutionary and Environmental Biology, University of Haifa, Haifa 3498838, Israel.
[6]Center for Data-Intensive Biomedicine and Biotechnology, Skolkovo Institute of Science and Technology, Moscow 143028, Russia.

Abstract: Modern research on aging and longevity is focused on inflammaging - a concept that considers aging primarily as a combination of systemic inflammatory processes causing various diseases, including cancer. This approach brings to the fore the role of immunological aging in organismal aging. From the perspective of adaptive immunity, immunological aging is expressed as changes in the functional properties and reduction of the antigen-specific diversity of T-lymphocytes. The subterranean blind mole rat (*Spalax*) is an exceptionally long-lived rodent that seems to present a promising model in the study of immune aging and its role in inflammaging. To date, there is no data on the relationship between aging and adaptive immunity in long-lived rodents, such as a naked mole rat or a blind mole rat.

We have developed a set of tools necessary for investigation of Spalax adaptive immunity: unbiased 5'RACE UMI protocol to generate Spalax TCR alpha and beta repertoires, a complete gene reference for the Spalax TCR repertoire-extracting MiXCR software, and a high-quality transcriptome reference. We have shown for the first time that the diversity of T-cell repertoire does not decrease with age, even in very old 14-18 year-old blind mole rats. Spalax, as well as the other rodent species, naked mole rats, are long-lived mammals that are resistant to hypoxia and occurrence of tumors. The phenomenon we discovered may indicate a unique organization of T-cell immunity, which presumably can cause a significant longevity of mole rats.

Recent studies, including non-model mammalian organisms, have shown that there is no universal mechanism for longevity. At the same time, the study of non-model animals reveals new genes, specific mutations in already known genes or even mechanisms that ensure longevity and resistance to diseases.

Funding: This work was supported by the Russian Science Foundation project №16-15-00149.

Title: BCR affinity differentially regulates colonization of the subepithelial dome and infiltration into germinal centers within Peyer's patches

Presenter: Ziv Shulman
Adi Biram[1], Anneli Strömberg[2], Eitan Winter[3], Liat Stoler-Barak[1], Ran Salomon[1], Yoseph Addadi[4], Rony Dahan[1], Gur Yaari[3], Mats Bemark[2] and Ziv Shulman[1]*
[1] Department of Immunology,
[2] Department of Microbiology and Immunology, Institute of Biomedicine, University of Gothenburg, Sweden.
[3] Faculty of Engineering, Bar Ilan University, Ramat Gan, Israel.
[4] Department of Life Science Core facilities, Weizmann Institute of Science, Rehovot, Israel.

Abstract: Gut-derived antigens trigger immunoglobulin A (IgA) immune responses that are initiated by cognate B cells in the Peyer's patch (PP). These cells colonize the subepithelial domes (SEDs) of the PP, and subsequently infiltrate into pre-existing germinal centers (GCs). Here, we defined the pre-GC events and the microanatomical site at which affinity-based B cell selection occurred in PPs. Using whole-organ imaging, we showed that the affinity of the B cell antigen receptor (BCR) regulated infiltration of antigen-specific B cells into GCs, but not clonal competition in the SED. Follicular helper-like T cells resided in the SED and promoted its B cell colonization, independently of the magnitude of BCR affinity. Imaging and immunoglobulin sequencing indicated that selective clonal-expansion ensued during infiltration into GCs. Thus, in PPs, in contrast to draining lymph nodes and spleen, T cells predominantly promoted expansion of B cells without clonal selection during pre-GC events. These findings have major implications for the design of oral vaccines.

Funding:

Title: A primer on patient-oriented research

Presenter: Tania Bubela
Simon Fraser University

Abstract:
Patient-oriented research is research that is done '*with*' or '*by*' patients or the public, rather than done 'for,' 'on,' or 'to' people. A key component is ensuring knowledge synthesis and transfer to clinical settings. This talk will discuss how a patient-oriented research approach can be designed to facilitate clinical research and improve health outcomes.

Funding:

Title: Dynamic consent: individual oversight and control of data and samples

Presenter: Harriet Teare
H Teare[1], M Prictor[2], M Haas[3], J Kaye[4]
[1]Centre for Health, Law and Emerging Technologies, University of Oxford
[2]Melbourne Law School, University of Melbourne
[3]Australian Genomics Health Alliance

Abstract: When participants donate samples to biomedical research it is often difficult to fully predict how these samples and associated data might be used. This creates challenges for the informed consent process, if participants are unclear what future research might entail when they are volunteering to take part. Dynamic consent has been developed to address these challenges by initiating an ongoing conversation between researchers and participants, allowing consent decisions to be reviewed and updated over time.

1 It provides an electronic portal through which researchers can communicate project progress, and engage and interact with participants, while extending the consent discussion beyond a single point in time.

2 Dynamic consent is being trialled in different research settings, including the RUDY study (UK),3 a rare disease network gathering experiential data directly from participants, and the Australian Genomics Health Alliance, a national research collaboration which aims to provide evidence for the equitable, effective and sustainable delivery of genomic medicine in healthcare.

By evaluating implementation of dynamic consent, we can better understand how it alters the participant experience and whether it improves peoples understanding of research compared with traditional forms of consent. This will help to establish whether it provides an appropriate vehicle to address some of the legal and ethical issues involved in sharing and using data.

Funding:

Title: We are relational: The challenge of Indigenous collective rights in bio-ethics

Presenter: Namaste Marsden
A/Director, Research and Knowledge Exchange
First Nations Health Authority

Abstract:
Indigenous peoples' cultures, languages and world views are diverse, however there are many aspects of Indigenous ways of knowing and being that are seen across cultures. For example, spirituality, connection to land, collective and tribal identity, and relational connections to all beings and between humans are important aspect of many Indigenous ways of knowing and being. This presentation will provide a brief overview of the inter-face of biological sampling in health research with Indigenous peoples in Canada and it's interconnection with the colonialist imperative. Through use of examples and stories of projects that have violated Indigenous peoples rights, and, looking at contemporary projects that are co-led by Indigenous people with Indigenous governance models, themes of ethics will be illustrated that demonstrate good engagement and healthy partnerships with communities, nations and their members. The First Nations Health Authority (FNHA) has a mandate and transformational role in health service delivery for First Nations peoples in British Columbia, Canada. This is exemplified by FNHA's governance model, the Seven Directives from First Nations leadership and communities, leadership in systemic changes to ensure cultural safety and humility in trauma-informed health care, and in research. Creation of knowledge grounded in First Nations perspectives on health and wellness are also transforming FNHA's partnerships in research, ethics and data governance. Collective rights of First Nations and Indigenous peoples are affirmed in agreements, partnerships, constitutional law, and in the United Declaration of Rights of Indigenous Peoples. Nations, communities and organized groups will be emphasized in research partnerships that are respectful of collective rights. Identifying ways of approaching research as relationship and seeing bio-ethics within the cultural context of collective rights and values of relational being will be encouraged.

Funding:

Title: Applied Repertoire Analysis - Computational ImmunoEngineering of Therapeutic Antibodies and TCRs

Presenter: Jacob Glanville
Distributed Bio

Abstract: Modern medicine increasingly utilizes adaptive immune receptors as biologic therapeutics. Over the past decade, advances in the high-throughput sequence analysis of adaptive repertoires has enabled multiple practical advances into computationally-guided engineering of biologics. Here we review specific example applications in antibody engineering, nanobody engineering, and TCR engineering.

Funding:

Title: The Adaptive-Microsoft TCR-Antigen Map Project

Presenter: Jonathan Carlson
Microsoft

Abstract: In January of 2018, Adaptive Biotechnologies and Microsoft launched the Antigen Map Project, the aim of which is to learn to map T-cell receptors to antigens as a basis for new tools to help diagnose and treat cancers, pathogens, and autoimmune disorders. To this end, we have initiated two data generation efforts:

the first is to immunosequence 25,000 TCR (beta) repertoires for individuals affected by one of several initial target diseases;

the second is to sequence millions of TCRs that are specific to thousands of antigens using multiplexed assays. In this talk, I will describe the vision and aims of the partnership and some early results.

Funding:

Title: Bridging the gap: manage, analyze, visualize and interpret repertoire sequencing data using the IGX platform

Presenter: Nicola Bonzanni
ENPICOM B.V., ’s-Hertogenbosch, The Netherlands

Abstract: Repertoire analysis is rapidly gaining traction as an essential tool for immunotherapy development, patient stratification before start of treatment and monitoring of patients on treatment. Repertoire analysis combines the unique sequences of T- and B-cell receptors and the power of high-throughput sequencing to identify specific T- and B-cell clones. However, the implementation, analysis and interpretation of repertoire data in research and clinical workflows is challenging and requires specialized expertise across various disciplines like immunology, bioinformatics and software engineering.

To date, no integrated, end-to-end solution exists. To address this unmet need, we developed the ImmunoGenomiX (‘IGX’) platform: a new versatile platform for the analysis and interpretation of high-throughput repertoire sequencing data. IGX is an end-to-end platform, enabling users to perform state-of-the-art analyses, starting from raw sequencing files, up to the generation of report- or publicationready figures. The powerful, tag-based, annotation system of the IGX platform allows users to extensively annotate repertoire sequencing data with structured metadata at great level of granularity, including annotations for single clones such as affinity or avidity.

Sequencing data, as well as metadata, can be used to perform powerful searches, create collections of clones, or subset and filter data for follow-up analysis. The IGX platform features a userfriendly interface that helps researchers and clinicians bridge the gap between the staggering amounts of generated sequencing data and its interpretation, without requiring any computational or bioinformatic expertise.

Funding:

AIRR-seq Tool Demonstrations

AIRR Community Meeting IV – AIRR-seq Tool Demonstrations		
	Presenter	**Abstract**
1.	Nicola Bonzanni, Enpicom B.V.	**ImmunoGenomiX** Intuitively manage, store, analyze, and visualize repertoire sequencing data using the ImmunoGenomiX ('IGX') platform Repertoire analysis is rapidly gaining traction as a tool to monitor patients receiving immunotherapy treatment. Repertoire analysis can be used for, amongst others, MRD detection in lymphoid cancers and monitoring specific CTL responses in peripheral blood. However, the analysis, interpretation, and management of repertoire data in research and clinical workflows is challenging. The ImmunoGenomiX ('IGX') platform is a new versatile platform to manage, store, analyze, visualize and interpret high-throughput repertoire sequencing data. Through a user-friendly graphical interface, users can execute innovative end-to-end analyses, starting from raw sequencing (fastq) files, up to the production of publication-ready figures containing highly accurate results. In addition to its analytical capabilities, state-of-the-art clone management is an integral part of the IGX platform: samples, repertoires, as well as single clones can be annotated with user-defined tags to include metadata into IGX analysis. Using this powerful tag- and text-based annotation system, users can quickly search, retrieve, subsets, and categorize data for follow-up analysis. In this tutorial, we will demonstrate how the analytical capabilities of the IGX platform, paired with its intuitive and powerful data management layer, can facilitate the analysis, interpretation, and management of repertoire sequencing data.
2.	Bryan Briney, Scripps Research	**ab[x]** Massively scalable genetic analysis of antibody repertoires" With technical breakthroughs in the throughput and read-length of next-generation sequencing platforms, antibody repertoire sequencing is becoming an increasingly important tool for detailed characterization of the immune response. There is a need for open, scalable software for the genetic analysis of repertoire-scale antibody sequence data. To address this gap, we have developed the ab[x] package of software tools. There are three core components of the ab[x] toolkit, all of which are freely available: abcloud (github.com/briney/abcloud) for deployment and management of computational resources on Amazon's Elastic Compute Cloud; abstar (github.com/briney/abstar) for pre-processing, germline gene assignment and primary annotation of antibody sequence data; and abutils (github.com/briney/abutils), which provides utilities for interactive downstream analysis of antibody repertoire data.
3.	Brian Corrie, Simon Fraser University	**iReceptor** iReceptor – A platform for querying and analyzing antibody/B-cell and T-cell receptor repertoire data across federated repositories The iReceptor Platform (www.ireceptor.org) is a distributed data management system and scientific gateway for mining "Next Generation" sequence (NGS) data from the

		adaptive immune (antibody/B-cell or T-cell) receptor repertoire. The main goal of iReceptor is to connect a distributed network of AIRR-seq repositories (the AIRR Data Commons), allowing queries across multiple projects, labs, and institutions. Currently, the iReceptor Platform integrates two international AIRR-seq repositories (VDJServer (www.vdjserver.org) and the iReceptor Public Archive), providing researchers access to over 1 Billion annotated sequences across 805 samples, 16 research labs and 18 studies. Integrating these important AIRR-seq repositories will result in improvements to the design of vaccines, therapeutic antibodies and cancer immunotherapies. iReceptor provides a technology platform that lowers the barrier to immune genetics researchers who need to federate large, distributed, immune genetics data repositories in order to answer complex questions about the immune response. The iReceptor Gateway (http://gateway.ireceptor.org) is a scientific web platform which enables the exploration, analysis, and downloading of curated AIRR-seq across a number of distributed data repositories. In this tool demonstration we will demonstrate the use of the iReceptor Scientific Gateway to find, explore, and federate AIRR-seq data across the AIRR Data Commons. Funding: CANARIE, Canada Foundation for Innovation, BC Knowledge Development Fund, Canadian Institute for Health Research, EU Horizon 2020 (No. 825821).
4.	Kenneth Hoehn, Yale University	**IgPhyML** Phylogenetic analysis of B cell repertoires with IgPhyML Affinity maturation in B cells is an evolutionary process of descent from a germline sequence through somatic hypermutation (SHM), and clonal selection. Because of the similarity between affinity maturation and evolution in natural populations, phylogenetics has a long history of use in studying B cell clonal lineages. However, certain features of affinity maturation violate important assumptions in standard models. Further, most phylogenetic models characterize single lineages, while B cell repertoires often contain thousands of lineages. We recently developed a phylogenetic framework in the program IgPhyML (https://bitbucket.org/kbhoehn/igphyml) that addresses both of these issues by introducing a substitution model that incorporates SHM biases and shares information among lineages within a B cell repertoire to precisely estimate model parameters. This framework has been integrated into the Immcantation suite (http://immcantation.org), a widely-used start-to-finish BCR repertoire analysis toolkit. In this tutorial, demonstrate how to use IgPhyML to build and display B cell clonal lineage trees, as well as quantify signatures of natural selection and SHM acting on B cell repertoires.
5.	Steven Kleinstein, Yale School of Medicine	**Immcantation** The Immcantation Framework for analysis of AIRR-seq data Next-generation sequencing (NGS) technologies have revolutionized our ability to carry out large-scale adaptive immune receptor repertoire sequencing (AIRR-seq) experiments. AIRRseq is increasingly being applied to profile B cell receptor (BCR) repertoires and gain insights into immune responses in healthy individuals and those with a range of diseases. As NGS technologies improve, these experiments are producing ever larger datasets, with tens- to hundreds-of-millions of BCR sequences. Although promising, repertoire-scale data present fundamental challenges for analysis requiring the development of new techniques and the rethinking of existing methods that are not scalable to the large number of sequences being generated [1]. To address the challenges of AIRR-seq analysis, we make available the Immcantation tool suite (http://immcantation.org), a start-to-finish BCR repertoire analysis ecosystem. Immcantation includes pRESTO, which handles all stages of sequence processing from raw reads up to the task of V(D)J gene assignment, including support for unique molecule barcodes (UMIs). To facilitate advanced analysis of repertoire properties, Immcantation also includes methods for: novel V gene allele detection (TIgGER), subject-specific germline genotype identification, B

		cell clone assignment (Change-O and SCOPer), lineage tree construction and analysis (IgPhyML), somatic mutation profiling and selection analysis (BASELINe). These methods and others are fully integrated through a standardized file format. Immcantation can start from raw data or read the output of common V(D)J assignment tools, including IMGT/HighV-QUEST and IgBLAST. It also supports MiAIRR, the AIRR Community data standard, and includes tools to facilitate MiAIRR-compliant submissions to NCBI repositories. To facilitate use on computing clusters and promote reproducibility, we have encapsulated the entire Immcantation framework, along with a set of accessory scripts, V(D)J germline databases, and applicable third-party tools into a set of meta-versioned Docker containers that are compatible with Singularity. Standalone software and documentation are available online at: http://immcantation.org. Containers are available through docker hub at https://hub.docker.com/r/kleinstein/immcantation. 1. Yaari G, Kleinstein SH. Practical guidelines for B-cell receptor repertoire sequencing analysis. Genome Med. 2015 Nov 20;7:121. doi: 10.1186/s13073-015-0243-2. Funding: National Institutes of Health grant R01AI104739
6.	William Lees, Birkbeck College, University of London	**OGRDB – the Open Germline Receptor Database** Even for well-studied species such as humans and mice, our knowledge of immune receptor genes and their allelic variation is incomplete. In recent years, methods have been developed to infer novel genes and alleles from NGS repertoires. In 2017, the AIRR Community established the Inferred Allele Review Committee (IARC) to evaluate inferred alleles for inclusion in relevant germline databases. IARC has worked, together with colleagues at IMGT and the US National Institute of Health, to establish a systematic submission and review process. OGRDB (https://www.ogrdb.airr-community.org) was created and designed to support that process and provides a real-time record of affirmed sequences. The presentation will demonstrate the novel allele submission and review process, as well as the use of the tool to browse and download affirmed sequences.
7.	Vadim Nazarov, National Research University	**Immunarch** Bioinformatics software problems and how to make immune data analysis effortless with immunarch There has been a sharp increase in the number of published works in immunomics, however a lack of common data analysis tools and approaches renders challenges for research. Numerous research pipelines have spawned an enormous amount of annotation formats, yet not a single tool can parse them all. As a consequence, there has been developed a vast collection of them. Coupled with the lack of regular updates this slows research drastically, urging scientists to check each and every package. Finally, most of the software does not provide them with publication-ready plots. To overcome these challenges researchers require fast, practical and interpretable software. In this software demo session, we would like to introduce immunarch R package to bioinformaticians and immunobiologists with at least basic R programming understanding. Immunarch is an improved version of tcR package that gained popularity among immunomics researchers. In the course of the session we will cover all three steps in AIRR data analysis: file parsing, data analysis and publication-ready plots creating. Eventually, we will recap the major hurdles that are being encountered during bioinformatics software development and we will present our solutions.
8.	Duncan Ralph, Fred Hutchinson Cancer Research Centre	**Partis: A tool for accurate and efficient annotation of germline and clonal family inference** Partis is a tool for accurate, efficient annotation and germline and clonal family inference on deep sequencing B cell receptor data.

		It also contains a detailed simulation engine. We will begin with a short demonstration of the main modes of operation, covering recommended options for annotation, partitioning, and simulation, with a particular emphasis on the newer automatic per-sample germline inference. If you are planning to try out partis during the session, please install beforehand (instructions at
9.	Mikael Salson, Universite de Lille - VidjlNet	**Vidjil: an open-source platform for interactive repertoire analysis** Vidjil is an open-source platform for the interactive analysis of high-throughput sequencing data from lymphocyte recombinations, licensed under the GPLv3 open-source license. Source code, binaries and a public web server are available at http://www.vidjil.org. At the heart of the platform, Vidjil-algo, implemented in C++, processes high-throughput sequencing data to extract V(D)J junctions and gather them into clones, from both immunoglobulins and T-cell receptors, as well as some incomplete or uncommon rearrangements. It is extremely fast because, in the first phase, no alignment is performed with database germline sequences. The algorithm works on recombined reads coming from either amplicon-based or capture-based deep sequencing strategy. It exports data both in a native .json format and in the AIRR format. The web application, written in Python and Javascript, contains an interactive visualization for the analysis of clonotypes along the time that can be linked to a sample, experiment and patient database. Using the Vidjil web application consists of four steps: 1. uploading raw sequence files; 2. running Vidjil-algo or other RepSeq analysis software; 3. visualizing the results; 4. annotating the results and saving them for future use. The interactive visualization allows the user to investigate the clonotypes on several dimensions (V, J genes, number of insertions, deletions, read lengths, sequence similarity, CDR3. . .). DNA sequences from clonotypes can be aligned to identify mutations and can also be directly submitted to other software for further inspection (IMGT/V-QUEST, IgBlast, Blast). For the end-user, the Vidjil web application needs no specific installation and just requires a connection and a modern web browser. Vidjil is used by labs in hematology or immunology for research and clinical applications. More than 40 labs in Europe and in the world regularly use Vidjil through the web application, including hospitals in daily clinical practice on ALL and CLL diagnosis samples

Pipeline & Data Repository Tutorials

AIRR Community Meeting IV – Pipeline & Data Repository Tutorials		
	Presenter	**Abstract**
1.	Bryan Briney, Scripps Research	**AbCloud** Massively scalable genetic analysis of antibody repertoires With technical breakthroughs in the throughput and read-length of next-generation sequencing platforms, antibody repertoire sequencing is becoming an increasingly important tool for detailed characterization of the immune response. There is a need for open, scalable software for the genetic analysis of repertoire-scale antibody sequence data. To address this gap, we have developed the ab[x] package of software tools. There are three core components of the ab[x] toolkit, all of which are freely available: abcloud (github.com/briney/abcloud) for deployment and management of computational resources on Amazon's Elastic Compute Cloud; abstar (github.com/briney/abstar) for pre-processing, germline gene assignment and primary annotation of antibody sequence data; and abutils (github.com/briney/abutils), which provides utilities for interactive downstream analysis of antibody repertoire data.
2.	Nicola Bonzanni, ENPICOM B.V.	**IGX** Intuitively manage, store, analyze, and visualize repertoire sequencing data using the ImmunoGenomiX ('IGX') platform Repertoire analysis is rapidly gaining traction as a tool to monitor patients receiving immunotherapy treatment. Repertoire analysis can be used for, amongst others, MRD detection in lymphoid cancers and monitoring specific CTL responses in peripheral blood. However, the analysis, interpretation, and management of repertoire data in research and clinical workflows is challenging. The ImmunoGenomiX ('IGX') platform is a new versatile platform to manage, store, analyze, visualize and interpret high-throughput repertoire sequencing data. Through a user-friendly graphical interface, users can execute innovative end-to-end analyses, starting from raw sequencing (fastq) files, up to the production of publication-ready figures containing highly accurate results. In addition to its analytical capabilities, state-of-the-art clone management is an integral part of the IGX platform: samples, repertoires, as well as single clones can be annotated with user-defined tags to include metadata into IGX analysis. Using this powerful tag- and text-based annotation system, users can quickly search, retrieve, subsets, and categorize data for follow-up analysis. In this tutorial, we will demonstrate how the analytical capabilities of the IGX platform, paired with its intuitive and powerful data management layer, can facilitate the analysis, interpretation, and management of repertoire sequencing data.
3.	Scott Christley, UT Southwestern Medical Center	**VDJServer** A Cloud-Based Analysis Portal and Data Commons for Immune Repertoire Sequences and Rearrangements VDJServer (vdjserver.org) is a cloud-based analysis portal for immune repertoire sequence data that provides access to a suite of tools for a complete analysis workflow, including modules for pre-processing and quality control of sequence

		reads, V(D)J gene segment assignment, repertoire characterization, and repertoire comparison, and provides sophisticated visualizations for exploratory analysis. It is accessible through a standard web browser via a graphical user interface designed for use by immunologists, clinicians, and bioinformatics researchers. VDJServer provides a data commons for public sharing of repertoire sequencing data, as well as private sharing of data between users. VDJServer utilizes the high-performance computing resources at the Texas Advanced Computing Center (www.tacc.utexas.edu) and is a free, publicly available, and open-source licensed resource.
4.	Brian Corrie, Simon Fraser University	**iReceptor** A platform for querying and analyzing antibody/B-cell and T-cell receptor repertoire data across federated repositories The iReceptor Platform (www.ireceptor.org) is a distributed data management system and scientific gateway for mining “Next Generation” sequence (NGS) data from the adaptive immune (antibody/B-cell or T-cell) receptor repertoire. The main goal of iReceptor is to connect a distributed network of AIRR-seq repositories (the AIRR Data Commons), allowing queries across multiple projects, labs, and institutions. Currently, the iReceptor Platform integrates two international AIRR-seq repositories (VDJServer (www.vdjserver.org) and the iReceptor Public Archive), providing researchers access to over 1 Billion annotated sequences across 805 samples, 16 research labs and 18 studies. Integrating these important AIRR-seq repositories will result in improvements to the design of vaccines, therapeutic antibodies and cancer immunotherapies. iReceptor provides a technology platform that lowers the barrier to immune genetics researchers who need to federate large, distributed, immune genetics data repositories in order to answer complex questions about the immune response. The iReceptor Gateway (http://gateway.ireceptor.org) is a scientific web platform which enables the exploration, analysis, and downloading of curated AIRR-seq across a number of distributed data repositories. During this tutorial we will walk users through the two possible workflows on the iReceptor Scientific Gateway, the Sequence Annotation “Quick Search” and the Metadata/Sequence Annotation “Exploration” work flows. During these walk throughs we will explore the range of MiAIRR metadata and sequence annotation attributes on which users can search. In addition, we will discuss and demonstrate the iReceptor web API for querying repositories in the AIRR Data Commons.
5.	Jacob Glanville, Distributed Bio	**AbGenesis** A cloud based immune repertoire analysis platform AbGenesis is the world's leading platform for antibody and immune repertoire sequence analysis. In a simple to use environment users can interrogate sequence data and derive meaningful and decision oriented insights from their data. From identifying candidate clones, analyzing CDR diversity and guiding engineering decisions AbGenesis enables all facets of antibody discovery. Common uses include analyzing library selections, hybridoma and immunization data for discovery or monitor population immune responses. The platform accepts read data from any common platform including Sanger, MiSeq, HiSeq, 454, Ion Torrent, and public patent and structural databases. By using profile Hidden Markov Models as the core immunoglobulin-recognizing algorithm, AbGenesis is able to analyze human, mouse, rat, camelid, shark, dog, cat, cow, horse, pig, rhesus, chicken, and multiple other model organisms or engineered immunoglobulin constructs without constraints of codon usage or species of origin. Built in 2012, it immediately attracted multiple clients and enabled us to become profitable in our first year of business without ever needing venture funding. AbGenesis is now used at over 35 institutions and biotechnology companies, including 7 of the 10 top pharma. Internally at Distributed

		Bio, AbGenesis acts synergistically to empower our antibody library design, antibody discovery, and vaccine optimization programs.
6.	Kenneth Hoehn, Yale University	**IgPhyML** Phylogenetic analysis of B cell repertoires with Immcantation and IgPhyML Affinity maturation in B cells is an evolutionary process of descent from a germline sequence through somatic hypermutation (SHM), and clonal selection. Because of the similarity between affinity maturation and evolution in natural populations, phylogenetics has a long history of use in studying B cell clonal lineages. However, certain features of affinity maturation violate important assumptions in standard models. Further, most phylogenetic models characterize single lineages, while B cell repertoires often contain thousands of lineages. We recently developed a phylogenetic framework in the program IgPhyML (https://bitbucket.org/kbhoehn/igphyml) that addresses both of these issues by introducing a substitution model that incorporates SHM biases and shares information among lineages within a B cell repertoire to precisely estimate model parameters. This framework has been integrated into the Immcantation suite (http://immcantation.org), a widely-used start-to-finish BCR repertoire analysis toolkit. Immcantation includes the Repertoire Sequencing Toolkit (pRESTO), which handles all stages of sequence processing from raw reads up to the task of V(D)J gene assignment, including specific support for unique molecule barcodes (UMIs). To facilitate advanced analysis of the resulting B cell repertoire properties, Immcantation also includes methods for: novel V gene allele detection (TIgGER), subject-specific germline genotype identification, B cell clone assignment (Change-O and SCOPer), lineage tree construction, somatic mutation profiling and selection analysis (BASELINe). In this tutorial, we will first demonstrate how the Immcantation framework can be used to identify B cell clonal clusters from Adaptive Immune Receptor Repertoire sequencing (AIRR-seq) data. We will then demonstrate how to use IgPhyML to build and display B cell clonal lineage trees, as well as quantify signatures of natural selection and SHM acting on B cell repertoires.
7.	William Lees, Birkbeck College, University of London	**OGRDB** Allelic variants and OGRDB – the Open Germline Receptor Database Even for well-studied species such as humans and mice, our knowledge of immune receptor genes and their allelic variation is incomplete. In recent years, methods have been developed to infer novel alleles from NGS repertoires. Alleles discovered via these methods, and the associated techniques used to determine a subject's genotype, can be used to improve and refine the annotation of NGS repertoires. The tutorial will introduce the principles and benefits of using genotyping and allele discovery in a sequencing pipeline. It will cover in depth the use of OGRDB to submit novel alleles for review, and to examine those previously published. The tutorial will cover: - Using genotypes in NGS immune repertoire sequencing – principles, benefits and tools - Using OGRDB and other sources to explore novel alleles - What makes a good inference? - Supporting inferences through the use of haplotyping - The IARC review process - Pre-requisites: submitting putative alleles to NCBI - Submitting putative alleles to OGRDB

		- How OGRDB supports the review process
8.	Mikhail Shugay Skolkovo Institute of Science and Technology	**VDJtools** Analyzing TCR data with VDJtools and VDJdb The tutorial will introduce the audience to post-analysis of T-cell repertoire sequencing data. The major aim of this tutorial is to learn how to infer T-cells specific to certain epitopes and extract T-cell receptor (TCR) sequence motifs from high-throughput immune repertoire sequencing data (AIRR-seq). The audience will get familiar with the VDJdb database that lists TCRs with known antigen specificities, VDJtools software designed for statistical analysis of AIRR-seq data and some useful R templates for TCR sequence analysis.

Posters

All abstracts received by March 15th were considered for oral presentation, if so indicated on the submission. All abstracts received after this date were accepted only for poster presentation.

List of Posters - Session 1 (Monday, May 13th)

Poster Number	Presenter	Title
101	Kenneth B Hoehn	Repertoire-wide phylogenetic models of B cell molecular evolution reveal evolutionary signatures of aging and vaccination
102	Pieter Meysman	On the viability of unsupervised T-cell receptor sequence clustering for epitope preference
103	Ivana Mikocziova	AIRR-seq reveals a large number of germline polymorphisms in the variable immunoglobulin genes
104	Eric Waltari	Functional enrichment and analysis of memory B cell repertoires in PBMCs using an Immcantation-based pipeline
105	Nicola Bonzanni	Bridging the gap: manage, analyze, visualize and interpret repertoire sequencing data using the IGX platform
106	Anna Lorenc	Messenger RNA (cDNA) or genomic DNA (gDNA) for studying T cell receptor (TCR) repertoire; template selection impacts upon observed characteristics
107	Tahel Ronel	A Dynamic Bayesian Network Model for T cell receptor classification by antigen specificity
108	Samuel Schmitz	Human immunome repertoires facilitate B-Cell cDNA recovery and human-likeness evaluation of amino acid immunoglobulins
109	Sofie Gielis	TCRex: a webtool for the prediction of T-cell receptor sequence epitope specificity
110	Lucia Csepregi	Uncovering the physiological network of B cell clonal lineages
111	Cédric Weber	immuneNET: a computational framework for high-dimensional similarity comparison of immune repertoires
112	Milena Pavlović	ImmuneML: an open-source platform for large-scale machine learning on immunereceptor data
113	Valentin Von Niederhaeusern	Development of synthetic standards for error correction in B-cell receptor repertoire sequencing
114	Simon Schäfer	Computational analysis of human heavy chain CDR3 repertoires: The paradox of tyrosine rich antibodies in memory

		B cell repertoires
115	Cecilie Boysen	Immune Receptor Repertoire Analysis and Comparisons – Geneious Biologics NGS Pipeline
116	Bjoern Peters	Epitope Specific Antibodies and T Cell Receptors in the Immune Epitope Database
117	Anna Obraztsova	Antibody repertoire simulation for benchmarking RepSeq data analysis and lineage inference methods
118	Aleksandr Kovaltsuk	Observed Antibody Space: A Resource for Data Mining Next-Generation Sequencing of Antibody Repertoires
119	Konrad Krawczyk	Looking for Therapeutic Antibodies in Next Generation Sequencing Repositories

List of Posters - Session 2 (Tuesday, May 14th)

Poster Number	Presenter	Title
201	Xihao Hu	Landscape of B cell immunity and related immune evasion in human cancers
202	Felix Breden	iReceptor Plus Facilitates Sharing and Analysis of AIRR-seq Data
203	William Lees	Antibody repertoire analysis of inhibitory immune response in hemophilia A
204	Néstor Vázquez Bernat	High-Quality Library Preparation for NGS-Based Immunoglobulin Analysis in Non-Human Primates
205	Duncan Ralph	A method for Identification of new antibodies with high affinity for, or neutralizing ability against, antigens of interest
206	Natasha Smith	Characterisation of a Plasmodium chabaudi infection using Adaptive Immune Receptor Repertoire sequencing
207	Caroline Grönwall	Repertoire Studies In Rheumatoid Arthritis Reveal B-cell Distortions
208	Anastasios Spiliotopoulos	Phage Display Technology at UCB: Ongoing innovation with a mature platform
209	Marie Ghraichy	Functional maturation and self-reactivity of the human B cell system assessed by B-cell receptor repertoire sequencing
210	Maren Lindner	The role of T cells in Neuromyelitis optica disease pathogenesis
211	Sonal Henson	Characterization of the bovine adaptive immune repertoire responses to a candidate vaccine for East Coast Fever
212	Wei Zhang	Dynamics of age-based T-cell receptor repertoire derived from large-scale healthy population
213	Aaron Chevalier	Analysis of transgenic mouse antibody repertoires
214	Keshav Motwani	T-cell receptor repertoires in peripheral blood encode type 1 diabetes status
215	Vadim Nazarov	Prediction of cytomegalovirus serostatus from the T cell repertoire using deep learning
216	Khalil El Mazouari	Antibody-Extractor®, the in-silico antibody sequence discovery platform

217	Govinda Sharma	Novel in vitro methods for the discovery of functional T-cell receptor epitopes from large peptide-coding libraries
218	Sandra Nielsen	IgE B cell responses to environmental factors in infancy and childhood
219	Ayelet Peres	RAbHIT: R Antibody Haplotype Inference Tool

Poster Abstracts - Session 1 (Monday, May 13th)

101

Repertoire-wide phylogenetic models of B cell molecular evolution reveal evolutionary signatures of aging and vaccination

Kenneth B Hoehn

To produce effective antibodies, B cells undergo somatic hypermutation (SHM) and selection during affinity maturation. Phylogenetic models are versatile tools for studying evolution, and may be useful in characterizing this process. However, certain features of affinity maturation violate important assumptions in standard models. Further, most phylogenetic models characterize single lineages, while B cell repertoires often contain thousands of lineages. Here, we introduce a phylogenetic framework that incorporates unique features of SHM, and integrates information between multiple lineages to precisely estimate model parameters. We demonstrate the power of this approach first by using it to show evidence of age and sex-biased changes in SHM. We next show a consistent negative relationship between lineage tree length and signs of negative selection in healthy and recently vaccinated subjects, suggesting that lineages shift towards negative selection over time. Overall, this represents a powerful new phylogenetic framework for characterizing mutation and selection during affinity maturation.

102

On the viability of unsupervised T-cell receptor sequence clustering for epitope preference

Pieter Meysman, Nicolas De Neuter, Sofie Gielis, Danh Bui Thi, Benson Ogunjimi, Kris Laukens

Antwerp Unit for Data Analysis and Computation in Immunology and Sequencing (AUDACIS), University of Antwerp, Antwerp, Belgium

The T-cell receptor (TCR) is responsible for recognizing epitopes presented on cell surfaces. Linking TCR sequences to their ability to target specific epitopes is currently an unsolved problem, yet one of great interest. Indeed, it is currently unknown how dissimilar TCR sequences can be before they no longer bind the same epitope. This question is confounded by the fact that there are many ways to define the similarity between two TCR sequences. In this study, we investigate both issues in the context of TCR sequence unsupervised clustering [1]. We provide an overview of the performance of various distance metrics on two large independent data sets with 412 and 2835 TCR sequences respectively. Our results confirm the presence of structural distinct TCR groups that target identical epitopes. In addition, we find that several different techniques are able to generate clusters with high epitope-specificity purity at an equal rate. For example, it can be shown that a basic approach using a Hamming distance of one, i.e. a single amino acid substitution, was among the best performing methods.

[1] Meysman P, De Neuter N, Gielis S, Bui Thi D, Ogunjimi B, Laukens K. On the viability of unsupervised T-cell receptor sequence clustering for epitope preference. Bioinformatics.-Oxford. 2018:1-7.

103

AIRR-seq reveals a large number of germline polymorphisms in the variable immunoglobulin genes

Ivana Mikocziova[1†], Moriah Gidoni[2†], Ida Lindeman[1], Omri Snir[1], Gur Yaari[2*], Ludvig M. Sollid[1*]

[1]K.G.Jebsen Centre for Coeliac Disease Research and Department of Immunology, University of Oslo and Oslo University Hospital, 0372 Oslo, Norway
[2]Faculty of Engineering, Bar Ilan University, Ramat Gan 5290002, Israel
[†]These authors contributed equally to this work
*Shared senior authors

Polymorphisms in immunoglobulin (Ig) genes may have a profound impact on immune responses in health and disease. Characterisation of germline alleles is especially important for correct estimations of somatic hypermutation. Nonetheless, existing databases of germline Ig genes with annotation of germline polymorphisms are largely incomplete. To discover Ig polymorphisms from adaptive immune receptor repertoire sequencing (AIRR-seq) data, various software packages have been developed. Inferring Ig polymorphisms from AIRR-seq data is more common due to technical difficulties in aligning short reads of non-rearranged genomic DNA to this highly repetitive locus. Therefore, it is difficult to determine whether a novel allele candidate is truly present in the genomic DNA or whether it is an artefact. Here, we used a previously generated AIRR-seq data set from naïve B-cells of 100 individuals, which was analysed by two software tools, TIgGER and IgDiscover, to generate a personalised germline IgV reference for each individual. The availability of genomic DNA from non-B cells of the same subjects allowed us to compare and verify the findings using Sanger sequencing. As a result, novel alleles were confirmed. This work expands our knowledge about the diversity of IgV germline genes across the Caucasian population.

104

Functional enrichment and analysis of memory B cell repertoires in PBMCs using an Immcantation-based pipeline

Eric C. Waltari[1], A. McGeever[1], Peter S. Kim[1,2,3], Krista M. McCutcheon[1]

[1]Chan Zuckerberg Biohub
[2]Department of Biochemistry, Stanford University School of Medicine
[3]Stanford ChEM-H

Phenotypic screening of antigen-specific antibodies in human blood is a common diagnostic test for infectious agents and a correlate of protection after vaccination. In addition to long-lived antibody secreting plasma cells residing in the bone marrow, memory B cells are a latent source of antigen-experienced, long-term immunity that can be found at low frequencies in circulating PBMCs. Assessing the genotype, clonal frequency, quality, and function of antibodies
residing in an individual's persistent memory B cell repertoire can help inform the success or failure of immune protection. We are refining methods to functionally expand the memory repertoire from PBMCs and clonally map monoclonal antibodies from this population, using the Dockerized Immcantation framework as the backbone of a cloud computing-based pipeline. We show how the combined methods of deep sequencing stimulated memory B cell repertoires and retrieving single, activated and expanded antigen-specific cells is a promising approach in evaluating the latent, functional B cell memory in PBMCs.

Funding: Chan Zuckerberg Biohub

105

Bridging the gap: manage, analyze, visualize and interpret repertoire sequencing data using the IGX platform

Henk-Jan van den Ham, Lorenzo Fanchi, Alvise Trevisan, Nicola Bonzanni

ENPICOM B.V., ’s-Hertogenbosch, The Netherlands

Repertoire analysis is rapidly gaining traction as an essential tool for immunotherapy development, patient stratification before start of treatment and monitoring of patients on treatment. Repertoire analysis combines the unique sequences of T- and B-cell receptors and the power of high-throughput sequencing to identify specific T- and B-cell clones. However, the implementation, analysis and interpretation of repertoire data in research and clinical workflows is challenging and requires specialized expertise across various disciplines like immunology, bioinformatics and software engineering. To date, no integrated, end-to-end solution exists.

To address this unmet need, we developed the ImmunoGenomiX (‘IGX’) platform: a new versatile platform for the analysis and interpretation of high-throughput repertoire sequencing data. IGX is an end-to-end platform, enabling users to perform state-of-the-art analyses, starting from raw sequencing files, up to the generation of report or publication-ready figures. The powerful, tag-based, annotation system of the IGX platform allows users to extensively annotate repertoire sequencing data with structured metadata at great level of granularity, including annotations for single clones such as affinity or avidity. Sequencing data, as well as metadata, can be used to perform powerful searches, create collections of clones, or subset and filter data for follow-up analysis. The IGX platform features a user-friendly interface that helps researchers and clinicians bridge the gap between the staggering amounts of generated sequencing data and its interpretation, without requiring any computational or bioinformatic expertise.

106

Messenger RNA (cDNA) or genomic DNA (gDNA) for studying T cell receptor (TCR) repertoire; template selection impacts upon observed characteristics

Anna Lorenc[1], Iria Gomez-Tourino[2], Mark Peakman[1]

[1]Departement of Immunobioogy, Kings's College London
[2]Immunobiology Laboratory, Centro Singular de Investigación de Galicia, University of Vigo

Adaptive immune receptor repertoire sequencing (AIRR-seq) is a promising way to gain insight into the role of T cells in the immunopathology of autoimmune diseases, cancer and infection, through high throughput T cell receptor (TCR) sequencing . The decision as to whether to use RNA or DNA as the starting material for sequencing is often dictated by the availability of cells or tissue, sequencing protocol or assay availability; where there is a choice of template, it remains unclear what the impact of cDNA versus gDNA may be. We therefore examined how the selection of template might influence the constitution and statistical analyses describing TCR repertoire following deep repertoire sequencing.

We isolated total RNA and DNA in parallel from two cell types differing in clonotype distribution –naïve CD4+ T cells, with a very diverse repertoire, and central memory CD4+ T cells, the repertoire of which is pruned and shaped by antigen selection. After reverse transcription of RNA, cDNA and gDNA were sequenced in the same experimental run with the same protocol (Adaptive ImmunoSEQ assay).

We observe important differences between matched samples, that may be attributable to TCR expression levels and the use of reverse transcription, and demonstrate how these affect repertoire characteristics and overlap. Our findings will assist with experimental protocol selection and help with comparison and interpretation of results based on different starting material.

107

A Dynamic Bayesian Network Model for T cell receptor classification by antigen specificity

Tahel Ronel[1], Matthew Harries[2,3], Kate Wicks[2], Imran Uddin[1], Annemarie Woolston[1], Theres Oakes[1], Mazlina Ismail[1], Connor Husovsky[1], Rebecca Dearman[2], Gavin Maxwell[4], Benny Chain[1,5]

[1] Division of Infection and Immunity, University College London, London, UK
[2] Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK
[3] Salford Royal NHS Foundation Trust (Dermatology Centre), Salford, UK
[4] SEAC, Unilever, Colworth Science Park, Bedford, UK
[5] Department of Computer Science, University College London, London, UK

Functional annotation of TCR sequences according to antigen specificity remains a major challenge in interpreting the rapidly increasing amount of TCRseq data being generated. TCRs that recognise the same antigen are thought to share some sequence features, but these are not well understood. State-of-the-art algorithms can typically characterise only around 15% of a repertoire's sequences. We have developed a Dynamic Bayesian Network (DBN) model to represent and hence classify TCR sequences. DBNs are probabilistic graphical models, composed of a directed acyclic graph with variables of the system as nodes, and edges depicting conditional dependencies between the variables. The TCR DBN we have developed incorporates prior knowledge which suggests that short motifs of amino acids within the CDR3 are important in conferring antigen specificity. We have tested our DBN model on two sets of data: a set of annotated CMV-related TCRs from the public VDJdb database, and a set of TCRs that we sequenced from peripheral blood of a cohort of patients sensitised by diphenylcyclopropenone (DPC) for treatment of alopecia areata [study approved by NRES Ethics Committee East of England – Norfolk [14/EE/1067]]. The models are trained on a subset of TCR sequences functionally annotated for antigen specificity (CMV or DPC), or a balanced set of randomly selected control TCRs which we have sequenced. The output of the model is the class of TCR (antigen specific or control). The DBN is flexible and allows incorporating into the model information from different data sources. Moreover, decision-making in the model is fully traceable, which enables learning structural sequence information, beyond probabilistic classification. The DBN models performed at least as well as current methods, and could be used to infer sequence features which may contribute to antigen recognition. Our results suggest they will be a useful addition to the TCRseq analysis toolbox.

108

Human immunome repertoires facilitate B-Cell cDNA recovery and human-likeness evaluation of amino acid immunoglobulins

Samuel Schmitz[1], Cinque Soto[2], James E. Crowe[3] and Jens Meiler[1],*

[1] Department of Chemistry, Vanderbilt University, Nashville, TN 37212
[2] Department of Pediatrics, Infectious Diseases Division, Vanderbilt University Medical Center, Nashville, TN 37212
[3] Department of Pathology, Microbiology, and Immunology, Vanderbilt University Medical Center, TN 37212

The human fragment variable region (Fv) of antibodies consist of the heavy chain (HC) and light chain (LC) variable regions. The underlying germline gene rearrangement of variable (V), diversity (D) and joining (J) segments defines the sequence diversity of human antibodies that ultimately gives rise to their ability to bind to a wide variety of targets. Assignment and validation of complete V(D)J germline gene rearrangements to amino acid sequences remains a challenge due to high mutation rates. Here, we developed a method to assign V(D)J germline gene rearrangements to amino acid sequences, and suggest a probable human B-Cell cDNA from a statistical model. We captured position specific single nucleotide polymorphisms from human immunomes of 308 million sequences in form of position specific frequency matrices (PFMs). Our statistics cover almost the complete antibody variable region (Fv) with germline gene dependent V, D, and J PFMs as well as PFMs for untemplated junctions at the beginning and end of CDRH3 loops. Recovered nucleotide sequences of human immunoglobulins (Ig) are 91.4% (heavy chain) and 94.5% (light chain) identical in our native sequence recovery benchmark. We introduce the pfmS metric, which scores the similarity of the nucleotide sequence to human immunome repertoires of healthy blood donors. The pfmS metric can assist to determine if an antibody exhibits untypical sequence features indicating humanization, engineering efforts, or auto-immune diseases. We are hosting our algorithm.

Supported through the Vanderbilt Trans-Institutional Program (TIP) “Integrating Structural Biology with Big Data for next Generation Vaccines”.

109

TCRex: a webtool for the prediction of T-cell receptor sequence epitope specificity

Sofie Gielis[1,2], Pieter Moris[1], Nicolas De Neuter[1,2], Wout Bittremieux[1], Benson Ogunjimi[2], Kris Laukens[1,2], Pieter Meysman[1,2]

[1] Adrem Data Lab, Department of Mathematics and Computer Science, University of Antwerp, Antwerp, Belgium
[2] AUDACIS, Antwerp Unit for Data Analysis and Computation in Immunology and Sequencing, University of Antwerp, Antwerp, Belgium

To date, multiple immunoinformatics tools have been created with the goal to achieve a better understanding of immunological processes. Although great tools exist for the prediction of epitopes and their binding to MHC molecules, we are still lacking useful tools for the prediction of epitope-MHC recognition by TCRs. Hence, we propose TCRex, a tool to investigate TCR recognition of epitopes. This tool is based on our prior work related to the feasibility of predicting TCR-epitope recognition using TCRβ sequences. In this study, we showed that a random forest classifier trained to predict TCR-epitope interactions from TCR amino acid physicochemical properties can achieve a high accuracy. We extended this work into a toolbox trained on a large dataset containing epitopes from different viruses and tumour cells. To this end, we collected epitope-specific human TCRβ sequence data containing information about the CDR3 sequences and the corresponding V-and J-genes. Random forest classifiers were trained on this data and kept if they report a sufficiently high performance in a cross-validation setting. These classifiers are freely available in a webtool, called TCRex, at tcrex.biodatamining.be TCRex is useful to make predictions on newly gathered experimental TCRβ sequence data. As such, it will aid researchers in the elucidation of T cell repertoire targets.

Reference:
S. Gielis, P. Moris, N. De Neuter, W. Bittremieux, B. Ogunjimi, K. Laukens and P. Meysman, “TCRex: a webtool for the prediction of T-cell receptor sequence epitope specificity,” bioRxiv, 2018.

110

Uncovering the physiological network of B cell clonal lineages

Lucia Csepregi and Sai Reddy

Department Biosystems Science and Engineering, Basel
ETH Zurich, Switzerland

While clonal selection, expansion and diversification (via somatic hypermutation) are well established concepts of humoral immunity, the spatial distribution of B cell clonal lineages remains largely unexplored. Understanding the physiological network of B cell clonal selection could provide new ways of tracking and targeting clones, thereby yielding more insight on vaccine design. In this regard, we perform immunization studies in wild type mice using the pre-fusion glycoprotein of respiratory syncytial virus. We aim to determine how different immunization schemes influence clonal distribution of B cells. We accomplish this by quantitatively interrogating the antibody repertoires of B cells in various lymphoid organs, bone marrow and blood by high-throughput sequencing (HTS). We make use of an advanced protocol that uses molecular barcodes to correct for error and bias introduced during library preparation and sequencing. HTS coupled to bioinformatic analysis of clonal lineages will give insights into the hierarchical relationships between B cells, allowing us to build a map of clonal networks within and between physiological compartments. We hypothesize that analyzing the network formation of clonal lineages will lead to the detection of clones that extensively share sequence variants among different organs, which might enable the identification of antigen-specific antibody sequences.

111

immuneNET: a computational framework for high-dimensional similarity comparison of immune repertoires

Cédric R. Weber[1*], Sai T. Reddy[1#] and Victor Greiff[2#]

[1]ETH Zürich, D-BSSE, Mattenstrasse 26, 4058, Basel, Switzerland.
[2]Departement of Immunology, University of Oslo, Sognsvannsveien 20, Rikshospitalet, 0372, Oslo, Norway.
[*]Presenting author
[#]Corresponding authors

It is thought that infections continually shape immune repertoires by frequency-dependent clonal expansion and sequence diversification thereby generating a repertoire-based record of past and current immune states. If true, these records are stored in a high-dimensional repertoire space. This space, however, has remained inaccessible to previous approaches that have focused almost exclusively on unidimensional features. Specifically, unidimensional features fail to capture the high-dimensional complexity of immune repertoires spanned by both frequency and sequence-dependent features, thus limiting our ability to evaluate the true similarity of immune repertoires across disease states. Therefore, we sought to build a networks-based framework (immuneNET), for quantifying immune repertoire similarity based on the multidimensional combination of immunological features that capture repertoire architecture. These repertoire features cover the whole spectrum from total frequency to total sequence-dependence thus encompassing the full dimensionality of the repertoire space. Considering today's challenges of ever growing repertoire datasets, our networks-based definition of similarity explicitly allows not only pairwise, but also one-to-many and many-to-many repertoire comparisons. This implies that we can embed an individual's repertoire within a population's similarity space. In summary, we have defined and implemented a novel multi-feature repertoire similarity measure that accounts for repertoire dimensionality and achieves individual similarity embedding.

112

ImmuneML: an open-source platform for large-scale machine learning on immune receptor data

Milena Pavlović[1,2] , Keshav Motwani[3] , Rahmad Akbar[2] , Cédric Weber[4] , Igor Snapkov[2] , Sai T. Reddy[4] , Todd Brusko[3] , Victor Greiff[2] , Geir K. Sandve[1]

[1] Department of Informatics, University of Oslo
[2] Department of Immunology, University of Oslo
[3] Department of Pathology, Immunology and Laboratory Medicine, University of Florida
[4] Department of Biosystems Science & Engineering, ETH Zürich

Immune B- and T-cell receptors are natural diagnostics and therapeutics as they encode disease-specific immune information. This information is complex, high-dimensional, and hidden in large amounts of receptor data, thus rendering machine learning approaches necessary for its recovery. However, the amount and the intricacy of the data render machine learning analyses challenging. To facilitate the systematic application of machine learning on immune receptor data, we present our progress towards ImmuneML, an open-source platform for large-scale machine learning on immune receptor data. ImmuneML supports the study of experimental immune receptor data, as well as synthetic benchmarking data. It enables training, assessment and feature recovery of machine learning algorithms in order to expedite the future in silico discovery of immunodiagnostics and immunotherapeutics.

113

Development of synthetic standards for error correction in B-cell receptor repertoire sequencing

Valentin von Niederhäusern[1,2], Jacob D. Galson[1,2], Marie Ghraichy[1,2], Johannes Trück[1,2]*

[1]Division of Immunology, University Children's Hospital Zurich
[2]Children's Research Center, University Children's Hospital Zurich, University of Zurich

Despite the broader use of immune repertoire sequencing techniques in recent years, standardization and bias-correcting approaches are mainly focusing on bioinformatic filtering and sequence collapsing strategies while the use of biological standards is not part of routine workflows.

A synthetic B-cell receptor (BCR) repertoire consisting of 62 diverse BCR sequences including all functional variable (V) and constant region genes was PCR amplified using our standard set of V family primers (multiplex) or amplified using a single universal primer (singleplex). Data were processed in a standardized way using an in-house bioinformatic pipeline with output analyzed based on input concentrations (1e4–1e12 transcripts) and compared between expected and observed frequencies as well as between types of PCR amplification (singleplex vs. multiplex).

There was substantial variation in V family usage in both singleplex and multiplex samples indicating not only PCR amplification but also sequencing bias. This effect was less marked when only high-quality reads were considered. The comparison of singleplex with multiplex repertoires indicated a several-fold variability of V gene usage irrespective of input concentrations. The evaluation of isotype subclasses revealed overrepresentation of sequences annotated as IgD while IgA2 and IgG2 sequences were underrepresented in all samples.
We have developed a biological standard that allows to assess and ultimately correct PCR and sequencing bias across and within sequencing runs.

114

Computational analysis of human heavy chain CDR3 repertoires: The paradox of tyrosine rich antibodies in memory B cell repertoires

Simon Schäfer[a,b], Thomas Winkler[a] and Heinrich Sticht[b]

[a]Chair of Genetics and
[b]Division of Bioinformatics, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany

Next generation sequencing analysis of distinctive B cell subsets from leukapheresis products of hematopoietic stem cells from healthy donors [1] proved to be a powerful tool to characterize the human antibody repertoire. The antibodies produced by human naïve and memory B cells show significant differences in their heavy chain CDR3 repertoires. We prove that human heavy chain CDR3 loops are shorter in their memory repertoire compared to the naïve repertoire of the same donor and show a shift in the usage of an important J gene responsible for the creation of long CDR3 (J6) [2].

In addition, the quantification of specific amino acid motifs observed among all healthy donors hint that positive and negative selection shape the antibody memory CDR3 repertoire as opposing principles very specifically. These effects are highlighted by the observation of the paradox role of tyrosine repeats introduced by J gene 6. The usage of J6 is decreased in memory repertoire of healthy donors compared to the naïve repertoire of the same donor. For the CDR regions Tyr is described by previous studies to be beneficial for antigen recognition and is expected to be positively selected [3]. In contrast, we show the decrease of Tyr motifs and Tyr rich J genes (J6) in all analyzed memory repertoires. Previous and current studies describe the prevalence of antibodies containing J6 in patients with autoimmune disease and the usage of J6 in antibodies against neo antigens in cancer patients [4,5]. We therefore suggest that tyrosine rich antibodies with long CDR3 are negatively selected and are an important future research subject in understanding autoimmunity.

[1] Tittlbach, Hannes; Schneider, Andrea; Strobel, Julian; Zimmermann, Robert; Maas, Stefanie; Gebhardt, Bernd et al. (2017): Journal of translational medicine 15 (1), S. 228. DOI: 10.1186/s12967-017-1330-5.
[2] Briney, Bryan S.; Willis, Jordan R.; Crowe, James E. (2012): PloS one 7 (5), e36750. DOI: 10.1371/journal.pone.0036750.
[3] Birtalan, Sara; Zhang, Yingnan; Fellouse, Frederic A.; Shao, Lihua; Schaefer, Gabriele; Sidhu, Sachdev S. (2008): Journal of Molecular Biology 377 (5), S. 1518–1528. DOI: 10.1016/j.jmb.2008.01.093.
[4] Foster, Mary H.; Buckley, Elizabeth S.; Chen, Benny J.; Hwang, Kwan-Ki; Clark, Amy G. (2016): Molecular immunology 76, S. 123–133. DOI: 10.1016/j.molimm.2016.07.004.
[5] Liu, Song; Zhu, Ying; Lin, Lie-Wen; Ding, Shun-Kai; Lin, Xiao-Cong; Zhong, Ke-Li et al. (2018): Oncology letters 16 (1), S. 239–246. DOI: 10.3892/ol.2018.8677.

115

Immune Receptor Repertoire Analysis and Comparisons – Geneious Biologics NGS Pipeline

Cecilie Boysen, Matt Kearse, Geoffrey Gonzalez-Escobedo, Owen Bodley, Christian Stenvang, Megan Kennington, Alan Dragicevich, Alicia Lai, Jannick Bendtsen

Biomatters

Geneious Biologics is a cloud-based platform for analysis of antibody and TCR repertoires. Preprocessing, annotation and clustering pipelines, alignments of sequences or clusters of sequences, as well as comparisons between samples, can be run from a Graphical User Interface, and thus require no command-line or programming skills. In addition, as an enterprise solution with a developed set of reporting tools and graphical displays of results, collaboration and sharing of data across different stakeholders within an organization is easy without the need to send data back and forth.

Here we present NGS analyses and comparisons of BCR and TCR repertoire data, including preprocessing steps as needed, annotations and clustering of FR and CDR regions, identification of isotypes and V(D)J germline usage and variants from these.

116

Epitope Specific Antibodies and T Cell Receptors in the Immune Epitope Database

Swapnil Mahajan[1], Randi Vita[1], Deborah Shackelford[1], Jerome Lane[1], Veronique Schulten[1], Laura Zarebski[1], Martin Closter Jespersen[2], Paolo Marcatili[2], Morten Nielsen[2,3], Alessandro Sette[1,4], Bjoern Peters[1,4]*

[1]Center for Infectious Disease, La Jolla Institute for Allergy and Immunology, La Jolla, CA, United States
[2]Department of Bio and Health Informatics, Technical University of Denmark, Kongens Lyngby, Denmark
[3]Instituto de Investigaciones Biotecnológicas, Universidad Nacional de San Martín, Buenos Aires, Argentina
[4]University of California San Diego, La Jolla, CA, United States

The Immune Epitope Database (IEDB) is a free public resource which catalogs experiments characterizing immune epitopes. To accommodate data from next generation repertoire sequencing experiments, we recently updated how we capture and query epitope specific antibodies and T cell receptors. Specifically, we are now storing partial receptor sequences sufficient to determine CDRs and VDJ gene usage which are commonly identified by repertoire sequencing. For previously captured full length receptor sequencing data, we have calculated the corresponding CDR sequences and gene usage information using IMGT numbering and VDJ gene nomenclature format. To integrate information from receptors defined at different levels of resolution, we grouped receptors based on their host species, receptor type and CDR3 sequence. These data are accessible as full exports and through a new dedicated query interface. The later combines the new ability to search by receptor characteristics with previously existing capability to search by epitope characteristics such as the infectious agent the epitope is derived from, or the kind of immune response involved in its recognition. We expect that this comprehensive capture of epitope specific immune receptor information will provide new insights into receptor-epitope interactions, and facilitate the development of novel tools that help in the analysis of receptor repertoire data.

Funding: NIH contracts 75N93019C00001 and HHSN272201200010C

117

Antibody repertoire simulation for benchmarking RepSeq data analysis and lineage inference methods

Anna Obraztsova[1], Mikhail Shugay[1,2,3,4], Dmitry Chudakov[1,2,3,4,5]

[1] Skolkovo Institute of Science and Technology, Russia;
[2] Institute of Bioorganic Chemistry (RAS), Russia;
[3] Pirogov Russian National Research Medical University, Russia;
[4] Privolzhsky Research Medical University (PIMU), Russia;
[5] Central European Institute of Technology (CEITEC), Czechia

High-throughput sequencing of antibody repertoires is now widely applied in studies of the adaptive immune response both in health and disease. As the amount of data produced by Rep-Seq continually increases, the demand for bioinformatic utilities for analysis of this data is growing and there is a need for synthetic antibody repertoires with known properties to be used as a benchmark. We developed a framework for antibody repertoire simulation that can be ne-tuned to recapture features of immunoglobulin heavy chain sequencing datasets of various origin: from vaccinated donors to healthy donor memory B-cell repertoires. The ability to control and track cell division and mutation events, as well as the sampling effect, allowed us to validate existing methods for antibody repertoire processing and B cell lineage inference. We report a model parameter set that provides the best t to existing B-cell repertoire data, explore the accuracy issues of the existing software tools for AIRR-seq data processing, and propose a novel algorithm for fast and accurate antibody lineage Inference.

This work was supported by Russian Science Foundation grant No 17-15-01495.

118

Observed Antibody Space: A Resource for Data Mining Next-Generation Sequencing of Antibody Repertoires

Aleksandr Kovaltsuk[1], Jinwoo Leem[1], Sebastian Kelm[2], James Snowden[2], Charlotte M. Deane[1], Konrad Krawczyk[1,3]

[1]Department of Statistics, University of Oxford, Oxford OX1 3LB, United Kingdom
[2]UCB Pharma, Slough SL1 3WE, United Kingdom
[3]Natural Antibody, Hamburg, Germany

Antibodies are immune system proteins that recognize noxious molecules for elimination. Recently, it has become possible to query their immense natural diversity using next-generation sequencing of immunoglobulin gene repertoires (Ig-seq). However, Ig-seq outputs are currently fragmented across repositories and tend to be presented as raw nucleotide reads, which means non-trivial effort is required to reuse the data for analysis. To address these issues, we have created the Observed Antibody Space (OAS) resource that allows large-scale data mining of antibody repertoires (1). We have, so far, collected and cleaned the raw outputs of 60 Ig-seq experiments, covering over a billion sequences. We have organized the sequences by metadata, such as organism, isotype, B cell type, and source, and the immune status of B cell donors to facilitate bulk retrieval of specific subsets for comparative analyses. We have converted all of the Ig-seq sequences to amino acids while preserving the link to the respective original raw nucleotide sequences and numbered them using the IMGT scheme. The data are available for querying or bulk download at http://antibodymap.org/oas. We believe that OAS will facilitate data-mining Ab repertoires for improved understanding of the dynamics of the immune system and, thus, better engineering of biotherapeutics.

References

1. Kovaltsuk, A., J. Leem, S. Kelm, J. Snowden, C. M. Deane, and K. Krawczyk. 2018. Observed Antibody Space: A Resource for Data Mining Next-Generation Sequencing of Antibody Repertoires. J. Immunol. 201: 2502–2509.

Funding:
Biotechnology and Biological Sciences Research Council (BBSRC) [BB/M011224/1] and UCB Pharma Ltd
Industrial Fellowship from the Royal Commission for the Exhibition of 1851

119

Looking for Therapeutic Antibodies in Next Generation Sequencing Repositories

Konrad Krawczyk[1], Matthew Raybould[2], Aleksandr Kovaltsuk[2], Charlotte M. Deane[2]

[1] NaturalAntibody, Hamburg, Germany
[2] Oxford University Department of Statistics, Oxford, UK

Abstract**:** Recently it has become possible to query the great diversity of natural antibody repertoires using Next Generation Sequencing (NGS). These methods are capable of producing millions of sequences in a single experiment. Here we compare Clinical Stage Therapeutic antibodies to the ~1b sequences from 60 independent sequencing studies in the Observed Antibody Space Database. Of the 242 post Phase I antibodies, we find 16 with sequence identity matches of 95% or better for both heavy and light chains. There are also 54 perfect matches to therapeutic CDR-H3 regions in the NGS outputs, suggesting a nontrivial amount of convergence between naturally observed sequences and those developed artificially. This has potential implications for both the discovery of antibody therapeutics and the legal protection of commercial antibodies.

Funding source: Natural Antibody

Poster abstracts - session 2 (Tuesday, May 14th)

201

Landscape of B cell immunity and related immune evasion in human cancers

Xihao Hu, Jian Zhang, Jin Wang, Jingxin Fu, Taiwen Li, Xiaoqi Zheng, Binbin Wang, Shengqing Gu, Peng Jiang, Jingyu Fan, Xiaomin Ying, Jing Zhang, Michael C. Carroll, Kai W. Wucherpfennig, Nir Hacohen, Fan Zhang, Peng Zhang, Jun S. Liu, Bo Li, X. Shirley Liu

Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute and Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA.

Tumor-infiltrating B cells are an important component in the microenvironment with unclear anti-tumor impacts. We enhanced our previous computational algorithm TRUST to extract the B cell immunoglobulin (Ig) hypervariable regions from bulk tumor RNA-seq data. TRUST assembled over 30 million complementarity-determining region 3 (CDR3s) of the B cell heavy chain (IgH) from The Cancer Genome Atlas (TCGA). Widespread B cell clonal expansions and Ig subclass switch events were observed in diverse human cancers. Prevalent somatic copy number alterations in MICA and MICB genes related to antibody-dependent cell mediated cytotoxicity (ADCC) were identified in tumors with elevated B cell activity. IgG3-1 subclass switch interacts with the B cell receptor affinity maturation and defects in the ADCC pathway. Comprehensive pan-cancer analyses of tumor-infiltrating B cell receptor repertoires identified novel tumor immune evasion mechanisms through genetic alterations. The IgH sequences identified here are potentially useful resources for future development of immunotherapies.

Note:
The full manuscript is in press in Nature Genetics

202

iReceptor Plus Facilitates Sharing and Analysis of AIRR-seq Data

iReceptor Plus Consortium

The international iReceptor Plus Consortium, composed of more than 20 partners from 9 countries, will promote human immunological data storage, integration and controlled sharing for a wide range of clinical and scientific purposes. The four-year project will expand iReceptor, an innovative platform that integrates massive distributed repositories of Adaptive Immune Receptor Repertoire sequence (AIRR-seq) data (antibody/B-cell and T-cell receptor data).

Most AIRR-seq data are currently stored and curated by individual labs, using a variety of tools and technologies. The project will support the sharing of public AIRR-seq data across labs, diseases and institutions, guided by the protocols and standards developed by the AIRR Community. iReceptor Plus will develop and link to new analysis tools, integrate AIRR-seq data with other types of clinical and “omics” data, including single-cell data, and offer an advanced and collaborative immune-profiling hub. The project will add advanced security during transfer of public data, and provide a mechanism for users to protect private data when such protection is necessary. The platform’s software will be free through open-source licensing, making it possible for the research community to extend and adapt the tools and technologies used in the project. The iReceptor Plus project will lower the barriers to access, integrate and analyze large AIRR-seq datasets, in order to advance our understanding of immune responses, and thus provide new targets for immunotherapies and new methods for monitoring therapeutic efficacy.

Funded by the European Union’s Horizon 2020 research and innovation programme under grant agreement No 825821 and the Canadian Institutes of Health Research (CIHR).

203

Antibody repertoire analysis of inhibitory immune response in hemophilia A

William Lees[1], Eva Wozniak[3], Paul Batty[2,3], Charles Mein[3], Daniel Hart[2,3], Adrian Shepherd[1]

[1.] Department of Biological Sciences and Institute of Structural and Molecular Biology, Birkbeck, University of London, London, UK
[2.] The Royal London Hospital Haemophilia Centre, Barts Health NHS Trust, London, UK
[3.] Blizard Institute, Barts and The London School of Medicine and Dentistry, QMUL, London, UK

Inhibitory antibodies to infused factor VIII (fVIII) represent a serious complication in the treatment of hemophilia A. Although examples of inhibitory anti-fVIII monoclonal antibodies (mAbs) have been characterized (such as the fVIII inhibitor BO2C11), the diversity and development pathways of such antibodies is poorly understood. Characterizing the antibody repertoires of patients using next-generation sequencing (AIRRSeq) can provide insights into germline utilization, clonal development and isotype usage at an unprecedented level of detail. The application of this technique to time-series samples collected from patients as part of the GENA05 study provides the opportunity to compare repertoire development in patients with and without inhibitors, and to examine the development of inhibition. We anticipate that this study will aid the identification of prognostic markers that will facilitate the development of improved treatment regimens for hemophilia A patients.

To our knowledge, this represents the first application of NGS-based antibody repertoire analysis to the development of fVIII inhibition. In the course of this work we have developed and validated an experimental pipeline which facilitates the use of AIRRSeq in a clinical setting with minimal disruption.

This work was funded in part by Octapharma. The authors declare no conflicts of interest.

204

High-Quality Library Preparation for NGS-Based Immunoglobulin Analysis in Non-Human Primates

Néstor Vázquez Bernat, Martin M. Corcoran, Mateusz Kaduk, Gunilla B. Karlsson Hedestam

Department of Microbiology, Tumor and Cell Biology, Karolinska Institutet, SE-171 77 Stockholm, Sweden

Next generation sequencing (NGS) of antibody (Ab) repertoires (Rep-seq) enables examination of the adaptive immune system at an unprecedented depth. Applications include studies of repertoires induced by infection or vaccination, analysis of Ab somatic hypermutation levels, lineage tracing studies and identification of genetic variation within the immunoglobulin (Ig) loci through inference methods. All these applications require starting libraries that allow the generation of sequence data with low error rate and optimal representation of the expressed repertoire. Here, we provide protocols for the production of libraries suitable for all these applications. We describe an improved 5'RACE technique that reduces the length constraints of Illumina MiSeq based Rep-seq analysis and provides information about sequences upstream of Ab V genes. This approach allows for 5' primer design in species where limited genomic information is available, such as non-human primates (NHPs), a commonly used model for vaccine evaluation and studies of host responses to infection. We then describe a 5' multiplex method for library preparation, which yields full length V(D)J sequences suitable for genotype identification and Ab lineage tracing. Using the optimized 5' multiplex protocol, we produced IgM libraries for a number of macaques from difference species/regions and analyzed them using the germline inference tool IgDiscover (1) to identify expressed germline V alleles. This process uncovered a higher degree of allelic diversity in NHPs compared to in humans, as well as numerous alleles not currently described in the IMGT reference database or elsewhere. The library generation protocols and primer presented here enable robust means of analyzing expressed Ab repertoires, identifying novel alleles and producing individualized germline gene databases from any immunologically relevant animal model.

1. Corcoran, M. M., Phad, G. E., Vazquez Bernat, N., Stahl-Hennig, C., Sumida, N., Persson, M. A., Martin, M., Karlsson Hedestam, G. B. (2016). Production of individualized V gene databases reveals high levels of immunoglobulin genetic diversity. Nat Commun, 7, 13642. doi:10.1038/ncomms13642\

This work was funded by grants from the Swedish Research Council (grant 2017-00968), the National Institutes of Health (NIH) (P01 HIVRAD grant AI104722) and the European Union's Horizon 2020 Research and Innovation Programme (Grant agreement No. 681137) to GKH.g

205

A method for Identification of new antibodies with high affinity for, or neutralizing ability against, antigens of interest

Duncan Ralph

Identification of new antibodies with high affinity for or neutralizing ability against antigens of interest, particularly HIV, is an active area of current research. Most approaches, however, require expensive and time-consuming wet lab techniques. A computational approach to this problem might therefore be able to provide a useful adjunct, allowing experimental effort to be better focused on more promising antibodies. Here we describe such a method using the tree shape metrics local branching index (LBI) and local branching ratio (LBR), and show how they correlate to affinity and fitness on simulation and real data samples. We give particular focus to data samples from subjects previously identified as harboring broadly neutralizing anti-HIV antibodies (bNAbs), in particular subject BF520.

This work was supported by a 2018 University of Washington/Fred Hutch Center for AIDS Research Young Investigator award (P30 AI027757), and by NIH grants U19-AI117891, R01-GM113246, R01-AI120961, and R01-AI138709.

206

Characterisation of a Plasmodium chabaudi infection using Adaptive Immune Receptor Repertoire sequencing

Natasha L. Smith, Catherine L. Heylings, Wiebke Nahrendorf, Philip J. Spence, Graeme J. M. Cowan

Institute of Immunology and Infection Research, University of Edinburgh

Despite extensive research, there remains no effective vaccine licensed for use against malaria. This has in part been attributed to the lack of available knowledge regarding the development dynamics of the adaptive immune response, and the parasitic products that drive naturally acquired immunity. The advent of high-throughput sequencing has now provided the opportunity to study the immune repertoires that are generated following a Plasmodium infection, allowing detailed examination of both T and B cell repertoire clonal structure, and identification of expanded clones resulting from exposure to the parasite.

Using a comparative infection model of recently mosquito-transmitted (MT) parasites vs parasites that have been serial blood-passaged (SBP), we have sequenced the T-naïve (CD62L+ CD127+), T-effector (CD127-, CD62L-), T-effector memory (CD44hi, CD127+, CD62L-) and T-central memory (CD44hi, CD127+, CD62L+) CD4+ TCR immune repertoires of C57Bl/6 mice produced over the time-course of a P. chabaudi infection. We demonstrate that during the acute effector proliferation, although the expansion is highly polyclonal, there is a greater degree of similarity of clone usage between replicate repertoires. This indicates a conserved response between individual mice at the clonal level, a shared response that is also more enhanced in mice infected with MT parasites than those infected with SBP parasites. We also show that these repertoires have dominant TCRBV3 gene usage compared to unchallenged repertoires, and that TCRBV3 clones of challenged mice preferentially pair with TCRBJ1-1, TCRBJ2-7, and TCRBJ2-4. This, combined with these repertoires also having a greater degree of clones that are near-identical to each other, indicates these repertoires have the hallmarks of a public potentially antigen-specific response. This response, although also detectable in both memory populations, does not dominate nor expand within memory repertoires, which appear to be a private response within each individual replicate mouse.

Funding for this project was provided by the Wellcome Trust through the primary author’s PhD studentship.

207

Repertoire Studies In Rheumatoid Arthritis Reveal B-cell Distortions

Yan Wang[1], Katy A. Lloyd[1], Sunithi Gunasekera[2], Camilla Eriksson[2], Daniel Ramsköld[1], Karin Lundberg[1], Per-Johan Jakobsson[2], Ulf Göransson[2], Vivianne Malmström[1], Caroline Grönwall[1]

[1.] Rheumatology Unit, Department of Medicine Solna, Karolinska Institutet, Karolinska University Hospital, 171 76 Stockholm, Sweden
[2.] Division of Pharmacognosy, Department of Medicinal Chemistry, Biomedical Centre, Uppsala University, Uppsala, Sweden

The immunological hallmark of rheumatoid arthritis (RA) is autoreactivity to citrullinated proteins and rheumatoid factor activity and B cells are postulated to be central in RA pathogenesis. Here, we study overall B cell repertoire shifts in seropositive RA patients (nti13) compared to matched healthy individuals (nti6) as well as in citrulline specific B cell populations. We investigated the overall recombined BCR repertoire in out-of-frame (OOF) DNA, the expressed BCR repertoire in unmutated IgM after negative selection, the class-switched IgG/IgA after positive selection and peripheral tolerance, and an autoantigen specific anti-citrullinated fibrinogen (Cit-Fib) tetramer-positive repertoire. Several significant shifts in the RA B-cell repertoire could be observed. Strikingly, there was a higher frequency of VH with low somatic hypermutation (SHM) level in RA-derived B cells (<5 mutations, $p<0.0001$ 14.7% vs 8,7). This was especially prominent in IgG1 rearrangements (9.6% vs 18.8% low mutation, $p<0.0001$ ORti2.2 CI:2.0-2.35). In both RA and healthy donors, the frequency of VH3 are decreased from OOF to IgM but increased again in class-switched (HTL: OOF 64.5%; IgM 48.4%, 61%IgG/A) suggesting and intrinsic autoreactivity in VH3 genes. VH4 slightly increased with each selection step (healthy: OOF 7.11%; IgM 9.4%; IgG/A 11.7%) and VH4 were higher in RA compared to healthy OOF (8.4% vs 7.1%) and in the expressed BCR (10.6% vs 9.4%) and further so in anti-Cit-Fib (15.7%) positive cells. Furthermore, VH SHM associated N-linked glycosylation sites were increased in RA (IgG > 15 mutations, $p<0.0001$, 17.2% vs 13.8%) and especially in Cit-Fib populations (32.2%). Previous studies have shown that anti-citrulline autoreactivity in RA is characterized by high somatic mutations and N-glycosylation sites. However, here the largest B-cell distortions in seropositive RA were observed in the unmutated B cells. These differences could reflect baseline shifts and elevated natural autoreactivity as an underlying mechanism in RA pathogenesis.

208

Phage Display Technology at UCB: Ongoing innovation with a mature platform

Anastasios Spiliotopoulos, Anthony Scott-Tucker, Laura Starkie, Chris Grice, James Snowden, Michael Wright.

Antibody ‘display’ using filamentous bacteriophage is a well characterized discovery and engineering platform that has been used successfully for over 30 years by many companies and academic groups. Although the technology is ‘mature’, here at UCB we are developing new and innovative ways to apply phage display to impact both internal research efforts and foster collaborations with external groups. This poster highlights a few of the novel approaches being developed by the display technology group in Discovery Biology.

209

Functional maturation and self-reactivity of the human B cell system assessed by B-cell receptor repertoire sequencing

Marie Ghraichy[1,2], Valentin von Niederhäusern[1,2], Jacob D. Galson[1,2], Johannes Trück[1,2]

[1]University Children's Hospital Zürich, Switzerland
[2]Children's Research Center, University Children's Hospital, University of Zürich, Switzerland

Background:
B cells play a central role in adaptive immune processes, mainly through the production of antibodies. Children are born without having had much contact with foreign antigens and are initially protected by maternal antibodies. Through continuous antigen exposure, the human immune system builds a repository of cells bearing diverse antigen-specific adaptive immune receptors that enable a targeted, rapid and extensive secondary immune response.
Little is known about the speed and magnitude or the detailed characteristics of this antigen-driven maturation of the B-cell system throughout childhood.

Methods:
We investigated the naïve and antigen-experienced B-cell receptor (BCR) repertoire in 46 healthy individuals aged 6m to 50y. Heavy chain BCR transcripts were amplified and sequenced and data analysis was performed with an in-house bioinformatic pipeline to assess repertoire characteristics and the self-reactive and structural nature of BCR transcripts.

Results:
The final dataset consisted of ~7M unique sequences (~150K sequences/individual). In the first 10 years of life, frequencies of highly mutated transcripts greatly increased through positive selection. These changes were accompanied by an increased usage of more downstream constant region genes (IgG2, IgA2) and a decrease in the frequency of transcripts with self-reactive properties indicating that somatic hypermutation has driven specificity of these sequences away from self. Structural analysis showed a higher frequency of antibodies, whose shapes differed from germline, with increasing age.

Conclusions:
This study demonstrates an extensive maturation of the B-cell system in the first 10 years of life as a consequence of environmental antigen exposure. Further antibody repertoire alterations continue to be made thereafter, although at a lower rate. This study also provides a reference data set of BCR repertoires and stresses the importance of using well-selected, age-appropriate controls in future studies.

210

The role of T cells in Neuromyelitis optica disease pathogenesis

Maren Lindner[1], Urvashi Bhatia[1], Andreas Schulte-Mecklenbeck[1], Timo Wirth[1], Tyge Schmidt[1], Catharina Gross[1], Melanie Korsen[1], Nicholas Schwab[1], Tania Kümpfel[2], Ingo Kleiter[3], Norbert Goebels[4], Anne Winkler[5], Christine Stadelmann[5], Wolfgang Brück[5], Heinz Wiendl[1], Tanja Kuhlmann[6], Marius Ringelstein[4], Luisa Klotz[1]

[1] Department of Neurology, University Hospital Münster, Münster, Germany
[2] Department of Neurology, University of Munich, Munich, Germany
[3] Department of Neurology, St. Josef Hospital, Ruhr University Bochum, Bochum, Germany
[4] Department of Neurology, Medical Faculty, Heinrich Heine University, Düsseldorf, Germany
[5] Institute for Neuropathology, Göttingen, Germany,
[6] Institute for Neuropathology, University of Münster, Münster, Germany

Neuromyelitis optica (NMO) is a rare autoimmune disease of the CNS with lesion development predominantly in the optic nerve and spinal cord. Most of the patients also exhibit an auto-antibody response against Aqaporin-4 (AQP4), a water channel on astrocytes. The pathogenicity of the AQP4 antibody as well as B cells themselves have been in the focus of recent research and seem to be a major driving force of disease pathogenesis. However, it could be demonstrated, that T cells from NMO patients show a higher reactivity towards AQP4 peptides, but not much is known about the role of T cells during lesion development and disease pathogenesis. Therefore, we characterized in detail immune cells in the periphery and the CSF of NMO patients in order to gain more insight into their role in NMO disease pathogenesis.

PBMCs from NMO patients were immune - phenotyped by multi parameter flow cytometry and compared to healthy donors and MS patients. In addition, the T cell receptor (TCR) repertoire of CD4+ T cells and CD8+ T cells purified from peripheral blood, CSF and CNS tissue was analyzed by deep sequencing of the beta chain of the CDR3 region.

Immune phenotyping revealed distinct differences in immune cell subsets in the NMO cohort compared to MS patients or healthy controls. In detail, NMO patients have more terminal differentiated T cells, which also show higher cytolytic activity. Moreover, TCR sequencing showed a more clonal repertoire in the CD4 as well as in the CD8 compartment. TCR repertoires from NMO patients were highly perturbed and differed significantly from MS patients or healthy controls. Clonal expansion was also prominent in the CSF as well as in CNS tissue and seems to be caused by local antigen-driven processes.

Our data confirm, that the role of T cells in NMO disease pathogenesis is more substantial than previously thought and further studies providing more insight into the mechanisms will potentially lead to new treatment options in NMO.

211

Characterization of the bovine adaptive immune repertoire responses to a candidate vaccine for East Coast Fever

Sonal Henson, Samuel O. Oyola, Elizabeth Kibwana, Benjamin Nzao, Vish Nene

International Livestock Research Institute, Nairobi, Kenya

In order to develop effective vaccines against pathogens, there is a need to assess the immunogenicity, not only by quantifying antigen-specific antibodies, but also by characterizing the diversity of the activated B-cells that generate functional antibodies, through B-cell receptor characterization. High-throughput sequencing (HTS) technologies provide the ability to study antigen receptor repertoires at both single cell and single nucleotide resolution. We are using HTS to qualitatively and quantitatively study bovine immune responses upon immunization with vaccine candidates against Theileria parva, the causative agent of East Coast Fever in cattle, with the aim to identify antigen-specific antibodies. We compare the characteristics of naïve and immunized bovine B cell repertoires focusing on VH genes with ultra-long CDRH3. We observe clonality, preferential germline gene usage and skewed CDRH3 amino acid properties in the immunized repertoires. We also identify putative antigen-specific B cell sequences with ultra long CDHR3.

212

Dynamics of age-based T-cell receptor repertoire derived from large-scale healthy population

Wei Zhang[1,2], Jinghua Wu[1,2], Longlong Wang[1,2], Tao Li[1,2], Xiao Liu[1,2]

[1]BGI-Shenzhen, Shenzhen, 518083, China
[2]China National GeneBank-Shenzhen, BGI-Shenzhen, Shenzhen 518083, China

The diversity and dynamics of T cell receptor (TCR) repertoire largely determine the ability to effectively maintain homeostasis of individuals in health and disease. As people grow older, the repertoire would be altered due to the involution of thymus and more kinds of infection. However, it is still not known the exactly change of TCR repertoire with age. In this study, we have recruited thousands of healthy donors, age from 0 to 80. We have employed deep TCR repertoire sequencing from peripheral blood to explore the characteristics of repertoire. We found the diversity of TCRs is gradually declined as age increased, and there are different features between male and females. Based on our data, we propose a healthy baseline of TCR repertoire, which may provide the helps to immunity assessment and disease research.

213

Analysis of transgenic mouse antibody repertoires

Aaron Chevalier*, Chris Murawsky*, John Bergen*, Karyn McFadden*, Yax Sun*

*Departments of Therapeutic Discovery, Amgen Inc

Tracing lead candidate antibodies back through their precursor repertoires can help elucidate metrics indicative of discovery success. Using NGS, we sequenced antigen-experienced and antigen-enriched hybridoma pools generated from immunized transgenic mice as part of an antibody discovery campaign. Bioinformatic analysis of the heavy chains identified tens of thousands of unique antibody clones clustered into hundreds of antibody lineages. We observed half of those lineages enriched in the antigen selected pools. We traced lead candidate antibodies back to their respective lineages and found they most commonly arose from the largest, most evolved lineages. In a second study, we NGS-sequenced directly the primary B cells of both immunized and non-immunized transgenic mice and analyzed the differences. Incorporation of UMIs and spike-in controls resulted in significantly increased data reproducibility and calibration.

214

T-cell receptor repertoires in peripheral blood encode type 1 diabetes status

Keshav Motwani[1], Milena Pavlovic[2,3], Laura Jacobsen[4], Geir K. Sandve[3], Victor Greiff[2], Todd Brusko[1]

[1]Department of Pathology, Immunology and Laboratory Medicine, University of Florida, Gainesville, FL, USA.
[2]Department of Immunology, University of Oslo, Oslo, Norway.
[3]Department of Informatics, University of Oslo, Oslo, Norway.
[4]Department of Pediatrics, University of Florida, Gainesville, FL, USA.

Organ-specific autoimmune diseases are often characterized by histological evidence for lymphocytic infiltration in target tissues. The progressive destruction of pancreatic islet β-cells that occurs in type 1 diabetes (T1D) is mediated by a coordinated adaptive immune response. Autoreactive T and B cells have been detected within islets and the emergence of multiple β-cell directed autoantibodies are highly predictive of disease progression. Despite robust serological biomarkers, no biomarkers exist that can reliably monitor disease-associated autoreactive T cells during disease progression. We hypothesized that the T-cell receptor (TCR) repertoire may contain sufficient information to foretell disease risk. To test this hypothesis, we sequenced the PBMC TCRβ repertoire of 1600 individuals from a cross-sectional cohort comprised of individuals with, or at varying degrees of risk for the development of T1D. We were able to separate patients by presence of autoimmunity using hierarchical clustering on clonal overlap between TCR repertoires in samples based on Morisita index and quantification of clonal expansion based on TCR diversity profiles. These findings suggested that T1D patients share common TCR-encoded signatures on both frequency (clonal expansion) and the sequence level (clonal sequence). To substantiate these claims, we used Fisher's exact test to identify those clones that were disproportionately present in T1D samples. Notably, previously published clones reactive to T1D-antigens were present with equal incidence in the PBMC samples across T1D and control groups. Our preliminary findings suggest that PBMC CDR3β signatures exist that distinguish between autoimmune positive (T1D) and healthy controls. Furthermore, we showed that incidence of published autoreactive sequences is not increased in the peripheral blood of T1D samples versus controls. Most interestingly, even without feature selection, clonal overlap and clonal expansion statistics are able to distinguish autoimmunity thereby underlining the potential disease-discriminatory role of public TCR sequences.

Funding: Helmsley Charitable Trust Team Science Award, NIH P01 AI42288, and National Institutes of Health (NIH) Human Islet Research Network (HIRN) UC4 DK104194 and U01 DK104162.

215

Prediction of cytomegalovirus serostatus from the T cell repertoire using deep learning

Vadim I. Nazarov[a,b], Evgeny Ofitserov[b,c] and Vasily Tsvetkov[d,e]

[a]National Research University Higher School of Economics, Russia
[b]ImmunoMind, Tallinn, Estonia
[c]Tula State University, Russia
[d]Shemyakin-Ovchinnikov Ins#tute of Bioorganic Chemistry of the Russian Academy of Sciences, Russia
[e]Pirogov Russian Na#onal Research Medical University (RNRMU), Russia

Immune repertoire is the combination of BCR and TCR that constitute the adaptive immune system. Since there are up to 4 × 1011 only T cells in a human organism, immune sequencing or immune repertoire profiling yields large quantities of poorly structured data. Classical machine learning algorithms fail to learn due to the complexity of given data. On the other hand, Deep Learning algorithms allow for the analysis of huge amounts of disparate information. In this work we have developed a Deep Neural Network (DNN) to extract CMV-specific motifs and to infer the CMV serostatus from immune repertoire. The DNN exploits motif-based approach that has proven particularly effective in protein classification task.

Firstly, the DNN derives substrings of CDR3 sequences of TCR receptors, which are implied to target CMV antigens. Then it applies logistic regression to motif weights and finally predicts the probability of CMV infection. We trained our model on the TCR repertoire dataset with known CMV serostatus from previously published works. A series of experiments was conducted to tune the hyperparameters of the model, which are the number of motifs implied (32 to 512) and their length (4 to 6). The best hyperparameters were chosen using cross-validation. With 256 motifs of length 5 as hyperparameters, the model has achieved 0.97 AUC on train dataset, 0.90 on test data. Although the previously published classification model is slightly be]er with 0.99 and 0.94 respectively, our model has extracted a limited set of CMV-specific TCR motifs that can be further confirmed in wet lab experiment. Since our model has achieved 0.97 AUC, there is still room for improvement. The accuracy of prediction can be increased due to use of L1/L2 regularisation and dropout layers and the model can be further enhanced to detect gapped motifs.

216

Antibody-Extractor®, the in-silico antibody sequence discovery platform

Khalil El Mazouari

Antibody-Extractor is a suite for scientists and API for developers, active in the field of antibody discovery. Immunology is one of the key areas in which NGS / Deep Sequencing plays a significant influence. The sequencing of an entire Antibody Repertoire provides an in-depth view regarding the diversity of in-vivo produced and selected antibodies and can generate new insights into the results obtained by traditional expression system as phage display.
More than two decades of expertise in Antibody discovery was captured as algorithms and implemented in API using the most recent technologies from High-Frequency Trading industry to perform High-Frequency Sequences Analysis. Antibody-Extractor© algorithms are developed in-house and successfully used for antibody sequence discovery, rationalization of in-vitro maturation and repertoire NGS analysis. Typical selection processes of antibody leads are restricted by the biology of the expression system used, antibody discovery teams will most often focus on the most abundantly expressed clones and miss out on a significant part of the library diversity. Rarely expressed clones will often not be picked up or will vanish during subsequent panning rounds.
Based on our strong interactions with different antibody discovery teams, We present three antibody use cases :
1. efficient processing of phage-display sequences: annotation, families identification, multiple alignments, germaligning, hotspot identification, protein calculations, etc.
2. selecting binder candidate from NGS: CDR3 clustering, extracting highly expressed antibody chains.
3. combining phage-display and NGS for rationalization of in-vitro maturation: variants extraction, mutation sites identification, mutation results analysis.

217

Novel in vitro methods for the discovery of functional T-cell receptor epitopes from large peptide-coding libraries

Govinda Sharmaa, Z. Alia, C.M. Rivea, S. Sivasothy[a], R. A. Holt[a,b,c]

[a] Canada's Michael Smith Genome Sciences Centre, BC Cancer Agency, Canada
[b] Department of Medical Genetics, University of British Columbia, Canada
[c] Department of Molecular Biology and Biochemistry, Simon Fraser University, Canada

The ability to rapidly and deeply search peptide space to determine specific peptide epitopes that are naturally processed, presented, and capable of eliciting functional cytotoxic T-cell (CTL) responses is a critical unmet need in the study of adaptive immunity. We describe a novel method for high-throughput CTL epitope profiling that enables simultaneous in vitro interrogation of target cell populations encoding diverse peptide-coding "minigene" libraries with CTL populations-of-interest. Target cells that elicit CTL reactivity are selectively isolated by FACS, using a FRET-based sensor of T-cell cytotoxicity, and are subject to deep amplicon sequencing to reveal the putatively antigenic minigenes encoded within. We have extensively validated this approach using known murine T-cell receptor (TCR)/peptide-MHC pairs and have shown that this approach can unambiguously identify canonical minigenes from libraries of vastly more candidate antigens in parallel than would be tractable using conventional methods.

We extended the capability of this strategy by applying a synthetic biology approach. Using pairs of immortalized natural killer (NK)-like effector cell lines and naturally tolerated target cell lines, we have shown that fully reconstituting the TCR/CD3 complex in effectors and expressing relevant MHC-/minigene-coding sequences in targets is sufficient to direct the cytotoxic response of the host NK-like cell line in a TCR antigen-specific manner. These results provide indication that it should be possible to use an entirely synthetic framework for functionally screening recombinant TCR-of-interest against minigene libraries without the requirement to first isolate and expand primary CTL clones or donor-derived antigen-presenting cells. Current work is aimed at continuing to benchmark this system and, if successful, it will provide a means to make any CTL with a literature-documented TCR sequence accessible for TCR antigen discovery studies.

This work was supported by Genome Canada, Genome BC, the BC Cancer Foundation, BioCanRx (Biotherapeutics for Cancer Treatment), and the National Cancer Institute of the US National Institutes of Health under award number 1R21CA226321-01.

218

IgE B cell responses to environmental factors in infancy and childhood

Sandra C. A. Nielsen[1], Krishna M. Roskin[1], Katherine J. L. Jackson[1], Ramona A. Hoh[1], Ji Yeun Lee[1], Catherine Ley[2], Julie Parsonnet[2,3], Scott D. Boyd[1]

Departments of [1]Pathology, [2]Medicine and [3]Health Research and Policy, Stanford University, Stanford, CA 94305, USA

Background:
Antigenic exposures at epithelial sites in infancy and early childhood are thought to influence the maturation of humoral immunity and modulate the risk of developing IgE-mediated allergic disease. How different kinds of antigens influence class-switching to IgE, IgG, or IgA, and the somatic hypermutation (SHM) maturation of these antibody pools, is not fully understood.

Methods:
Stanford's Outcomes Research in Kids is a multiethnic cohort, which follows children from birth through age 3 years. We have studied a sub-cohort of 51 children with up to three longitudinal time points per subject. For each sample, libraries of IGHV gene rearrangements were prepared and high-throughput sequenced.

Results:
IgE antibody genes are already significantly mutated in the first year of life, distinguishing this isotype from IgG and IgA switched isotypes that undergo more gradual and progressive mutation increases during the first three years.
IgE SHM is primarily increased in infants with impaired skin barrier conditions such as eczema. We found that IgM was the most common isotype found in IgE-expressing clonal lineages (IgE+) in healthy subjects, whereas IgA was the most common isotype found in IgE+ lineages of children with eczema or allergic diagnoses.

Conclusions:
The results highlight the differences in the mutational responses of IgE+ B cells compared to those expressing other switched subtypes. IgE+ B cells show elevated SHM in children with eczema, suggesting a key mechanistic link between impaired skin barrier function and IgE SHM that could contribute to allergic sensitization and the development of pathogenic IgE in allergic individuals.

Funding sources:
Child Health Research Institute and the Stanford NIH-NCATS-CTSA (grant no. UL1 TR001085)
Ulla og Mogens Folmer Andersens Fond
Crown Family Foundation

219

RAbHIT: R Antibody Haplotype Inference Tool

Ayelet Peres, Moriah Gidoni, and Gur Yaari

Faculty of Engineering, Bar Ilan University, Ramat Gan, Israel

Analysis of antibody repertoires by high throughput sequencing is of major importance in understanding adaptive immune responses. Our knowledge of variations in the genomic loci encoding antibody genes is incomplete, mostly due to technical difficulties in aligning short reads to these highly repetitive loci. The partial knowledge results in conflicting V-D-J gene assignments between different algorithms, and biased genotype and haplotype inference. Previous studies have shown that haplotypes can be inferred by taking advantage of IGHJ6 heterozygosity, observed in approximately one third of the population.

We created a robust novel method for determining V-D-J haplotypes by adapting a Bayesian framework. Our method extends haplotype inference to IGHD, IGHV, IGKJ, IGKV, and IGLV based analysis, thereby enabling inference of complex genetic events like deletions and copy number variations in the entire population. Based on this method we developed an R package, which implements the method on sequences from naive B-cells, for both the heavy and the light chains. The package offers a haplotype and single chromosome deletion inference based on an anchor gene. The inferred haplotypes and deletion patterns may have clinical implications for genetic predispositions to diseases.

AIRR Community Meeting IV Attendees

Coming soon!

AIRR-C Community Sub-committees & Working Groups Members as of May 1, 2019

Sub-committees

Communications Sub-committee
Victor Greiff (Co-leader, University of Oslo, Norway)
Jean-Philippe “JP” Bürckert (Co-leader, BISC Global Inc., USA)
Chaim Schramm (Vaccine Research Center, NIAID, NIH, USA)
Jamie Scott (Simon Fraser University, Canada)

Executive Sub-committee
Felix Breden (Chair, Simon Fraser University, Canada)
Scott Christley (The University of Texas Southwestern Medical Center, USA)
Danny Douek (NIH NIAID Vaccine Research Center, USA)
Steven Kleinstein (Yale University Medical School, USA)
Nina Luning Prak (Chair-elect, University of Pennsylvania Medical School, USA)

Inferred Allele Review Sub-committee (IARC)
Andrew Collins (Co-leader, University of New South Wales, Australia)
Martin Corcoran (Karolinska Institute, Sweden)
Mats Ohlin (Lund University, Sweden)
Cathrine Scheepers (National Institute for Communicable Diseases, South Africa)

Meeting Sub-committee
Davide Bagnara (University of Genoa, Italy)
Pam Borghardt (Co-leader, Simon Fraser University, Canada)
Jean-Philippe “JP” Bürckert (BISC Global Inc., USA)
Ramit Mehr (Bar Ilan University, Israel)
Jamie Scott (Co-leader, Simon Fraser University, Canada)

Working Groups
*Ongoing update

Biological Resources Working Group*
Davide Bagnara (University of Genoa, Italy)
Maggie Bostick (Co-leader, Takara Bio, USA)
Christian Busse (German Cancer Research Center, Germany)
Uri Hershberg (University of Haifa, Israel)
Nina Luning Prak (University of Pennsylvania Medical School, USA)
Encarnita Marrioti-Ferrandiz (Co-leader, Sorbonne University, France)
Jacob Sherkow (New York Law School, USA)
Johannes Trück (University Children's Hospital Zürich, Switzerland)

Common Repository Working Group*
Syed Ahmad Chan Bukhari (Yale School of Medicine, USA)
Meredith Ashby (Pacific Bio Sciences, USA)
Felix Breden (Simon Fraser University, Canada)
Jamie Scott (Simon Fraser University, Canada)

Tania Bubela (Simon Fraser University, Canada)
Scott Christley (The University of Texas Southwestern Medical Center, USA)
Brian Corrie (Co-leader, Simon Fraser University, Canada)
Lindsay Cowell (Co-leader, The University of Texas Southwestern Medical Center, USA)
John Harting (Pacific Bio Sciences, USA)
David Klatzmann (Sorbonne University, France)
Uri Laserson (Icahn School of Medicine at Mount Sinai, USA)
Nishanth Marthandan (Simon Fraser University, Canada)
Bjoern Peters (Center for Infectious Disease and University of California San Diego, USA)
Adrien Six (Sorbonne University, France)
Adrian Thorogood (McGill University, Canada)
Corey Watson (University of Louisville, USA)
Yariv Wine (Tel Aviv University, Israel)
George Blanck (Morsani College of Medicine, USA)

Data Representation Working Group*
Felix Breden (Simon Fraser University, Canada)
Scott Christley (Co-leader, The University of Texas Southwestern Medical Center, USA)
Syed Ahmad Chan Bukhari (Yale School of Medicine, USA)
Brian Corrie (Simon Fraser University, Canada)
Jessica Finn (Vanderbilt Vaccine Center, USA)
Anna Fowler (University of Oxford, USA)
Daniel Gadala-Maria (Yale University, USA)
Jerome Jaglale (Simon Fraser University, Canada)
Steven Kleinstein (Yale University Medical School, USA)
Uri Laserson, Co-leader (Icahn School of Medicine at Mount Sinai, USA)
Susanna Marquez (Yale School of Medicine, USA)
Nishanth Marthandan (Simon Fraser University, Canada)
Duncan Ralph (Fred Hutchinson Cancer Research Center, USA)
Aaron Rosenfeld (Drexel University, USA)
Chaim Schramm (Vaccine Research Center, NIAID, NIH, USA)
Jason Vander Heiden (Yale University, USA)
Corey Watson (University of Louisville, USA)
Bojan Zimonja (Simon Fraser University, Canada)
Christian Busse (German Cancer Research Center, Germany)
Frederick A. Matsen IV (Fred Hutchinson Cancer Research Center, USA)

Minimal Standards Working Group
Christian Busse (Co-leader, German Cancer Research Center, Germany)
Florian Rubelt (Co-leader, Roche Sequencing Solutions, USA)
Brian Corrie (Simon Fraser University, Canada)
Bojan Zimonja (Simon Fraser University, Canada)
Corey Watson (University of Louisville, USA)
Eline Luning Prak (University of Pennsylvania Medical School, USA)
Felix Breden (Simon Fraser University, Canada)
Jamie Faison (Duke University, USA)
Jean-Philippe “JP” Bürckert (BISC Global Inc., USA)
Jerome Jaglale (Simon Fraser University, Canada)
John Graybeal (Stanford University, USA)
Lindsay Cowell (The University of Texas Southwestern Medical Center, USA)

Enkelejda Miho (aiNET GmbH, CH)
Nina Luning Prak (University of Pennsylvania Medical School, USA)
Nishanth Marthandan (Simon Fraser University, Canada)
Susanna Marquez (Yale School of Medicine, USA)
Scott Christley (The University of Texas Southwestern Medical Center, USA)
Steven Kleinstein (Yale University Medical School, USA)
Syed Ahmad Chan Bukhari (Yale School of Medicine, USA)
Uri Laserson (Icahn School of Medicine at Mount Sinai, USA)

Software Working Group*
Chaim Schramm (Co-Leader, Vaccine Research Center, NIAID, NIH, USA)
William Lees (Co-Leader, University of London, UK)
Branden Olson (University of Washington, USA)
Duncan Ralph (Fred Hutchinson Cancer Research Center, USA)
Frederick “Erick” Matsen (Fred Hutchinson Cancer Research Center, USA)
Inimary Toby (University of Dallas, USA)
Jason Vander Heiden (Yale University, USA)
Mikhail Shugay (Skolkovo Institute of Science and Technology, Russia)
Jian Ye (National Institutes of Health, USA)
Enkelejda Miho (aiNET GmbH, CH)
Pejvak Moghimi
Anna Obraztsova (Skolkovo Institute of Science and Technology, Russia)
Adrian Shepherd (University of London, UK)

Germline Database Working Group
Corey Watson (Co-leader, University of Louisville, USA)
Andrew Collins (Co-leader, University of New South Wales, Australia)
Scott Christley (UT Southwestern Medical Center, USA)
Bruno Gaeta (University of New South Wales, Australia)
Felix Breden (Simon Fraser University, Canada)
Brian Fritz (10X Genomics, USA)
Chaim Schramm (Vaccine Research Center, NIAID, NIH, USA)
Christian Busse (German Cancer Research Center)
Daniel Gadala-Maria (Yale University, USA)
Davide Bagnara (University of Genoa, Italy)
Deanna Church (10X Genomics, USA)
Florian Rubelt (Roche Sequencing Solutions, USA)
Duncan Ralph (Fred Hutchinson Cancer Research Center, USA)
Gur Yaari (Bar Ilan University, Israel)
Jamie Faison (Deloitte Consulting, USA)
Jean-Philippe “JP” Bürckert (BISC Global Inc., USA)
Jian Ye (National Institutes of Health, USA)
Jamie Scott (Simon Fraser University, Canada)
Justin Kos (University of Louisville, USA)
Katherine Jackson (Garvan Institute, Australia)
Kevin Wu (10x Genomics, USA)
Martin Corcoran (Karolinska Institute, Sweden)
Mats Ohlin (Lund University, Sweden)
Melissa Smith (Icahn School of Medicine at Mount Sinai, USA)
Nishanth Marthandan (Simon Fraser University, Canada)

William Rounds (UT Southwestern Medical Center, USA)
Steven Kleinstein (Yale University Medical School, USA)
Victor Greiff (University of Oslo, Norway)
William Gibson (University of Louisville, USA)
Susana Magadan (University of Vigo, Spain)
Pierre Boudinot (French National Institute for Agricultural Research, France)
Henk-Jan van der Ham (Enpicom, The Netherlands)
Oscar Rodriguez (Icahn School of Medicine at Mount Sinai, USA)

AIRR-C Committee Sub-committees and Working Groups Reports

AIRR-C Executive Sub-committee Report – May 1st, 2019

Current:
SC/WG Name: Executive Sub-committee
SC/WG Co-Leaders: Felix Breden (Chair) and Nina Prak (Chair-Elect)
SC/WG Members: Danny Douek, Steve Kleinstein, Scott Christley
Purpose/Mission: The Executive Sub-committee will provide leadership for the AIRR Community.

Current Goals:

- Review all publications presented to research public on behalf of AIRR Community. Conduct election of Chair-elect and members of AIRR Executive Sub-committee.
- Interact with international Immunological Societies, including IUIS (International Union of Immunogical Societies).
- Meet with leaders of Sub-committees and Working Groups as one body. Integrate AIRR Community with The Antibody Society (TAbS)

Progress report on current goals:

- Website, Meetings committees, and governance structure integrated with TAbS Oversaw publication of AIRR Community overview publication (Breden et al., 2017, Frontiers in Immunology:8:1418), MiAIRR standards (Rubelt et al., Nature Immunology, 18:1274-1278), Data Representation standards (Vander Heiden et al., Front. Immunol., 9:2206), and Protocols for gene inference from Germline Database Working Group and IARC (Inferred Allele Review Committee) (Ohlin et al., Front. Immunol. 10:435).
- Met with Working Group and Sub-committee leaders January 2019. Conducting elections for Executive Committee 2019.
- Drafting new governance structure

Future:
Proposed goals and work plan for the coming interval:

- Adopt new Governance Structure and integrate with TAbS
- Oversee change in Minimal Standards Working Group to Sub-committee performing oversight of standards for publishing and submitting AIRR-seq data sets to public repositories.
- Oversee formal ratification of IARC and Communications Sub-committee.
- Further integration with TAbS including adding representation of AIRR Community members on TAbS Board of Directors, Meetings Committee, and Budget/Financing Committee.

Proposed SC/WG leadership (co-leaders) and members:
Chair-elect and members of Executive Sub-committee to be determined by vote of AIRR Community conducted by Executive Sub-committee.

AIRR-C Meeting Sub-committee Report – May 1st, 2019

Current:
SC/WG Name: AIRR-C Meeting Sub-committee
SC/WG Co-Leaders: Jamie Scott and Pam Borghardt
SC/WG Members: Jamie Scott, Ramit Mehr, Davide Bagnara, Jean-Philippe Bürckert and Pam Borghardt
Purpose/Mission:
(Please refer to your SC's or WG's description of mission/purpose at: https://www.antibodysociety.org/the-airr-community/ and add revisions if necessary)

The AIRR-C Meeting Sub-committee (SC) is responsible for the initiation and planning of AIRR-C-related meetings.

Current Goals:
(Please refer to your SC's or WG's description of goals for 2018 at: https://www.antibodysociety.org/the-airr-community/ and add revisions if necessary)

- The AIRR-C Meeting SC will comprise 5 members, including 2 Co-Leaders who are chosen by the SC members after its membership is ratified by the AIRR-C. One of the Co-Leaders must be an AIRR-C member.
- The term of AIRR-C Meeting SC members can be renewed for up to 3 consecutive terms.
- The membership of the AIRR-C Meeting SC will be ratified at the AIRR-C Meeting.
- The AIRR-C Meeting SC is responsible for planning scientific sessions, workshops and demonstrations for all AIRR-C-related meetings.
- The AIRR-C Meeting SC is responsible for planning a meeting of the AIRR-C every 12-18 months.

At that Meeting:

- The AIRR-C Working Groups and SCs (including the Meeting SC) will report on their progress and seek ratification for their plans.
- The AIRR-C will vote on these and other Community-related matters.
    - The TAbS Scientific Advisory Board plans the Annual Antibody Engineering & Therapeutics Meeting (San Diego).
    - One member of the AIRR-C Meeting SC will represent the AIRR-C on the TAbS Meeting Committee and the Scientific Advisory Board.
    - The AIRR-C Meeting SC will plan one session of the Antibody Engineering & Therapeutics Meeting (San Diego) each year. The AIRR-C Meeting SC appointee to the TAbS Meeting Committee and Scientific Advisory Board will lead the planning of that session.

Progress report on current goals:

- In collaboration with the AIRR Community the AIRR-C Meetings SC has planned all aspects of the AIRR-C Meeting IV, to be held at the University of Genoa, 11-15 May 2019.
- Meetings – On average, the AIRR-C Meetings SC met weekly over the past 18 months in order to coordinate the Meeting.
- Agenda – As of May 1st, 2019, that Meeting will comprise: 134 Registrants, 2 Workshops, 6 Working Group Reports, 3 SC Reports, 3 Panel Discussions, 9 AIRR-seq Tool Demonstrations, 12 Pipeline & Data Repository Tutorials, 37 Posters, 2 Scientific

Sessions with each: 1 Keynote, 4 Oral Presentations, 3 Short Presentations, 3 challenge sessions with each 3 Oral Presentations, Industry/Sponsor reception; 14 Sponsors.

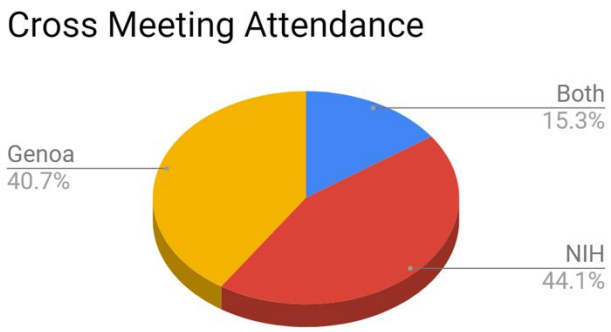
- Budget – Acquired sponsor contributions (USD): 2x Grant Funding (~$10k), 5x Silver Level ($5k), 8x Bronze Level (6x $2.5k + 1x $1k). With the collection of registration fees, the 2019 AIRR-C Meeting IV will be on budget.
- Statistics – See Appendix A.
- The AIRR-C Meetings SC has begun planning AIRR-C Meeting V.
- The AIRR-C Meetings SC has begun planning the AIRR-C session of the Antibody Therapeutics Meeting (Dec 2020, San Diego, CA.)

Future:
Proposed goals and work plan for the coming interval:

- Choose Co-Leaders after Meetings SC membership is ratified by the AIRR-C.
- Appoint the AIRR-C Meetings SC member who will sit on the TAbS Meeting Committee and the Scientific Advisory Board, and lead planning of an AIRR-C session for the annual Antibody Engineering & Therapeutics Meeting (Dec 2020, San Diego, CA.)
- Plan the upcoming AIRR-C Meeting V.
- Consider Fall 2020:
    - This time it should be in North America. Consider holding it at the University of Toronto.
    - Potential themes:
        - Engaging/integrating with other types of Big Immunological Data (Systems Immunology)
        - Provisional title: AIRR Community V: “Exploring New Frontiers”
        - Sustaining the AIRR Community initiatives in the future.
- Proposed SC/WG members:
    - Davide Bagnara, Pam Borghardt (Meeting Manager), Jean-Philippe Bürckert, Ramit Mehr, Jamie Scott.
    - May bring on/engage a local contact once the location is determined.

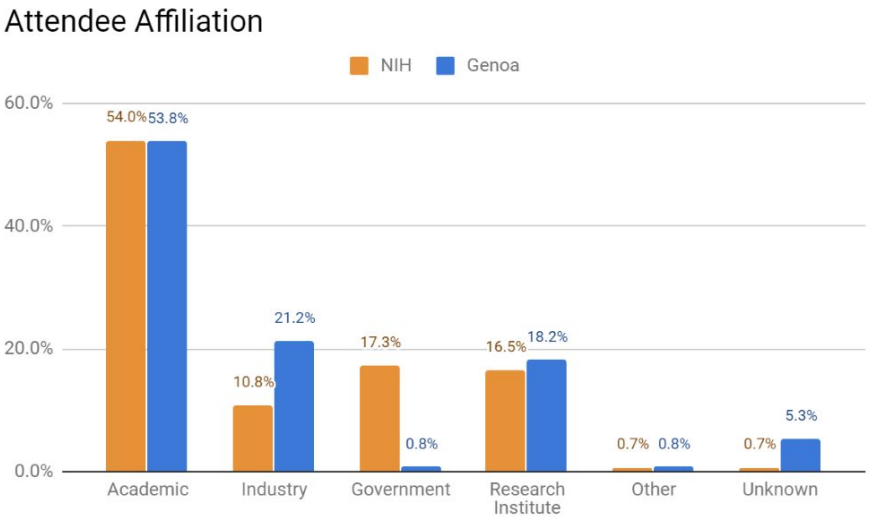
Appendix A:

AIRR-C Meeting IV Registration Statistics (as of April 30, 2019)

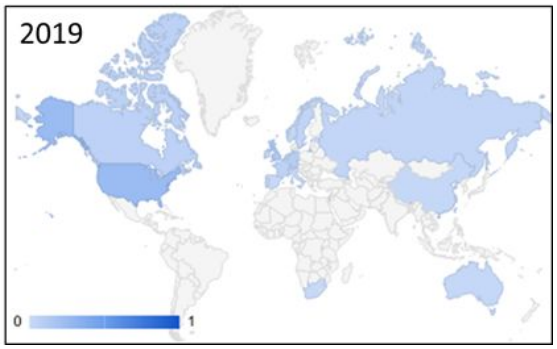


New members: The current registrants are largely new, non-AIRR-C members. This demonstrates a successful outreach together with the efforts of AIRR-C Communications SC. Of all attendees across the upcoming and previous meeting only 15% will have attended both

meetings. This corresponds to approximately 30 “core attendees” and thus about 100 new attendees.



Affiliation: Similar to 1. AIRR-C Meetings SC was successful in “bridging the gap” together with AIRR-C Communications SC, expanding attention towards Industry. We can report about twice as many attendees (2017: 21.2% vs 2019: 10.8%) from the industry sector this year. The transit was mostly from US Government institutions which are less likely to participate in a European-based meeting (see also 3.). The number of participants from both Academia and Research Institute are stable.



Country: While the last meeting’s attendees were largely US-based (2017: 73.2% vs 2019: 24.8%), this year’s meeting attracted more EU citizens (2017: 17.4% vs 2019: 51.1%, including the UK).
We are also delighted to report attendees from 10 new countries (Belgium, China, Switzerland, Taiwan, Russia, Portugal, Kenia, Ireland, Estonia and Spain) while we lost participation from Luxembourg and Japan compared to 2017.

AIRR-C Communications Sub-committee Report – May 1st, 2019

Current:
SC/WG Name: AIRR-C Communications Sub-committee
SC/WG Co-Leaders: Jean-Philippe Bürckert and Victor Greiff
SC/WG Members: Jean-Philippe Bürckert, Victor Greiff, Chaim Schramm, Jamie Scott
Purpose/Mission:
https://www.antibodysociety.org/airrc/working_groups/communications-sub-committee/

Current Goals:

- Establish the AIRR-C Communications Subcommittee
- Update and maintain AIRR-C web presence
- Facilitate AIRR-C internal and external communication and outreach
- Develop an AIRR-C "corporate identity"
- Ratification by the AIRR-C Community at the 4th AIRR-C Meeting in Genoa, Italy.

Progress report on current goals:

- Establishment of the AIRR-C Communications Subcommittee:
    - We've successfully put together a group of 4 AIRR-C members to lead the initiative of forming a communications subcommittee. This includes the setup of initially weekly, and later "as needed", meetings (8 in total so far since the first meeting on July 27, 2018) and mission statement.
    - Further the promotion of co-leaders and coordination of initial efforts. We've reported to the AIRR-C Executive Subcommittee and obtained preliminary ratification.

AIRR-C Web Presence:

- Organized the transition from previous web presence http://airr-community.org to the website of our umbrella society TAbS together with Brian Corrie. AIRR-C Comm SC made sure that the original content was transferred correctly and the new website is still reachable via airr-community.org, which was used in publications and thus ensuring link persistence.
- Restructured website to reflect the governance and community organization
- Established "AIRR Community" as visible section of TAbS and accessible part of the website, similar to "mAbs"
- Updated web content with up to date mission statements and WG/SC memberships
- Communication and Outreach:
    - Established the following email addresses together with AIRR-C Exec SC for each WG and SC and included them into the TAbS mail system. Provided "How-To" manuals for all WG/SC Co-leaders on how to use the system
        - exec@airrc.antibodysociety.org
        - comm@airrc.antibodysociety.org
        - meetings@airrc.antibodysociety.org
        - iarc@airrc.antibodysociety.org
        - datarep@airrc.antibodysociety.org
        - ministd@airrc.antibodysociety.org
        - commonrepo@airrc.antibodysociety.org
        - software@airrc.antibodysociety.org
        - germlines@airrc.antibodysociety.org
        - bioresources@airrc.antibodysociety.org

- Set up of AIRR-C Slack channel to facilitate internal communication (WIP) complementary to the BTCR forum (www.b-t.cr) which in return intersects internal and external communication. Internal sharing and coordination of news, code, positions, papers, outreach.
- Set up of AIRR-C twitter account to facilitate outreach and news propagation.
- Gathered all social media logins. All channels of AIRR-C on Youtube, Twitter, mailing lists and slack are managed by AIRR-C Comm SC
- Conducted direct outreach through presence and presentation at conferences and through collaborations:
  - Introducing companies to AIRRC and our work at conferences or through collaborations
  - Several biotech companies including top 5 pharma have been introduced to AIRR-C and data formats. Most of them are now also using
  - AIRR-compliant data analysis pipelines or are in the process of setting these up
  - Showed AIRR-slide at conferences to promote AIRR Conference in Genoa
  - Promotion of Genoa meeting and inviting industry sponsors

Future:
Proposed goals and work plan for the coming interval:
Now that main channels of communications for internal and external communications have been established, the AIRR-Communications SubCommittee has identified the following proposed goals.

- Extending and refining AIRR “corporate identity” and external dissemination and promotion:
  - Finalize AIRR-C logo (the logo will be chosen out of three suggestions by a vote from the AIRR-C present at the 4th AIRR Community Meeting in Genoa). The three logos from which the final one will be chosen are shown below.
  - Prepare slides and posters that disseminate the AIRR-C goals at relevant conferences | goal: one AIRR-C poster at every relevant conference
  - Standardize AIRR-C nomenclature on the website and in the AIRR-C in general: e.g. abbreviations (AIRRC, AIRR-C, AIRR C?!)
- Establish a quarterly/twice-a-year AIRR Newsletter:
  - A quarterly conference call with WG leaders for updates to get latest updates on progress made
- Transitioning of AIRR-C-Comm SC into TAbS-Comm SC:
  - Given the increasing importance of AIRR-Seq for antibody and TCR research as well as a missing communication structure within TAbS, we believe it beneficial for both TAbS and AIRR-C to join efforts and transition the AIRR-C-Comms SC into TAbS-Comms SC. Janice Reichert (TAbS) has already expressed interest in this direction.
- Update email communication structure:
  - The current emailing system used is archaic and needs updating. It is very hard to edit mailing lists and there is not an automatic way for people to subscribe or unsubscribe to the AIRR-C mailing list. A better emailing system is needed. The currently established one is fine for the time being but as the AIRR-C grows the current system will become unmaintainable.

- Extending the AIRR-C Comm SC team:
    - Pam Borghardt has played a pivotal role in the inter-SC communication between Exec, Meetings and Communications SCs. She is deeply involved in the administrative and governance tasks of the AIRR-C. Her knowledge and skills are crucial to our future missions such as the transit to a TAbS-Comm SC, website maintenance and outreach.
- Maintenance of communication structure:
    - One of the main tasks of the AIRR-C-Comms-SC will be to maintain current and future communication structures.
    - To facilitate transparency of the communication structure within AIRR-C, we also plan to devise a communication chart, which clearly outlines communication avenues and responsibilities.

Proposed SC leadership (co-leaders) and members:
Jean-Philippe Bürckert, Victor Greiff, Jamie Scott, Pam Borghardt

AIRR-C Committee Reporting Template for Sub-committees and Working Groups

Current:
SC/WG Name: Data Representation
SC/WG Co-Leaders: Uri Laserson, Scott Christley
SC/WG Members: Regularly active: Chaim Schramm, Ahmad Chan, Brian Corrie, Christian Busse, Jason Vander Heiden, Nishanth Marthandan, Susanna Marquez
In total there are 25 people on our mailing list.
Purpose/Mission:
(Please refer to your SC's or WG's description of mission/purpose at: https://www.antibodysociety.org/the-airr-community/ and add revisions if necessary)

This working group will be responsible for developing standardized names for data fields which can be understood and interpreted by all software tools, allowing interoperability between pipelines from different developers. Close collaboration with the Minimal Standards Working Group is expected.

The Data Representations working group is focused on developing standard file formats and schemas to represent annotated antibody and T cell receptor sequences and any downstream data representations. The proliferation of tools for processing raw AIRR data is making it more difficult to compare results between tools and to build modular data pipelines. We have been developing a CSV-like file format for representing annotated reads and clones, with the goal of having it implemented in multiple common AIRR pipelines (e.g., immcantation).

- Multiple WGs are designing implementation standards and could use technical input on data representation.
- Coordination with AIRR Working Groups to specify data models, e.g.,
- Common Repo defining minimal APIs for repositories and REST resources
- MinStd choosing ontologies for their fields
- Germline defining new germlines and annotations
- Ensure all AIRR groups are working in mutually compatible ways (in terms of data)
- Ensure we have liaisons on all other relevant working groups
- Work on representation of provenance of data sets

Current Goals:
(Please refer to your SC's or WG's description of goals for 2018 at: https://www.antibodysociety.org/the-airr-community/ and add revisions if necessary)
Our 2018 goals as presented in the last AIRR Community meeting (2017):

- Submit manuscript to publicize format
- Develop format for representing clones
- Finish integration of GitHub repository with MiAIRR
- Finish specifying metadata file format
- Public release of reference library to read/write/validate AIRR format files
- Initially targeting Python and R
- Releasing documentation incl. example output/use

Progress report on current goals:

- Submit manuscript to publicize format:

    - A manuscript was submitted and published in Frontiers in Immunology with AIRR Community as coauthor.
    - https://www.frontiersin.org/articles/10.3389/fimmu.2018.02206/full
- Develop format for representing clones/trees:
    - Currently handled purely by tracking clone ids in the rearrangement schema
    - We had some discussion around this and came up with a strategy for representing clones and trees. However, formalizing a spec has stalled, in part because we have not identified any champions to make use of such a standard.
- Finish integration of GitHub repository with MiAIRR:
    - This is complete. The repo continues to be actively developed for both purposes.
- Finish specifying metadata file format:
    - We have recently (early 2019) put together a draft spec for metadata closely tied to the CommonRepo APIs that are being developed.
- Public release of reference library to read/write/validate AIRR format files:
    - We have published Python and R packages for working with AIRR TSV data.
- Releasing documentation incl. example output/use:
    - Lots of work continues on the docs.airr-community.org site. Rudimentary example use is included with the libraries.
    - We have also built a general documentation resource that is used by more than just the DataRep group.

Future:
Proposed goals and work plan for the coming interval:

- Community outreach and documentation:
    - Goal: get people using our standards
    - Outreach to all common AIRR tools to promote/validate compliance
- Refinement of existing standards/documentation/tools:
    - Goal: "productionizing" our work to make it reliable and easy to use
    - Specs/software:
        - Refine/complete Repertoire Metadata file format specification
        - Integration tests to ensure latest versions of peoples' software packages remain AIRR-compliant
        - Contribute AIRR R package into Bioconductor
    - Documentation:
        - Ensure all relevant docs etc are auto-generated from the "master" YAML specification
        - Rewrite documentation with different target audiences in mind (e.g., data submitters, data analysts, tool developers).
        - User guide vs reference guide
        - Provide metadata templates for different experimental designs (e.g., cross-sectional, longitudinal, etc.). Provide some published datasets using our metadata format
- Enabling other working groups:
    - Goal: ensure consistency and interoperability across AIRR efforts
    - Contribute to germline file format representation (piggybacking on all the OGRE schema work). Downloadable reference databases.
    - Helping other groups contribute to central documentation site
- High-priority standards development
    - Clonality/tree representation
    - Integration of epitope/specificity data

Proposed SC/WG leadership (co-leaders) and members:
To be discussed at the meeting, but likely unchanged.

AIRR-C Committee Reporting Template for Sub-committees and Working Groups

Current:
SC/WG Name: Software Working Group
SC/WG Co-Leaders: Chaim Schramm, William Lees
SC/WG Members: Branden Olson, Duncan Ralph, Frederick “Erick” Matsen, Inimary Toby, Jason Vander Heiden, Mikhail Shugay, Jian Ye, Enkelejda Miho, Pejvak Moghimi, Anna Obraztsova, Adrian Shepherd
Purpose/Mission:
The goal of our working group is to encourage practices that enable software tools to work, and to work with one another. As such we have been assembling data sets that people can use to test functionality of various programs and are developing and promoting guidelines for software tools.

Our current first priority is standardized simulated data sets with known properties. We are working to define and implement summary statistics that can be used to characterize simulated data sets and compare them to real data sets. After that we can “benchmark the benchmarks” to decide how realistic the various simulations are.

Current Goals:
- Finalize list of summary statistics to use for validation of repertoire simulation tools and finish implementing these summary statistics in software
- Perform comparison of simulation tools using the summary statistics and write up results
- Publish guidelines for AIRR software tools

Progress report on current goals:
- Summary statistics code (Sumrep) is nearing completion (https://github.com/matsengrp/sumrep).
- WG members are applying Sumrep to simulated and real-world datasets.
- A paper on Sumrep is in preparation.
- Software tools guidelines have been developed. They emphasize the use of containers (e.g. Docker) to facilitate evaluation, and the use of AIRR standards (https://docs.airr-community.org/en/latest/swtools/airr_swtools_standard.html).

Future:
Proposed goals and work plan for the coming interval:
- Encourage better simulation via summary statistics:
    - Finish the initial release of Sumrep and publish a paper applying it to selected datasets.
- Evaluate annotation tools, using simulated and real-world data:
    - Identify simulated and real-world datasets that are useful for evaluation.
    - Develop a framework for the comparison of results.
- Encourage standard interchange formats:
    - Encourage tool providers to submit their tools for review against the guidelines. Promote by issuing a ‘badge’ that providers can use to confirm compatibility.

Proposed SC/WG leadership (co-leaders) and members:
Co-Leaders: Chaim Schramm, vacancy

Members: Branden Olson, Duncan Ralph, Frederick “Erick” Matsen, Inimary Toby, Jason Vander Heiden, Mikhail Shugay, Jian Ye, Enkelejda Miho, Pejvak Moghimi, Anna Obraztsova, Adrian Shepherd, William Lees

AIRR-C Committee Reporting Template for Sub-committees and Working Groups

Current:
SC/WG Name: Minimal Standards Working Group
SC/WG Co-Chairs: Christian Busse, Florian Rubelt
SC/WG Members: Ahmad Chan Bukhari, JP Brückert, Scott Christley, Brian Corrie, Lindsay Cowell, Jamie Faison, John Graybeal, Steven Kleinstein, Uri Laserson, Susanna Marquez Gargallo, Nishanth Marthandan, Enkelejda Miho, Nina Luning Prak, Corey Watson
Purpose/Mission: To develop, maintain and promote minimal reporting standards for AIRR-seq data

SC/WG Demographics:

- Country of residence: US 12, CA 2, CH 1, DE 1
- Female to male ratio: 4:12
- Members with non-academic affiliations: 3 Active members of mailing list recipients: 55%

The list includes all active members, i.e. individuals who either attended at least one WG call during the reporting period or had WG-related activity on the AIRR Standards Github repo. The mailing list contained a total of 29 unique individuals over the reporting period.

Current Goals:

- MiAIRR 1.1 (“Make it known, make it easy and demonstrate its utility”):
    - Make it known:
        - Reach out to and assist other labs to submit their data
    - Make it easy:
        - Develop toolkit/pipeline for submission and retrieval to/from NCBI
        - Evaluate submission to other INDSC repos (EBI/ENA)
        - Identify ontologies for a limited number of key data elements (6- 8 elements)
        - Bug-fix MiAIRR-NCBI implementation
        - Refine MiAIRR 1.0 data fields only if necessary, no addition of new data fields
    - Demonstrate utility:
        - Showcase with a meta-analysis using multiple data sets
- MiAIRR 2.0 (“Meeting the needs of the next decade”):
    - Develop mechanisms how AIRR-seq studies can report metadata related to clones or singles cells:
        - cell phenotypes (e.g. flow cytometry)
        - Ig/TCR reactivities and functional properties
        - structural information
    - Identify strategies to perform zero-knowledge analysis of restricted- access data

Progress report on current goals:

- Ontologies: The WG recognized that ontologies recommended by MiniStd would have relevant impact on other AIRR activities. Therefore a joint “Ontologies and Vocabularies Team” (OntoVoc) was established, which includes - among its 14 members - representative from the ComRepo, DataRep, GLDB and MiniStd WGs. The OntoVoc Team held its first “sprint” (a series of five calls on a weekly basis) in November/December 2018 to address the following topics:
    - Develop criteria for ontologies to be used in AIRR standards.

    - Identify ontologies for the five MiAIRR fields organism, disease_diagnosis, cell_subset, tissue and strain_name.
    - Provide an overview of current ontology implementations by other repositories and potential technical challenges.
    - The Team successfully completed all of these goals, with the exception of identifying an ontology for the strain_name field. The full report is available under: https://docs.google.com/document/d/115vclPVL_99xn1mUVOQNY55ax2SI7BNgpy-tTIDSQfc/edit
- MiAIRR refinements:
    - The definitions of several MiAIRR fields were clarified.
    - Introduced the nullable property and the respective state for all MiAIRR fields.
    - Following the OntoVoc report (see above), ontologies were introduced for the organism, disease_diagnosis, cell_subset and tissue fields.
    - Introduced a new set of keyword_* fields to increase findability of AIRR- seq data based on high-level properties.
- MiAIRR implementations:
    - CEDAR-AIRR, a web-based submission pipeline, implemented by groups in Stanford / Yale (https://doi.org/10.3389/fimmu.2018.01877).
    - NCBI, as published in the original report.
    - ENA has been undergoing “renovations”, therefore a mapping of the term NCBI implementation was put on hold for now.
- MiAIRR extensions:
    - An extension that will cover single-cell gene expression information (both flow cytometry and RNA-seq based) is currently drafted.
    - The WG recognized that both receptor reactivities and structural information is already covered by IEDB. Therefore an independent implementation of this information was not performed. However, to increase and/or accelerate searches (i.e. making sequences Findable) non-required fields holding aggregated information should be specified in the future.
- Metaanalysis: Given the high interconnectedness between the various AIRR WGs, MiniStd came to the conclusion that this activity should not be done by one WG alone but rather demonstrate the capability of the full AIRR Standards stack ("AIRR Ecosystem") involving ComRepo, DataRep and Software. It therefore took the first steps organize the establishment of an "Ecosystem demonstrator and meta-analysis" WG that would bundle these activities (discussed below).
- Zero-knowledge analysis of restricted data: The WG did not produce any tangible output on this topic.

Future:
Proposed goals and work plan for the coming interval:

- WG will transition to a standing committee to take over custody of the MiAIRR data standard, proposed chairs/members below.
- For the originally planned “Meta-analysis” activity the creation of a new “Ecosystems demonstrator and meta-analysis” WG will be proposed at AIRR2019. If this should not be successful, the activity will be shelfed until the identified roadblocks (e.g. missing infrastructure, missing data sets, missing analysis tools) have been resolved.
- The “Ontology” activities will be continued by the OntoVoc Team in a “Sprint” format, as the feedback from the team members on regarding this setup was very positive.

Proposed SC leadership (co-leaders) and members:
SC Co-Leaders: Christian Busse, N.N.
SC Members: Ahmad Chan Bukhari, Brian Corrie, Enkelejda Miho

AIRR-C Committee Reporting Template for Sub-committees and Working Groups

Current:
SC/WG Name: IG/TR Germline Database Working Group
SC/WG Co-Leaders: Andrew Collins and Corey Watson
SC/WG Members: Andrew Collins, Brian Fritz, Bruno Gaeta, Cathrine Scheepers, Chaim Schramm, Christian Busse, Corey Watson, Daniel Gadala-Maria, Davide Bagnara, Deanna Church, Duncan Ralph, Felix Breden, Florian Rubelt, Gur Yaari, Henk-Jan van der Ham, Jamie Faison, Jamie Scott, Jean-Philippe "JP" Bürckert, Jian Ye, Justin Kos, Katherine Jackson, Kevin Wu, Martin Corcoran, Mats Ohlin, Melissa Smith, Nishanth Marthandan, Oscar Rodriguez, Pierre Boudinot, Scott Christley, Steven Kleinstein, Susana Magadan, Victor Greiff, William Gibson, William Rounds, William Lees
Purpose/Mission: To promote the comprehensive and accurate identification, description, classification, annotation, curation, and consistent use of germline IG and TR genes/alleles across species, strains, and populations.

Current Goals:
Current Plans (GLDB):

- Finalize a formal process for depositing data associated with inferred novel alleles.
- Work with NCBI and IMGT to establish an efficient process for the submission of inferred allele data.
- Write up a "how-to" manual for the AIRR Community.
- Initiate consideration of issues relating to GLDBs of mouse and macaque sequences. In collaboration with IMGT, finalize the structure of a secondary "tiered" database of inferred alleles. (e.g., further define processes for the inner workings of the "stoplight" system and how exactly these will be represented and made available by IMGT for broader use by the AIRR Community).

Current Plans (Inferred Allele Review Committee (IARC))**:**

- Develop a system for how this group will function in the short-/long-term. (e.g., will they operate formally as part of the IUIS/IMGT nomenclature committee?).
- Decide who will serve on the committee and on what terms.
- Begin collating data for inferred human IGHV alleles from interested researchers, which will ultimately go through the formal submission process, once established.
- Begin to develop criteria for the review and acceptance of inferred alleles for AIRR-relevant loci other than human IGHV.

Progress report on current goals:

- IARC has held 36 meetings since January 2018:
    - Minutes of meetings are available at The Antibody Society website (https://www.antibodysociety.org/inferred-allele-review-committee-iarc/).
    - The IARC team was nearly always joined by William Lees, and often by Corey Watson.
- IARC and WG members published a paper in Frontiers in Immunology (Research Topic under the auspices of the IUIS Nomenclature Committee) in March 2019, detailing policies and procedures for the submission, evaluation, naming and publication of inferred IGHV genes. (https://www.ncbi.nlm.nih.gov/pubmed/30936866)
- IARC has affirmed 9 Human Level 1 IGHV sequences, and these have now been incorporated into the IMGT reference dataset (http://www.imgt.org/IMGTindex/IMGT-NC-Report_2019-11-

0418_Homsap_IGHV_180419.pdf). IARC has affirmed 12 Human Level 0 IGHV sequences.

- William Lees has developed OGRDB (https://ogrdb.airr-community.org) as a database to facilitate the management of submissions, and for documentation of the decision-making process.
- Subgroups have be established to address specific issues facing germline databases for non-human species. These include mouse, macaque, and salmonids.
    - The mouse subgroup is preparing a manuscript for submission to Frontiers in Immunology, proposing a new nomenclature to deal with the chaos of IG genes in inbred mouse strains.
    - The macaque subgroup is compiling existing AIRR-seq and genomic sequencing data/assemblies to formally assess the state of current rhesus databases. From these data, the group intends to work with the larger GLDB-WG and IMGT to establish a path forward for an improved nomenclature system and database. A manuscript will likely accompany this work, with a planned submission in 2019.
    - The salmonids subgroup (Susana Magadan [Chair] and Pierre Boudinot) was established in November 2018 and is working with the human IARC. A manuscript is also in preparation for submission to Frontiers in Immunology, describing a new Salmonid nomenclature. In addition, this subgroup has worked with IMGT to improve and update the the Salmo salar and Oncorhynchus mykiss databases.
    - A subgroup was also formed to focus on the topic of chimeric sequences that arise in the generation of AIRR-seq datasets. Some initial assessments of chimerism detection have been conducted.

Future

Proposed goals and work plan for IARC for the coming interval:

- Continue to refine the review process for novel inferences in human IGHV, while also developing additional mechanisms for the review of IGHD/J, IGKV/J, and IGLV/J genes. (Ultimately, this will be extended to efforts in the TR gene loci).
- Continue to work with IMGT toward the inclusion of novel inferred alleles reviewed and approved by IARC.
- Develop a sustainability model that can ensure longevity as well as establish a system of oversight for the creation of additional IARCs focused on data review and inferred gene/allele curation for non-human species.
- Establish a process that will define the set of alleles suitable for evolutionary/population genetics. The IARC process and OGRDB are possible routes towards the definition of such a set of alleles.

Proposed goals and work plan for OGRDB for the coming interval:

- Extend support to additional species and to additional segments and loci, as these become covered by IARCs.
- Provide support for genomically inferred genes and alleles.
- Enable more informed use of germline databases by attaching population/ethnicity-level metadata to curated genes and alleles.

Proposed goals and work plan for GLDB-WG for the coming interval:

- Propose and refine a new nomenclature for mouse IG genes, which addresses the shortcomings of the current NC and takes inferred genes as well into account as the convoluted breeding history of commonly used mouse strains.
- Propose and refine a new nomenclature for macaque IG genes, which leverages both existing genomic and AIRR-seq datasets.

- Provide a standardized and comprehensive nomenclature of IG genes from Salmo salar, O. mykiss and other salmonid species. Extend the efforts of salmonid IG gene annotation to TR loci.
- Develop and test tool/s to identify sequences with evidence of chimeric cross- over events in datasets used for germline gene inference and AIRR-seq analysis.
- Continue to work with IMGT on questions regarding versioning, programmatic access and licensing of germline reference databases.

Proposed SC/WG leadership (co-leaders) and members: Same as above.

AIRR-C Committee Reporting Template for Sub-committees and Working Groups

Current:
SC/WG Name: Common Repository Working Group
SC/WG Co-Leaders: Lindsay Cowell, Brian Corrie
SC/WG Members: Meredith Ashby, Felix Breden, Richard Bruskiewich, Tania Bubela, Syed Ahmad Chan Bukhari, Christian Busse, Scott Christley, John Harting, David Klatzmann, Uri Laserson, Nishanth Marthandan, Bjoern Peters, Adrien Six, Adrian Thorogood, Corey Watson, Yariv Wine, and George Blanck
Purpose/Mission: To promote and facilitate deposit, access, and sharing/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.

Current Goals:

- Define the API that AIRR Common Repositories will utilize to support standardized queries of AIRR-seq data.
- Have at least two repositories implement the API
- Demonstrate successful execution of queries that return repertoires and/or rearrangements identified by querying over a range of metadata

Progress report on current goals:

- API Definition:
    - Completed:
        - Defined with specification available here: https://github.com/airr-community/airr-standards/blob/master/specs/common_repository_api.yaml
- Have at least two repositories implement the API:
    - Partially Completed:
        - VDJServer draft implementation completed
        - iReceptor implementation underway (Completed summer 2019)
        - iReceptor Turnkey repository (freely available) will support the AIRR API (https://github.com/sfu-ireceptor/turnkey-service-php)
- Demonstrate successful execution of queries that return repertoires and/or rearrangements identified by querying over a range of metadata:
    - Completed:
        - Query documentation available here: http://docs.airr-community.org/en/metadata-docs/api/overview.html#search-and-retrieval

Other outcomes:

- Maintenance of the CRWG Recommendations:
    - Draft recommendations (under development) are here: https://github.com/airr-community/common-repo-wg/blob/issue-27/recommendations.md
    - Version to be approved at the General Assembly meeting

Future:
Proposed goals and work plan for the coming interval:

- Continue the development of the AIRR API.
- Transform VDJServer's draft code into an AIRR reference implementation that resides in the AIRR Community github.
- Work with the community to encourage the adoption of the AIRR API for repositories.
- Develop an AIRR Repository Registry to make AIRR compliant repositories "Findable" from a FAIR principles perspective.
- Work with Minimal Standards, Data Representation Working Groups (and others) to continue developing Ontology definitions for MiAIRR terms.
- Work with the emerging Meta-analysis Working Group, should it form, to develop a meta-analysis across the entire AIRR Community data/software Ecosystem.

Proposed SC/WG leadership (co-leaders) and members:
SC/WG Co-Leaders: Lindsay Cowell, Brian Corrie
SC/WG Members: Meredith Ashby, Felix Breden, Richard Bruskiewich, Tania Bubela, Syed Ahmad Chan Bukhari, Christian Busse, Scott Christley, John Harting, David Klatzmann, Uri Laserson, Nishanth Marthandan, Bjoern Peters, Adrien Six, Adrian Thorogood, Corey Watson, Yariv Wine, and George Blanck

AIRR-C Committee Reporting Template for Sub-committees and Working Groups

Current:
SC/WG Name: Biological Reagents WG
SC/WG Co-Leaders: Maggie Bostick and Encarnita Mariotti-Ferrandiz
SC/WG Members:
Name
Member since
Affiliation
City, Country
Research topic

Anne Eugster
07/12/2017
University of Dresden
Dresden, Germany
T cell repertoire; single cell; autoimmune disease; human

Christian Busse
06/12/2017
DKFZ
Heidelberg, Germany
B cell repertoire; single cell; bioinformatics

Cinque Soto
06/12/2017
Vanderbilt University
Nashville (TN), USA
B cell repertoire; vaccine design

Davide Bagnara
06/12/2017
University of Genoa
Genoa, Italy
B cell repertoire; oncoimmunology

Encarnita Mariotti-Ferrandiz (co-chair)
06/12/2017
Sorbonne University, Medical School
Paris, France
T cell repertoire; single cell; autoimmune disease; human

Jacob Sherkow
06/12/2017
New York Law School
New York (NY), USA
Regulatory and IP issues on biotechnologies

Johannes Trueck
06/12/2017

University of Zurich
Zurich, Switzerland
B cell repertoire; immunodeficiency; immune reconstitution; standards; vaccine design; bioinformatics

Kevin O'Connor
22/12/2017
Yale School of Medicine
New Haven (CT), USA
B cell; Autoimmune disease & tumor; Human Translational Immunology (HTI) consortium

Nina Luning
06/12/2017
University of Pennsylvania, Perelman School of Medicine
Philadelphia (PA), USA
B cell repertoire; autoimmune disease; human

Maggie Bostick (co-chair)
06/12/2017
Takara Bio
Mountain View (CA), USA
Molecular biology R&D; immune repertoire kit and tool development

Michal Or-Guil
22/12/2017
Humbold University
Berlin, Germany
Systems Immunology; B cell; Ig biomarker

Uri Hershberg
06/12/2017
Drexel University, School of Science and Engineering; University of Haifa
Philadelphia (PA), USA and Haifa, Israel
Immune repertoire data modeling

Aimee Payne
03/04/2019
University of Pennsylvania Perelman School of Medicine
Philadelphia (PA), USA
Human autoimmunity; B cell repertoire; cellular immunotherapy

Chris Tipton
20/02/2019
Emory University Department of Medicine
Atlanta (GA), USA
Immune Profiling Program

Hye-Won Song
20/02/2019

BD
San Jose (CA), USA
Single Cell Genomics Application

Lindsay Cowell
20/02/2019
UT Southwestern
Dallas (TX), USA
B cell repertoire modelling

Marie-Paule Lefranc
20/02/2019
IMGT
Montpellier, France
IMGT

Pierre Barennes
17/04/2019
Sorbonne University, Medical school
Paris, France
T cell repertoire; autoimmune disease; human

Purpose/Mission:
(Please refer to your SC's or WG's description of mission/purpose at:
https://www.antibodysociety.org/the-airr-community/ and add revisions if necessary)

The goal of our working group is to provide the AIRR Community with biological calibrators and reagents for standardization and harmonization of AIRR-seq experiments.

To reach that goal, the WG will work on identifying and recommending a set of biological standards that can be used for identification of bias and quality issues pertaining to AIRR- seq.

The Biological Resources Working Group will also coordinate the development and dissemination of reference samples that can be used as controls. The working group will reach out to established government and private organizations to help encourage ease of use and broad adoption.

Current Goals:
(Please refer to your SC's or WG's description of goals for 2018 at:
https://www.antibodysociety.org/the-airr-community/ and add revisions if necessary)

- Create a questionnaire to survey scientists regarding their needs for standards and calibrators in AIRR-seq experiments and analyse the results
- Write a manuscript reviewing biological standards and other analytical controls for immune repertoire profiling experiments by NGS using the questionnaire results
- Establish a collection of such standards for inter-lab comparisons and ultimately for communal use
- Write a second manuscript about one or more specific assay standards

Progress report on current goals:
The WG organized bi-weekly meetings to accomplish these goals.

- Survey reviewing current needs of standards for AIRR-Seq (completed):
    - The WG designed and disseminated the questionnaire through several mailing lists including members' personal contacts, as well as the AIRR-community and the Antibody society mailing lists. The survey was closed after 2 months, and we collected answers from 46 participants, out of which 43 were considered (2/46 were duplicated answers from the 2 responders and 1/46 was not completed as the responder indicated not being involved in AIRR-seq data production nor analysis).
- Manuscript about current needs and challenges on standards and survey results (in preparation):
    - The results of the survey have been compiled into the first manuscript with the aim to submit it to Frontiers in Immunology.
    - The manuscript is divided into four sections:
        - An introduction describing the applications of the variety of AIRR-seq techniques and the survey participants' demographics.
        - A presentation of the needs and requirements for control reagents for their applications.
        - A summary of currently available controls (including standards introduced during a dedicated session (see below)).
        - A proposal for areas of prioritization for current use and future development of controls.
    - A pre-submission letter has been sent to Deborah Dunn-Walter on April 30th 2019, and she replied with a positive response, encouraging us to submit on May 2, 2019.
- Scientific session on standards current use and development by three researchers (completed):
    - The WG invited three researchers to present their current development, each during one session:
        - Sai Reddy
        - Johannes Trueck (member of the WG)
        - Chris Tipton (member of the WG)
- Expansion of the WG members (completed):
    - The survey attracted new members from January 2019, expanding the community and reinforcing the interest for controls in the field.

Future:

Proposed goals and work plan for the coming interval:

- Short term:
    - Wrap-up and submit the first manuscript.
- Coming interval:
    - Subdivide the WG into subgroups, with the aim to address standards & calibrators issues for:
        - B cells
        - T cells
    - Initiate a comparative study for B-cell AIRR-Seq methods to work-out existing differences between approaches and define what calibrators need to be developed.
- Long term:

- Identify and collect suitable DNA/RNA calibrators/controls and cell mixes for AIRR- seq
- Perform outreach to academic, government, and commercial partners to disseminate the identified and validated controls
- Identify members of the community able to define statistical measurement for AIRR- Seq experiment validation (in collaboration with other WGs)
- Write a second manuscript about one or more specific assay standards

Proposed SC/WG leadership (co-leaders) and members:
As for the co-leadership, current co-leaders are willing to continue. This needs to be agreed on a vote during the AIRR-C 2019 meeting with members (present or remotely).
Member list needs to be updated according to AIRR-C rules. To be discussed with the Executive committee and other WG during the AIRR-C 2019 meeting.