2017 Adaptive Immune Receptor Repertoire (AIRR) Community Meeting

Prepared by the 2017 AIRR Community Meeting Organizing Committee:
Jamie Scott, Florian Rubelt, Chaim Schramm, Danny Douek, and Felix Breden

Welcome to the 3rd annual meeting of the Adaptive Immune Receptor Repertoire (AIRR) Community! AIRRs are characterized through the use of large data sets generated by high-throughput sequencing. The mission of the AIRR Community (**AIRR-C**) is to address issues involving the production, analysis, sharing, and use of this type of data (**AIRR-seq** data) to profile the immune system in health and disease. Thus, the AIRR-C comprises a diverse range of immunogeneticists who sequence AIRRs, and bioinformaticians and statisticians who analyze AIRR-seq data, along with computer-security experts, industrial partners, and scholars in the aspects of ethics, law, and policy whose expertise is relevant to the use of AIRR-seq data. The main outcome of our first AIRR-C Meeting (May 2015) was the formation of three Working Groups (**WGs**) whose goals were to develop recommendations for: (i) a Common Repository for AIRR-seq data, (ii) minimal standards for publishing or depositing AIRR-seq data, (iii) resources and guidelines for the evaluation of cellular, molecular and statistical methods for obtaining and analyzing AIRR-seq data. In addition, an Executive Committee was formed to plan annual meetings, and suggest options on whether and how the AIRR-C could become a society or other more formalized organization. The second AIRR-C Meeting (June 2016) comprised a mix of scientific presentations, reports from the AIRR-C WGs and Executive on their progress and plans. Attendees voted on the WGs' plans for the year; the Executive was tasked with finding an existing society to join, and developing a governance plan for the AIRR-C.

This 2017 AIRR-C Meeting will once again comprise scientific presentations in combination with progress reports from the WGs and Executive, followed by approval by attendee votes. Over the past year, the WGs and Executive have made significant progress on several fronts, including publication of a commentary describing the AIRR-C's mission and goals (Breden *et al*., 2017, *Front Immunol*, 8:1418), and recommendations for minimal standards for depositing AIRR-seq data to repositories at NIH's National Center for Biotechnology Information (**NCBI**) (Rubelt *et al*., 2017 *Nat Immunol*, 18(12):1274-1278). Progress in developing processes for (i) inference of germline V genes from AIRR-seq data, (ii) a distributed network of Common Repositories for AIRR-seq data, (iii) joining The Antibody Society and maintaining the AIRR-C WGs and governance structure, are highlights of the progress that will be reported in the 2017 meeting. Targeted plans, such as how germline V genes can best be inferred, and issues, such as if and how AIRR-seq data should be protected, will be discussed in depth by a diverse audience. An informal evening session will comprise demonstrations of software, platforms, and methods, which have been a hit at our previous meetings. This meeting will also provide mentoring for researchers who are interested in this technology, but do not have ready access to or training in it: in particular, new investigators and those from middle- and low-income countries.

The 3rd annual AIRR-C meeting will once again be held at NIAID's Fishers Lane Facility in Rockville, MD; we are grateful to ***Jon Warren*** and ***Joe Breen*** for their leadership in procuring both the venue, and funding to support our speakers. We are also grateful to our ***industry sponsors*** (listed below) for their financial and intellectual contributions to this meeting. We look forward to the participation by interested NIH staff members, and members of the biomedical research and therapeutics industries. ***The participation of all stakeholders in the AIRR-C*** has been central to our progress (*e.g.*, creation of guidelines and infrastructure at NCBI for annotation and storage of AIRR-seq data in an accessible and sharable form; see Rubelt *et al*. above.

Overview of the 2017 AIRR Community Meeting: The meeting agenda that follows is based on recommendations from the first and second annual AIRR-C Meetings. (See the final report for the May 2015 and June 2016 AIRR-C meetings at airr-community.org).

The goals of the 2017 meeting are: ***(i) Outreach and inclusiveness:*** To maximize the involvement of investigators performing or interested in AIRR research in forming a community and developing a framework for the production and use of AIRR data. As with the May 2015 meeting, this is accomplished through both face-to-face and virtual attendance by participants; our email list currently comprises >900 people. ***(ii) Education:*** To provide an educational environment for all those working in the highly diverse areas involving AIRR-seq data acquisition, analysis and use. Attendees will learn about the broad range of research in this area, meet new colleagues and form new collaborations, all of which are key to the development of truly interdisciplinary research. ***(iii) Finalization and approval of recommendations developed over the past year by the AIRR-C WGs and Executive Committee:*** These will be discussed, revised, and endorsed with input from the face-to-face and virtual attendees.

Proposed recommendations will be presented on Day 1, providing the opportunity for attendees to familiarize themselves with the issues, and initiate discussions that will last throughout the meeting. Days 2 and 3 will comprise four sessions – one for each WG and one for the Executive Committee – with each session comprising presentations by leaders in the field, followed by a continuation of the discussion on that WG's recommendations (or those of the Executive Committee). On the last day of the meeting, the recommendations will be discussed by the participants as a whole, amended as required, and voted on by all in attendance (including virtual attendees).

This inclusiveness has worked well in meeting the needs of the AIRR-C as a whole, and has aided the publication of standards, as well as their incorporation at NCBI. We are now looking to work with other international communities (*e.g.,* IUIS and FOCIS) in developing internationally accepted standards for reporting and sharing AIRR-seq data. We envision this process as a "best practice" in the development of burgeoning research areas, such as those involving Big Data and systems biomedicine.

Overall schedule: The 2017 meeting takes place Sunday - Wednesday**, 3-6 December 2017**. On the morning of **Day 1**, the Working Groups will meet and finalize their draft recommendations, which will be introduced in the afternoon to the meeting of the whole. **Day 2** will begin with sessions organized around two of the three AIRR WGs. Each WG has a representative on the Executive Committee and has primary responsibility for setting the agenda in their respective session, including inviting speakers. ***Evening sessions*** will be held on Days 1 and 2; attendees will get together at the Washington DC/ Rockville Hilton Hotel for a reception and discussion of inferring germline AIRR genes, and demonstrations of software, pipelines and platforms for the analysis of AIRR-seq data. **Day 3** will comprise a third WG session, followed by a session from the Executive Committee on the integration of AIRR-seq data with other systems-level data, and a final discussion on whether the AIRR-C should join The Antibody Society, the AIRR-C governance structure within that society, and initial plans for the 4th annual meeting. ***Poster sessions*** will be held during lunchtime on Days 2 and 3 at the NIAID Fishers Lane Facility. On the morning of **Day 4**, the Working Groups and Organizing Committee will present their final recommendations, and suggest work for the coming year, all of which will be finalized and ratified by participants.

Regarding food: In order to keep registration as low as possible, attendees will be responsible for paying for their meals. Boxed lunches will be available in the Fishers Lane Facility. (*You must sign up for them in advance!*). During the evening of ***Tuesday, December 5, 2017*** participants are invited to attend dinner at a local restaurant; the cost will be ~$60 per person.

Transportation: Most or all attendees will walk from their hotel to the NIAID Fishers Lane Facility. To enter the facility each day, all attendees will need to show some form of identification (US citizens: a driver's license or passport; non-US citizens: a passport). Attendees who require transportation to and from the Fishers Lane Facility should contact the hotel to see if the hotel shuttle can be scheduled; otherwise they should hire a cab.

Support: Besides providing the facility for the meeting, and support for virtual attendance, NIH has been very generous in providing funds to help cover expenses for speakers.

Our Sponsors: Adaptive Biotechnologies, 10X Genomics, Genentech, Grifols, and The Antibody Society have very kindly provided funds for the meeting rooms and evening receptions at the Washington DC/Rockville Hilton Hotel.

Day 1: Sunday, December 3, 2017 ~ Washington DC/Rockville Hilton Hotel

Mentorship Activity & Opening of AIRR Community Meeting
& Presentation of Working Group Progress

Time	Activities
9:00 – 12:00	**Working Groups meet concurrently** Minimal Standards Working Group convenes ~Twinbrook Room Common Repository Working Group convenes ~ Montrose Room Tools and Resources Working Group convenes ~ Democracy Room Germline Gene Working Group convenes ~ Lincoln Room **Mentorship Activity for New Investigators:** Victor Greiff, University of Oslo **Steps in AIRR Data Capture and Analysis ~** Jackson Room
12:00 – 12:45	*Lunch (on own)*
12:45 – 1:00	MEETING BEGINS: **Introduction of Conference Organizers, Working Group Leaders, and Review of the Agenda ~** Washington Theater
1:00 – 3:00 1:00 – 2:00 2:00 – 3:00	**All Working Groups convene with conference attendees to discuss recommendations including:** **Implementation Details from Minimal Standards Working Group** **Implementation Details from Common Repository Working Group**
3:00 – 3:30	*Break*
3:30 – 3:50 3:50 – 4:00 4:00 – 4:15 4:15 – 4:30	**Tools and Resources – Introduction and Software WG update** **Tools and Resources – File Formats WG update** **Tools and Resources – Biological Standards WG update** **Toward standards for TCR repertoire studies –** Encarnita Marrioti-Ferrandiz
4:30 – 5:30	**Presentation of Proposed Governance Structure:** **Discussion & Vote on Joining TAbS (The Antibody Society)**
5:30 – 6:00	***Reception ~ snacks provided ~ cash bar ~*** *Eisenhower Room* *No government funds will be used for food and drinks*
6:00 – 7:30	**Recommendations from Germline Gene Working Group:** **AIRR Community's Role in IG Gene Naming and Its Relationship to the IUIS IG/TR/MH Nomenclature Committee**
	Attendees are on their own for dinner

Day 2: Monday, December 4, 2017 ~ NIAID's Fishers Lane Facility

Scientific Presentations: Sessions 1 – 3

Webinar Links for Virtual Attendance

Monday, December 4 8:30 AM - 1:00 PM
https://attendee.gotowebinar.com/register/7041796218829310979

Monday, December 4 1:45 PM - 5:00 PM
https://attendee.gotowebinar.com/register/4636391332208890371

Time	Activities
7:45 AM	*Attendees gather in Washington DC / Rockville Hilton Hotel lobby to walk to NIAID Fishers Lane Facility ~ arrive at facility by 8:00 to go through security*
8:30 – 8:40	**Introduction to the Day's Agenda ~** Room 1D13 **Session 1: Germline Gene Inference Tools** **Moderator's Introduction:** Andrew Collins, University of New South Wales
8:45 – 9:05	**Presentation 1 – TIgGER for IGHV Gene Inference** Steve Kleinstein, Yale University
9:05 – 9:25	**Presentation 2 - IgDiscover for IGHV Gene Inference** Martin Corcoran, Karolinska Institute
9:25 – 9:35	**Shared Q&A**
9:35 – 10:00	**Presentation 3 – Evaluation and Validation of Germline Gene Inference** Mats Ohlin, Lund University
10:00 – 10:10	**Q&A**
10:10 – 10:30	**Panel Discussion: Validation & Comparison of Inference Tools**
10:30 – 10:45	*Break*
10:45 – 10:55	**Session 2: IGH Genomic Haplotypes** **Moderator's Introduction:** Marie-Paule Lefranc, IMGT, University of Montpellier
10:55 – 11:20	**Presentation 1 – Characterizing Germline Immunoglobulin Diversity in Human and Mouse** Corey Watson, University of Louisville Medical School
11:20 – 11:30	**Q&A**
11:30– 11:55	**Presentation 2 – Role of Novel Regulatory Elements in the Formation of the Ig-κ Repertoire** Ann Feeney, The Scripps Research Institute
11:55 – 12:05	**Q&A**
12:05 – 12:30	**Presentation 3 – Immunogenomics of the Rhesus macaque, an animal model for HIV vaccine development** Akshaya Ramesh, University of California, San Francisco **Q&A**

12:30 – 12:40 12:40 – 1:00	**AIRR Community Discussion** **Discussion Leaders:** Marie-Paule Lefranc & Corey Watson **IGH Haplotype and Germline Gene Nomenclatures**
1:00 – 1:45	***Boxed Lunches & Poster Session I ~*** *LD10 Lower Level*
1:45 – 2:15	**Minimal Standards Implementation Demos** Steven Kleinstein, Nina Prak, Florian Rubelt, Christian Busse, & Syed Ahmad Chan Bukhari
2:15 – 2:25	**Session 3: IP Issues in Data Sharing across Shared Repositories** **Moderator's Introduction:** Tania Bubela, Simon Fraser University
2:25 – 2:45	**Presentation 1 - Learnings from Publicly Available Immune Receptor Datasets** Harlan Robins, CEO, Adaptive Biotechnologies
2:45 – 3:05	**Presentation 2 - On the Human Immunome Project** Jim Crowe, Vanderbilt University
3:05 – 3:25	*Break*
3:25 – 3:45	**Presentation 3 - Intellectual Property Law** Jacob Sherkow, New York Law School
3:45 – 4:30	**Panel – IP Issues in Data Sharing** Harlan Robins, Jim Crowe & Jacob Sherkow
4:30 – 5:00	**Common Repository WG: Resolving IP and Protected Sequences Discussion and Community Vote** Lindsay Cowell & Corey Watson
5:00	*End of Day at NIAID's Fishers Lane Facility ~* *Attendees return to Washington DC/Rockville Hilton Hotel*
5:30 – 7:45 5:30 – 6:30 6:30 – 7:00 7:00 – 7:45	***Reception and Tool Demos at the Washington DC/Rockville Hilton Hotel*** ***Snacks provided ~ cash bar ~*** *Eisenhower Room* *No government funds will be used for food and drinks* **5 minute Tool Demonstrations** **Panel on validation and improvement of AIRR analysis tools** **Face-to-face tool demonstrations by teams**
	Attendees are on their own for dinner

Day 3: Tuesday, December 5, 2017 ~ NIAID's Fishers Lane Facility
Scientific Presentations: Sessions 4 and 5

Webinar Links for Virtual Attendance
Tuesday, December 5 8:30 AM - 12:15 PM
https://attendee.gotowebinar.com/register/7134530129027597827

Tuesday, December 5 2:00 PM - 5:00 PM
https://attendee.gotowebinar.com/register/2846600806957000963

Time	Activities
7:45 AM	*Attendees gather in Washington DC/Rockville Hilton Hotel lobby to walk to NIAID's Fishers Lane Facility ~ arrive at facility by 8:00 to go through security*
8:30 – 8:40	**Introduction to Day 3 Agenda ~** Room 1D13 **Session 4: AIRR Data in the Study of Disease** **Moderator's Introduction** Steve Kleinstein, Yale University
8:40 – 9:05 9:05 – 9:15	**Presentation 1 - Cancer Immunotherapy** John O'Malley, Brigham and Women's Hospital, Boston **Q & A**
9:15 – 9:40 9:40 – 9:50	**Presentation 2 - Neoantigen-specific T-cell Receptor Repertoires from Healthy Donors** Johanna Olweus, Oslo University Hospital **Q & A**
9:50 – 10:05	*Break*
10:05 – 11:30 11:30 – 11:40	**Presentation 3 - BCR repertoire analysis in human SLE** Iñaki Sanz, Emory University School of Medicine **Q & A**
11:40 – 12:05 12:05 – 12:15	**Presentation 4 - BCR Repertoire Sequencing in Patients with Primary Immunodeficiency Analysis** Johannes Trück, University Children's Hospital, Zurich, Switzerland **Q & A**
12:15 – 2:00	***Boxed Lunches and Poster Session II ~*** *LD10 Lower Level* concurrent with ***New Investigators' Lunch with Mentors***
2:00 – 2:05	**Session 5: Integration of AIRR Data with Complex Biological and Population Data** **Moderator's Introduction:** George Georgiou, University of Texas – Austin
2:05 – 2:30 2:30 – 2:40	**Presentation 1 – Molecular Analysis of the Serum Antibody Repertoire** George Georgiou, University of Texas – Austin **Q & A**

2:40 – 3:05 3:05 – 3:15	**Presentation 2 - Genetic Diversity of Non-Human Primate & Human Germline Ig Genes** Gunilla Karlsson Hedestam, Karolinska Institute, Stockholm **Q & A**
3:15 – 3:25	*Break*
3:25 – 3:50 3:50 – 4:00	**Presentation 3 - The Antibody Response of Celiac Disease** Ludvig Sollid, University of Oslo **Q & A**
4:00 – 5:00	**AIRR Community Discussion** **Discussion Leaders:** Felix Breden & Jamie Scott **Proposed AIRR Community Governance Structure and Working Group Organization**
5:00	*End of Day at NIAID's Fishers Lane Facility ~* *Attendees return to Washington DC/Rockville Hilton Hotel*
6:30 – 10:00	*Attendees convene for Group Dinner at Seasons52 in North Bethesda*

Day 4: Wednesday, December 6, 2017 ~ NIAID's Fishers Lane Facility
Community Initiative Working Day

Webinar Links for Virtual Attendance
Wednesday, December 6 8:40 AM - 12:00 PM
https://attendee.gotowebinar.com/register/4385066063843514883

Time	Activities
7:45	*Attendees gather in Washington DC/Rockville Hilton Hotel lobby to walk to NIAID Fishers Lane Facility ~ arrive at facility by 8:00 to go through security*
8:40 – 9:30	**Conclude discussion and vote on:** **AIRR Community Governance Structure ~** Room 1D13
9:35 – 10:30	**Conclude discussion and vote on:** **Working Group Organization**
10:35 – 11:00	**Conclude discussion and vote on:** **Germline Gene Working Group Recommendations**
11:00 – 12:00	**Elect AIRR Community Leadership**
12:00	**Full Meeting Adjourns**
12:00 – 1:00	*Lunch (on own)*
1:00 – 4:00	*Working Groups will convene at the Washington DC/Rockville Hilton Hotel to finalize membership & develop plans for the coming year ~ Montrose and Twinbrook rooms*

Washington DC/Rockville Hilton Hotel
1750 Rockville Pike, Rockville MD, USA
Phone: 1-301-468-1100

Directions to Washington DC/Rockville Hilton Hotel:

From Washington National Airport
Head Northeast to a slight left, following signs for B/C Arrivals-Baggage Claim, take ramp to George Washington Memorial Parkway North. Merge onto Capital Beltway I-495 North via ramp to Maryland. Follow Exit at I-270 Spur for I-270 North

Drive Time: 35 min.

Type	Typical Minimum Charge
Super Shuttle	29.00 USD
Rental Car	45.00 USD
Subway/Rail	8.00 USD
Taxi	55.00 USD

From Washington Dulles International Airport
Head West on Saarinen Circle towards Copilot Way to slight left on Saarinen Way (follow signs for arrivals). Continue on Dulles Access Road following signs for Washington. Take Exit toward Richmond/Exit 18-19/I-495/Baltimore/Va-123, merge
Drive Time: 35 min.

Type	Typical Minimum Charge
Super Shuttle	30.00 USD
Rental Car	45.00 USD
Taxi	60.00 USD

From Baltimore/Washington International
Take Route 195 West (signs for hourly parking). Turn left I-195 West to Exit 4B/I-95 South for about 19.6 miles. Then head take Exit 27 West on I-495 towards Silver Spring/Bethesda. Take Exit 34/Route 355 North (Wisconsin Avenue/Rockville).
Drive Time: 40 min.

Type	Typical Minimum Charge
Super Shuttle	42.00 USD
Rental Car	45.00 USD
Taxi	75.00 USD

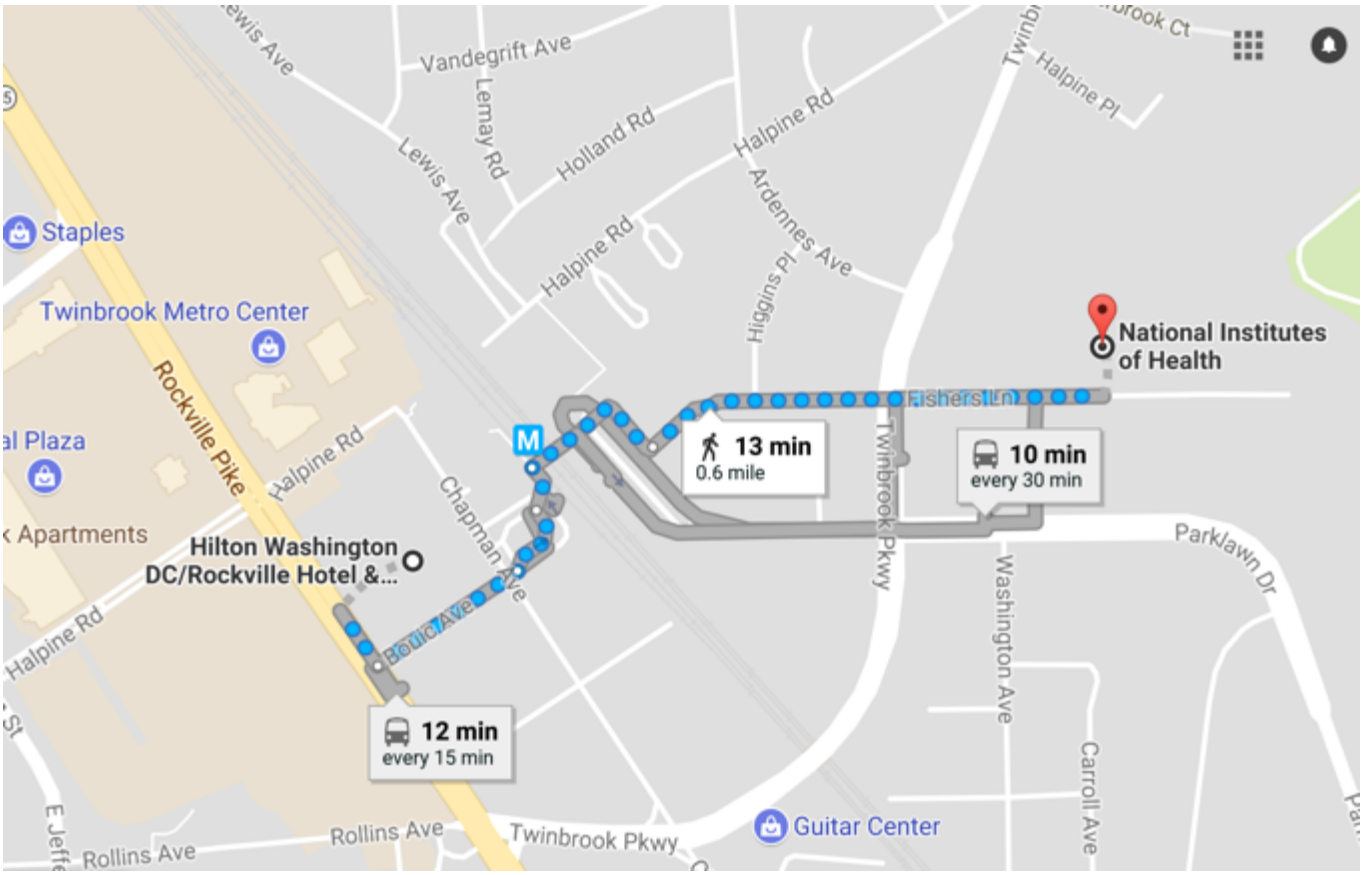
From Ronald Reagan National Airport – Metro Rail System
Take the Yellow Line towards Georgia Avenue/Petworth. Get off at Gallery Place/Chinatown, transfer station to the Red and Green Lines. Take the Red Line towards Shady Grove. Exit train at the Twinbrook Stop. The Hilton is located across the street from the station.

Airport Super Shuttle ride bookings are available at https://www.supershuttle.com/

NIAID Fishers Lane Facility
5601 Fishers Lane, Rockville, MD 20852
Rooms: 1D13 and LD10

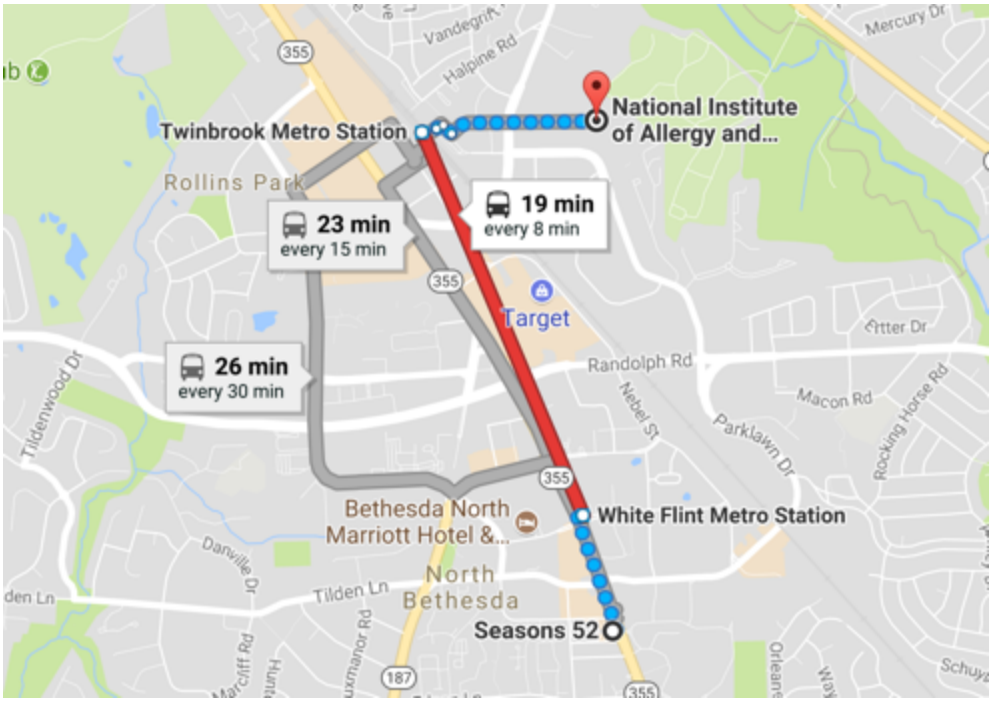
Walking Direction from the Washington DC/Rockville Hilton Hotel to the NIAID Fishers Lane Facility.

*Due to security times at the NIAID Fishers Lane Facility, please arrive before 8:00 am at the security desk. US citizens can present a US driver's license; all others please bring a valid passport. All non-US citizens should bring their passports.



Local Taxi (Rockville, MD): Action Taxi, +1 301-840-1000

Nearby Public transportation:
Metro Rail Red Line
Twinbrook Metro Station
1600 Chapman Avenue
Rockville, MD 20852

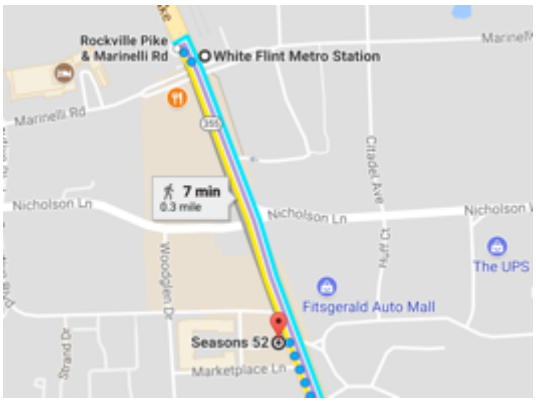


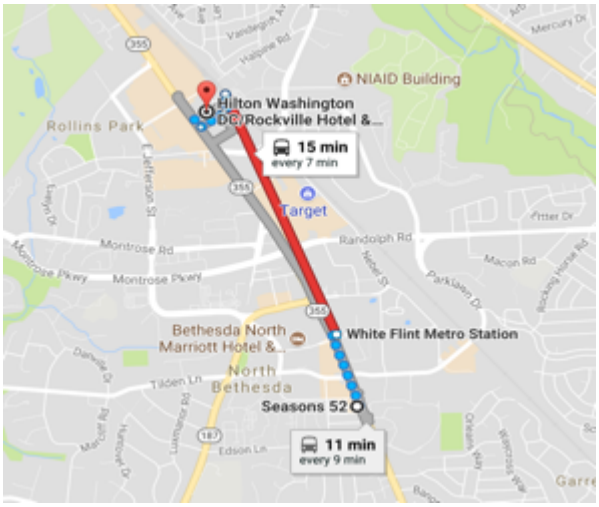
Seasons52 Restaurant
Group Dinner Location, Tuesday December 5, 7:30 pm
11414 Rockville Pike
North Bethesda, MD 20852, USA

From the Washington DC/ Rockville Hilton Hotel:

Walking Directions: Head southeast on Rockville Pike

Public Transit: Take the Metro Red Line from Twinbrook Station to White Flint Station, head south along Rockville Pike Road to Seasons52





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

Authors: Victor Greiff[1,2]

[1]Department of Immunology, University of Oslo, Oslo, Norway
[2]Department of Biosystems Science and Engineering, ETH Zürich, Basel, Switzerland

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 how to design, analyze and perform AIRR studies for answering fundamental immunological questions as well as discovering novel (therapeutic) antigen receptors. Specifically, I will address approaches to AIRR data capture 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 antigen receptor repertoires. For example, (i) diversity-, (ii) phylogenetic-, (iii) machine learning- and (iv) network-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 and limitations and highlight promising avenues of future research in basic and applied systems immunology. The workshop will be designed to be a helpful resource for all levels of expertise.

Funding: Centre for Immune Regulation (University of Oslo World-Leading Research Community)

Title: TIgGER for IGHV Gene Inference

Author: Steve Kleinstein, Yale University

Abstract: NA

Funding:

Title: The IgDiscover approach: Individualized immunoglobulin germline database production in multiple species

Authors: Martin M. Corcoran[a], Ganesh E. Phad[a], Néstor Vázquez Bernat[a], Marcel Martin[b], and Gunilla B. Karlsson Hedestam[a]
[a]Department of Microbiology, Tumor and Cell Biology, Karolinska Institutet, Sweden,
[b]Department of Biochemistry and Biophysics, SciLifeLab, Stockholm University, Sweden

Abstract: The ability of the naïve B cell repertoire to recognize structurally diverse epitopes requires an enormous range of possible antibody binding specificities. These are constrained, in part, by the availability of a limited set of germline immunoglobulin gene segments that can be combined to form functioning heavy and light chain sequences. In this study, we show that multiplexed antibody library preparation, Next Generation Sequencing (NGS) and analysis with the IgDiscover software tool enables the production of individualized immunoglobulin heavy and light chain germline databases. IgDiscover analysis identifies both known and novel allelic sequences of variable (V), diversity (D), and junction (J) segments. In addition, alleles used at either low or high frequency can be identified. The IgDiscover approach utilizes a database-replacement model and is therefore not limited by lack of a current germline database in the species being studied. In humans we have shown that IgDiscover both validates previously described V and J alleles, in addition to identifying novel undescribed alleles. In other species the use of 5'RACE library production followed by multiplex primer design has resulted in the production of species specific databases from a variety of primate and rodent species currently utilized in immunology research, including rhesus, and cynomolgus monkeys, and guinea pig and enabled the production of primer sets for monoclonal antibody isolation. The ability to rapidly define individualized germline immunoglobulin gene databases using IgDiscover enables improved assignment of cloned antibodies to their correct germline genes, facilitating accurate estimation of somatic hypermutation of antibody sequences. The production of individualized germline databases from large population groups will both enable epidemiological analysis of germline immunogobulin variation associated with disease susceptibility and will facilitate germline targeting approaches in vaccination studies.

Funding:

Title: Evaluation and Validation of Gene Inference Tools

Authors: Linnea Törnqvist[1], Ufuk Kirik[1], Mats Ohlin[1]
[1]Dept. of Immunotechnology, Lund University, Lund, Sweden

Abstract: Large NGS-derived immunoglobulin variable domain-encoding data sets offer new opportunities for the study of antibodies and immune responses. Analysis of such data is greatly facilitated by a detailed knowledge of the germline genes from which they are generated. They inherently offer opportunities to computationally infer the germline genes from which they were derived through use of *e.g.* TIgGER (1), Partis (2), or IgDiscover (3).

We have considered a diversity of approaches to support the validity of inferred genes/alleles to promote sensitivity and specificity of the inference processes (4-5; unpublished observations). We studied datasets generated from IgM-encoding transcriptomes. Inferred genes were assessed for their distribution between the subjects' haplotypes as defined by their different association to alleles of IGHJ or IGHD genes, or to IGHD genes uniquely present in only one of a subject's haplotypes. We demonstrated that many alleles, previously known as well as those currently not recognized by IMGT (6), segregated as expected from similar genes/alleles through use of haplotype analysis. However, we also demonstrated that some inferred alleles did not associate with a unique haplotype separate from that of previously known alleles, suggesting that they had been inferred in error or were derived from duplicated genes. Importantly, the allele-differentiating base of reads establishing the basis for erroneously inferred alleles were often uniquely of low quality. We also specifically defined the possibilities to, with confidence, infer the bases following codon 104 of germline genes. Altogether a thorough analysis of inference output provides confidence in properly inferred immunoglobulin genes, and identifies those inferred genes/alleles that require additional validation if they are to be accepted as real entities.

1. Gadala-Maria D, Yaari G, Uduman M, Kleinstein SH (2015) Automated analysis of high-throughput B-cell sequencing data reveals a high frequency of novel immunoglobulin V gene segment alleles. **Proc Natl Acad Sci USA** 112, 862-870.
2. Ralph DK, Matsen IV FA (2016) Consistency of VDJ rearrangement and substitution parameters enables accurate B cell receptor sequence annotation. **PLoS Comput Biol** 12, e1004409.
3. Corcoran MM, Phad GE, Vázquez Bernat, Stahl-Hennig C, Sumida N, Persson MAA, Martin M, Karlsson Hedestam GB (2016) Production of individualized V gene databases reveals high levels of immunoglobulin genetic diversity. **Nat Commun** 20, 13642
4. Kirik U, Greiff L, Levander F, Ohlin M (2017) Parallel antibody germline gene and haplotype analyses support the validity of immunoglobulin germline gene inference and discovery. **Mol Immunol** 87, 12-22.
5. Kirik U, Greiff L, Levander F, Ohlin M (2017) Data on haplotype-supported immunoglobulin germline gene inference. **Data Brief** 13, 620-640.
6. Lefranc MP (2011) IMGT, the International ImMunoGeneTics Information System. **Cold Spring Harb Protoc** 2011, 595-603.

Funding: The study was supported by a grant from the Swedish Research Council (grant number 2016-01720).

Title: Characterizing Germline Immunoglobulin Diversity in Human and Mouse

Author: Corey Watson, University of Louisville Medical School

Abstract: Genes within the immunoglobulin (IG) loci encode expressed antibodies (Abs), critical components of adaptive immunity. Largely due to technical challenges caused by locus complexity, which have hindered the characterization of variation using standard high-throughput technologies, our knowledge of germline haplotype diversity in IG regions in many species remains limited. This knowledge gap has many downstream impacts, hindering our understanding of potential impacts of IG polymorphism on Ab repertoire variation and function in health and disease. We are currently leveraging long-read sequencing approaches to explore and catalog the extent of IG germline diversity in diverse human populations and inbred mouse strains.

Funding:

Title: Role of novel regulatory elements in the formation of the Igk repertoire

Authors: E. Mauricio Barajas-Mora[1], Eden Kleiman[1], Jeffrey Xu[1], Nancy M. Choi[1], Zhaoqing Ba[2], Frederick W. Alt[2] and Ann J. Feeney[1]
[1]The Scripps Research Institute, Department of Immunology and Microbiology, La Jolla, CA, and [2]Howard Hughes Medical Institute, Boston Children's Hospital and Harvard Medical School, Boston, MA

Abstract: Using genomic DNA-based deep sequencing of the Igk repertoire, we show that the functional Vk genes rearrange at very different frequencies, but nonetheless, the frequently and moderately rearranging Vk genes are spread across the ~3 Mb locus. We are interested in understanding how distal Vk genes get close enough to the Jk genes to rearrange. It is known that long-range chromatin looping and contraction of the Igk locus are essential to this process. Published data infer the existence of hubs of long-range interactions within the Igκ locus. We analyzed those hubs, and found that they bound many transcription factors and were enriched with enhancer-associated epigenetic marks. We proposed that these novel enhancer-like elements regulate the 3D conformation of the Igκ locus, and that alterations in the architecture of the locus would alter individual Vκ gene utilization. We used CRISPR/Cas9 to delete the most prominent of these elements (E88) in an inducible pro-B-cell line and subsequently in mice. Consistent with our hypothesis, we observed changes in rearrangement frequency throughout the Vκ locus in the E88-deleted B cell precursors, with profound reduction in rearrangement of the Vκ genes near E88, more modest reduction for a bigger swath of the locus, and increase in rearrangement in the far distal and proximal regions. Our 4C data shows that these domains may correlate with the 3D conformation of the locus. Thus, our data demonstrate that this novel regulatory element contributes to the proper conformation of the 3D chromatin structure of the locus and consequently to the generation of the normal Vκ diversity in the antibody repertoire in B-cells.

Funding: R56 AI 119092, R03 AI 1154861

Title: Immunogenomics of the Rhesus macaque, an animal model for HIV vaccine development

Authors: Akshaya Ramesh[1], Sam Darko[2], Axin Hua[4], Glenn Overman[3], Amy Ransier[2], Ashley Trama[3], Georgia D. Tomaras[3], Barton F. Haynes[3], Daniel C. Douek[2,*], Thomas B. Kepler[4,5,*]
[1]Graduate Program in Genetics and Genomics, Boston University School of Medicine, Boston, MA, USA
[2]NIH Vaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA
[3]Duke Human Vaccine Institute, Duke University Medical Center, Durham, NC, USA.
[4]Department of Microbiology, Boston University School of Medicine, Boston, MA, USA
[5]Department of Mathematics and Statistics, Boston University, Boston MA, USA
*Joint senior authors

Abstract: The rhesus macaque is a critically important animal model in biomedical research, most recently playing a key role in the development of vaccines against HIV-1. Nevertheless, the immunoglobulin (Ig) loci of macaques are as yet incompletely determined and our understanding of differences between human and macaque humoral immunity remains deficient. We completed a high-coverage, high-quality whole genome sequencing and assembly project with a single rhesus macaque of Indian origin, and partial genome assemblies using genomic molecular targeting of the Ig loci in nine other rhesus macaques of Indian origin. These data indicate that the macaque Ig loci are substantially more diverse than those in humans, including greater sequence diversity and copy-number variation between individuals. It appears likely that such copy-number variation even occurs between allelic loci within individuals. Different Ig gene families in the macaque show distinct relationships to the corresponding human gene families and appear to evolve under different mechanisms. These results raise intriguing questions about the evolution of antigen receptors in primates but also have important practical implications for the design and interpretation of biomedical studies.

Funding: This research was supported by the NIAID Division of AIDS at the NIH under grants UM-1 AI100645: Center for HIV/AIDS Vaccine Immunology-Immunogen Discovery (CHAVI-ID) and HHSN272201100023C (Ronald Brown). This work was also funded in part through the intramural program of NIAID/NIH.

Title: Learnings from publicly available immune receptor datasets

Author: Harlan Robins

Abstract: NA

Funding:

Title: On the Human Immunome Project

Author: Jim Crowe

Abstract: NA

Funding:

Title: Priority: Immunotherapy and the CAR-T Patent Dispute

Authors: Jacob S. Sherkow[1]
[1]Associate Professor, Innovation Center for Law and Technology New York Law School; Visiting Assistant Professor of Health Policy and Management, Columbia University Mailman School of Public Health

Abstract: One of the most critical issues in patent disputes is priority—determining when the inventor first invented, and filed a patent application for, his or her invention. Beyond deciding whether one inventor is entitled to a patent over another, priority also cabins the scope of the "prior art" used to assess a patented invention's novelty and significance, and freezes in time assessments over the patent application's sufficiency and scope. Against this backdrop, Kite Pharma, Inc. v. Sloan-Kettering Institute for Cancer Research—currently being appealed to the U.S. Court of Appeals for the Federal Circuit from the Patent Trial and Appeal Board--presents significant and interesting challenges to importance of priority for cancer immunotherapies. One of Sloan-Kettering's patents covering chimeric antigen receptor T-cell therapy (CAR-T), U.S. Patent No. 7,446,190, was unsuccessfully challenged in inter partes review before the Patent Trial and Appeal Board by Kite Pharma, even though Kite presented evidence from almost a decade earlier of similar technology. The presentation will discuss the case, the Patent Office's view of the technology involved, and its implications for priority assessments for future iterations of the technology.

Funding: Travel paid for by presenter and CAIRR; lodging provided by presenter's in-laws.

Title: High-throughput T-cell receptor sequencing improves the diagnosis and predicts outcome in patients with Cutaneous T-cell Lymphoma

Authors: John T. O'Malley[1], Adele De Masson[1,2], Ilan Kirsch[3], Christopher P. Elco[1], Elizabeth Lowry[1], Jessica E. Teague[1], Ahmed Gehad[1], Nicole R. LeBoeuf[1], David C. Fisher MD[4], Philip M. Devlin MD[5], Harlan Robins[3], TS Kupper[1] and RA Clark[1]

[1]Department of Dermatology, Brigham and Women's Hospital and DF/BWH Cancer Center, Harvard Medical School, Boston, MA, USA;

[2]Affiliation

[3]Adaptive Biotechnologies, Seattle, WA, USA.

[4] Department of Hematologic Oncology, Dana Farber Cancer Institute, Harvard Medical School, Boston, USA

[5]Department of Radiation Oncology, Dana Farber / Brigham and Women's Cancer Center, Harvard Medical School, Boston, USA

Abstract: Mycosis fungoides (MF), the most common variant of cutaneous T-cell lymphoma (CTCL), is a malignancy of skin-tropic T cells. Two major clinical challenges in caring for MF patients are attaining earlier diagnosis and identifying the 20% of early stage patients at risk for disease progression to a fatal systemic lymphoma. Major goals of our Center are to develop and implement high-throughput T cell receptor (TCR) sequencing (HTS) as a powerful tool to provide more rapid and definitive diagnoses and to discriminate early stage patients who are at high risk for eventual disease progression. With respect to establishing a diagnosis of CTCL, we found that HTS of the T-cell receptor beta gene (TCRB) was more sensitive and specific than TCRγ PCR, detected T cell clones in 46/46 CTCL patients, and successfully discriminated CTCL from psoriasis, eczematous dermatitis and healthy skin. To determine if HTS could discriminate the 20% of early stage MF patients who will ultimately develop progressive disease, we performed HTS of the TCRB gene in lesional skin biopsies from a discovery cohort of 208 patients with cutaneous T-cell lymphoma from a longitudinal observational clinical trial initiated 15 years ago. We performed an identical analysis on an independent validation cohort of 101 CTCL patients from the same observational clinical trial. We found that the strongest predictor of progression in stage I patients was the malignant clone frequency (MCF) in lesional skin. In the discovery cohort patients with early-stage MF, an MCF of >25% was strongly associated with decreased progression-free survival (PFS, HR 4·9, 95% CI 2·5-9·7, p<0·0001) and decreased overall survival (OS, HR 4·8, 95% CI, 2·0-12, p=0·0006). These findings were validated in the independent cohort. HTS is a powerful clinical and research tool that can enhance diagnosis and identify patients at high risk for disease progression.

Funding: These studies were supported by generous charitable contributions from Edward P. Lawrence, Esq. from the Lubin Family Foundation, National Institute of Health (NIH) Specialized Program of Research Excellence (SPORE P50 CA9368305) grant to T.S.K., NIH R01CA203721-01 to R.A.C. and T.S.K., T32 AR-07098-36 to T.S.K. (J.T.O.) and R01 AR063962 NIH/National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS) (to R.A.C.) and NIH/NIAMS P30 AR069625 (to R.A.C.). A.d.M. was supported by grants from the Societe Francaise de Dermatologie, CEDEF, Association pour la Recherche contre le Cancer, Fondation Rene Touraine and the Philippe Foundation. This work was conducted with support for J.T.O from Harvard Catalyst | The Harvard Clinical and Translational Science Center (National Center for Research Resources and the National Center for Advancing Translational Sciences, National Institutes of Health Award UL1 TR001102) and

Title: Neoantigen-specific T-cell receptor repertoires from healthy donors

Authors: Johanna Olweus[1, 2], Weiwen Yang[1, 2], Saskia Meyer[1, 2], Eirini Giannakopoulou[1, 2], Bala Siddaiah Anangi[1, 2], Sachin Kumar Singh[3]
[1]Oslo University Hospital Radiumhospitalet; [2]K.G. Jebsen Center for Cancer Immunotherapy, University of Oslo, [3] Centre for Molecular Medicine Norway, Oslo, Norway

Abstract: Neoantigens encoded by somatic mutations have emerged as attractive targets of cancer immunotherapy. However, although tumors like malignant melanoma on average harbour several hundred mutations per tumor, patient T cells spontaneously respond to only 1-2% of these *in vivo*. We recently demonstrated that T cell repertoires from healthy donors can recognize manifold more mutations from patient tumors when primed *in vitro*. Thus, healthy donor T cell repertoires can be used to broaden the immune response to neoantigens independently of tissue or blood sampling from the patient and unbiased by the immunosuppressive environment of the tumor. We next demonstrated that transfer of T cell receptor (TCR) genes isolated from healthy donor T cells to peripheral blood T cells efficiently mediated recognition of patient cancer cells, providing a rationale for potential use in adoptive T cell therapy of cancer. To facilitate therapeutic development, a major hurdle is to find strategies for selection of the best TCRs. As a next step, we have embarked on the characterization of the T cell receptor repertoires responding to neoantigens in different healthy donors. A high throughput approach yields the sequences of paired alpha and beta TCR chains from several hundred single cells per run. Initial experiments suggest that certain neoantigens select for TCRs containing motifs that are shared between donors. The repertoire diversity is highly variable depending on the reactive neoantigen. Using this approach, we have furthermore identified and characterized a TCR recognizing a recurrent mutation in acute myeloid leukemia.

Funding: This work was funded by the Research Council of Norway, the Regional Health Authorities of South-Eastern Norway, the Norwegian Cancer Society, the University of Oslo and Oslo University Hospital Radiumhospitalet.

Title: B cell Differentiation Pathways in Human Disease

Authors: Tipton, C[1], Tomar, D[1], Hom, J[1], Wei, C[1], Cashman, K[1], Lee, E-H[2], Fucile, C[3], Rosenberg, A[3], Jenks, S[1], Sanz, I [1]. [1] Division of Rheumatology & Lowance Center for Human Immunology. Emory University School of Medicine. Atlanta, GA; [2] Division of Pulmonary and Critical Care & Lowance Center for Human Immunology. Emory University School of Medicine. Atlanta, GA; [3] Department of Microbiology & Immunology. University of Alabama at Birmingham. Birmingham, AL.

Abstract: The cellular pathways and mechanisms of B cell activation and differentiation in human autoimmunity remain to be defined. In part, this limited knowledge is due to imprecise definition of different B cell types representing discreet stages of B cell differentiation. We have developed a comprehensive B cell immunomics approach to address these questions with the limited numbers of cells usually available for human experimentation. This approach includes cellular B cell receptor repertoire analysis coupled with serum autoantibody proteomics; multidimensional flow cytometry; and integrated transcriptomics and epigenetic analysis. We will discuss the use of these tools to define B cell and plasma cell subsets; their activation pathways; and their participation in human autoimmune responses.

Funding: NIAID Emory Autoimmunity Center of Excellence U19; NIAID R37 AI049660; NIAID PO1 AI052689; Georgia Research Alliance; Lowance Center for Human Immunology

Title: BCR repertoire sequencing in patients with primary immunodeficiency analysis

Authors: Johannes Trück[1,2,3], Marie Ghraichy[1,2], Jacob D. Galson[1,2], Dominic F. Kelly[3]

[1]Division of Immunology, University Children's Hospital Zurich, Switzerland
[2]Clinical Research Center, University Children's Hospital Zurich, Switzerland
[3]Oxford Vaccine Group, University of Oxford, United Kingdom

Abstract: Next-generation B-cell receptor (BCR) repertoire sequencing now allows a detailed assessment of the adaptive immune system in health and disease. However, it is mostly unknown how the BCR repertoire is altered in the context of primary immunodeficiencies (PID) and whether findings are consistent across phenotypes and genotypes. In addition, there is the potential of applying BCR repertoire sequencing to patients with (suspected) PID to evaluate B-cell function. In an ongoing study, we investigate heavy chain BCR repertoires of PID patients recruited from the outpatient clinic at the University Children's Hospital Zurich as the national referral centre for primary immunodeficiency in Switzerland. Preliminary findings of our study show that BCR repertoires from PID patients are different compared with those of control samples. However, it is somewhat surprising that BCR repertoires even from severe clinical phenotypes often show only mild abnormalities and that diversity or immunoglobulin gene segment usage is generally preserved to some extent. Repertoire characteristics that are commonly and severely altered in PID patients are isotype subclass usage and rates of somatic hypermutation. There are also repertoire properties that seem to be specific to certain PIDs such as an increased usage of IGHV4-34, which has been linked to self-reactive antibodies. In addition, BCR repertoire features seem to correlate with the severity of the underlying immunodeficiency. Ultimately, BCR repertoire sequencing opens the opportunity to add to current management of PID patients by estimating the functionality of the B-cell system and thereby improving diagnosis and clinical management.

Funding:
Swiss National Science Foundation (JT)
Bangerter-Rhyner Foundation (JT)
Olga-Mayenfisch Foundation (JT)
NIHR Oxford Biomedical Research Centre (DFK)

Title: Molecular Analysis of the Serum Antibody Repertoire

Author: George Georgiou
Depts of Chemical Engineering, Biomedical Engineering and Molecular Genetics and Biology
University of Texas, Austin

Abstract: We have now developed an integrated technology workflow that combines high mass accuracy LC-MS/MS proteomics, with microfluidics and bioinformatics tools for the molecular-level deconvolution of the identities, relative amounts and functions of antibodies in human blood and also for the delineation of the relationships between antibody production and the relevant B cell immunological mechanisms. This technology is providing invaluable insights on: (i) how the composition of the serum antibody repertoire is shaped by infection or vaccination; (ii) the discovery of highly potent therapeutic antibodies directly from the analysis of the serum repertoire in convalescent patients and (iii) the environmental and genetic factors that determine antibody immunity in humans. This presentation will summarize key findings on the composition of the serological antibody repertoire (i) following vaccination with seasonal flu, IPV or an experimental NoV vaccine: (ii) in persistent infection and (iii) in cancer, with emphasis on how the antibody repertoire is informing on vaccination strategies and on the discovery of novel therapeutic antibodies.

Funding:

Title: Diversity of Non-Human Primate and Human Germline Ig Genes

Authors: G. B. Karlsson Hedestam[1], Néstor Vázquez Bernat[1], G. E. Phad[1], Cathrine Scheepers[2,3], Lynn Morris[2,3], Mats Ohlin[4]; M. Martin[5] and Martin M. Corcoran[1]

[1]Department of Microbiology, Tumor and Cell Biology, Karolinska Institutet, Sweden, [2]Centre for HIV and STIs, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa; [3]School of Pathology, University of the Witwatersrand, Johannesburg, South Africa; [4]Department of Immunotechnology, Lund University, Sweden [5]Department of Biochemistry and Biophysics, SciLifeLab, Stockholm University, Sweden

Abstract: B lymphocytes of the adaptive system express highly polymorphic receptors that allow recognition of large numbers of antigens. The specific germline-encoded variable (V), diversity (D), and junction (J) gene segments that form the basis for all rearranged mature B cell receptor repertoires vary between individuals, which impacts antigen-specific antibody responses. By applying Next Generation Sequencing (NGS) and analysis using the IgDiscover software, we are able to produce complete individualized immunoglobulin heavy and light chain germline databases of multiple individuals in a matter of weeks. By analyzing humans of different ethnic background, we are uncovering significant heterogeneity in antibody germline genes, both at the allelic and structural levels, diversity that cannot be easily investigated with conventional whole genome sequencing approaches due to the complexity of the Ig loci. Further, because IgDiscover is based on the analysis of expressed VDJ repertoires, it allows the identification of haplotypes, offering additional information from each individual. While we observe a high level of variation in human antibody germline genes, we record even higher diversity for rhesus and cynomolgus macaques, with only few antibody V alleles shared between animals. This has implications for vaccine studies performed in macaques and for the use of current public databases of antibody germline genes, which so far are limited for macaques. By applying IgDiscover, we have produced individualized databases of germline V alleles from over 20 macaques of different origin, forming the basis for an improved macaque Ig germline gene database. The generation of individualized databases from vaccinated subjects (humans or non-human primates) offers a robust starting point for antibody lineage tracing studies aimed at defining how antigen-specific B cell responses evolve in response to infection and immunization, or during disease development.

Funding: This work was supported by the Swedish Research Council (Grant #C0331501), the European Union's Horizon 2020 Research and Innovation programme (Grant #681137) and the National Institutes of Health (Grants # AI104722 and #1U01AI136677-01).

Title: The antibody response of celiac disease

Authors: Rasmus Iversen, Roberto Di Niro, Ralf S. Neuman, Bishnuedo Roy, Omri Snir, Jorunn Stamnaes, Øyvind Steinsbø, Ludvig M. Sollid. Centre for Immune Regulation and KG Jebsen Coeliac Disease Research Centre, Department of Immunology, University of Oslo and Oslo University Hospital, Oslo, Norway.

Abstract: Celiac disease is prevalent disorder caused by an inappropriate immune response to cereal gluten proteins. The disease has a strong genetic basis. The chief genetic determinants are HLA-DQ2 and HLA-DQ8 encoding genes. CD4 T cells of the patients recognize gluten peptides presented by the disease-associated HLA-DQ molecules following posttranslational modification by the enzyme transglutaminase 2 (TG2). The celiac patients, when consuming gluten, have highly disease-specific antibodies to TG2 as well as to deamidated gluten peptides (DGP). The dual role of TG2 creating gluten T-cell epitopes and being the target of an autoantibody response is hardly coincidental. Conceivably, this can be explained by a hapten carrier like mechanism that involves complexes of TG2 and gluten peptides and which enables gluten-specific T cells to provide help to TG2-specific B cells.

The gut celiac lesion is characterized by plasmacytosis. Characterization of plasma cells of the lesion has revealed abundance of plasma cells being specific for TG2[1] (approx. 10%) and DGP[2] (approx 1%). The plasma cells specific for TG2 or DGP display a restricted VH/VL gene usage with low degree of somatic hypermutation[1-5] indicating similarities in development of responses to the two antigens. Proteomic analysis of serum IgA specific for either TG2 or DGP has revealed antigen-specific V-gene preferences that match those found in gut plasma cells.[6] Further, gut plasma cell CDR-H3 sequences are abundant in serum IgA, but also detectable in serum IgG suggesting that the same B cell clones that give rise to gut plasma cells also contribute to the serum antibody pool. Notwithstanding, as the antigen-specific serum IgA antibodies are mostly monomeric IgA and antibodies secreted in the gut are mostly dimeric, this suggests that individual B-cell clones give rise to different plasma cell populations.

References:

1 Di Niro, R. *et al. Nat. Med.* **18**, 441-445 (2012).
2 Steinsbo, Ø. *et al. Nat. Commun.* **5**, 4041 (2014).
3 Snir, O. *et al. JCI Insight* **2** (2017).
4 Snir, O. *et al. J. Immunol.* **194**, 5703-5712 (2015).
5 Roy, B. *et al. J. Immunol.* **199**, 782-791 (2017).
6 Iversen, R. *et al. Cell Rep.* **20**, 2357-2367 (2017).

Funding: The Research Council of Norway through its Centre of Excellence funding scheme (project number 179573/V40), the South-Eastern Norway Regional Health Authority (projects 2011050, 2014045, and 2016113), the European Commission (projects ERC-2010-Ad-268541 and FP6-2005-FOOD-4B-36383) and Stiftelsen KG Jebsen (SKGH-MED-017).

Title: *iReceptor*: A Platform for Storing and Sharing AIRR-seq Data

Authors: Felix Breden[1,2], Nishanth Marthandan[1,3], Bojan Zimonja[1], Jerome Jaglale[1], Nicole Knoetze[1], Emily Barr[1], Frances Breden[1], Yang Zhou[1], Richard Bruskiewich[1], Jamie K. Scott[1,3,4], and Brian Corrie[2]
[1]The IRAMCS Centre, [2]Dept. of Biological Sciences, [3]Dept. of Molecular Biology and Biochemistry, and [4]Faculty of Health Sciences, Simon Fraser University, Burnaby, British Columbia, V5A 1S6, Canada

Abstract: Recent advances in next generation sequencing (NGS) technologies have made it possible to sample the human antibody (Ab)/B-cell (BcR) and T-cell receptor (TcR) repertoires in tremendous detail. A typical human has ~10^9 B cells at any given time, with a very large percentage of these cells changing very rapidly over time. It is this rapid adaptation of our immune system over time that makes it so effective at fighting pathogens. NGS technologies allow investigators to sequence the immune receptor genes in essentially all the B and T cells in an individual, allowing researchers to explore how the immune system adapts. Adaptive Immune Receptor Repertoire sequencing (AIRR-Seq) has enormous promise for understanding the immune repertoire dynamics in vaccinology, infectious disease, autoimmunity, and cancer biology.

The iReceptor system is a platform to integrate and analyze these immense data sets by combining: 1) an international network of AIRR-Seq data repositories; 2) the ability to federate AIRR-Seq data across these distributed repositories; 3) advanced analytical tools unique to AIRR-Seq data; and 4) a scientific gateway that hides the complexity of performing research queries and advanced analyses across these federated data. iReceptor enables data federation across many distributed data sets by defining a web based API that implements the emerging AIRR Community (airr-community.org) data standard. In this tools demonstration, we will demonstrate the iReceptor Scientific Gateway and its ability to federate and explore data across multiple, distributed, AIRR compliant repositories.

Funding: Current: Canada Foundation for Innovation – Challenge 1
Past: CANARIE Network Enabled Platforms Program

Title: High-throughput, single cell T-cell receptor profiling by SMART technology

Authors: Magnolia Bostick[1], Sarah Taylor[2], Ishminder Mann[1], Nao Yasuyama[3], Alain Mir[1], Andrew Farmer[1] [1]Takara Bio USA, Inc. Mountain View, CA; [2]currently 10X Genomics Pleasanton, CA; [3]currently Takara Bio, Inc. Kusatsu, Shiga, Japan.

Abstract: Single-cell T-cell receptor (scTCR) clonotype analysis permits the determination of the specific TCR alpha-beta (α/β) chain pairing expressed on each cell. This pairing information allows researchers to gain insight into T-cell heterogeneity and plasticity, determine the contribution of the pairing to antigen specificity of the individual TCR, and design therapeutic antibodies. Here we employ a 5'-RACE-like approach and SMART® technology, in conjunction with two novel next-generation sequencing (NGS) library preparation kits, using the same primer pairs, to capture full-length variable regions of TCR-α and -β chains. Method 1, using the SMARTer® Human scTCR a/b Profiling Kit, permits NGS library preparation of FACS-sorted cells in 96-well plates. We present data showing α/β pairing from Jurkat, CCRF, PBMCs, and CD4+ T cells. In addition to the sensitivity of this method, the ability to pool the cDNA from 96 wells into 12 sequencing libraries adds to the ease of use. Consistent with immunology reports, unstimulated CCRF-CEM cells examined with this kit expressed aTCR-β but not aTCR-α chain. Method 2 is an adaptation of the process above that scales to ~1,200 single cells using the SMARTer™ ICELL8® platform, which enables single-cell isolation and nanoliter PCR in a nanowell chip. For proof-of-principle studies, Jurkat cells and CCRF-CEM cells were processed using an ICELL8 chip preprinted with barcoded oligos. Paired TCR α/β Jurkat clonotypes were detected in 77% and 87% cells in mixed and single cell populations, respectively. The ability of the core biochemistry and PCR components of these kits to be used with either FACS-sorted cells in 96-well plates or >1,000 cells in novel ICELL8 chips (in development) points to the general utility and scalability of this approach in understanding paired scTCR clonotype diversity.

Funding:

Title: SHMPrep: A Fast and User-Friendly Program for Preprocessing Paired-end Immunoglobulin Sequence Data from Illumina Platforms

Author: Stephen Meier, Postdoctoral Associate, Stony Brook University

Abstract: Next generation sequencing platforms, predominantly the Illumina MiSeq, are now routinely used to sequence immunoglobulin (Ig) genes to measure effects such as somatic hypermutation (SHM). There is a need for software that can be used widely by experimentalists who may not have computational expertise or access to specialized computers. We have developed SHMPrep, a very fast and user-friendly program for preprocessing paired-end immunoglobulin data from the Illumina MiSeq. The program aligns paired-end reads, performs error correction and collapses sequences. Implementation of efficient string matching algorithms such as suffix automata and string alignment algorithms such as k-band global alignment, enable processing of an entire MiSeq run of ~20 million reads in approximately 30 minutes using a standard desktop computer. SHMPrep can also handle barcodes (UMI) and multiple sets of primers, making it ideal for working with Ig repertoire data. The program produces outputs in FASTA/Q format that can be further processed using tools such as IMGT. The software is composed of a kernel written in the high-performance C++ language and a user-friendly Graphical User Interface (GUI) developed in Java that is designed for use by non-computational experts. The software is available for Windows, Mac and Linux. The software is available at: http://www.ams.sunysb.edu/~maccarth/software.html

Funding:

Title: Performance-optimized partitioning of clonotypes from high-throughput immunoglobulin repertoire sequencing data

Author: Nima Nouri, Department of Pathology, Yale School of Medicine

Abstract: During adaptive immune responses, activated B cells expand and undergo somatic hypermutation of their immunoglobulin (Ig) receptor, forming a clone of diversified cells that can be related back to a common ancestor. Identification of B cell clonotypes from high-throughput Adaptive Immune Receptor Repertoire sequencing (AIRR-seq) data relies on computational analysis. Recently, we proposed an automate method to partition sequences into clonal groups based on single-linkage clustering of the Ig receptor junction region with length-normalized hamming distance metric. This method could identify clonally-related sequences with high confidence on several benchmark experimental and simulated data sets. However, this approach was computationally expensive, and unable to provide estimates of accuracy for new data. A new method is presented that address this computational bottleneck and also provides a study-specific estimation of performance, including sensitivity and specificity. The method uses a finite mixture modeling fitting procedure for learning the parameters of two univariate curves which fit the bimodal distributions of the distance vector between pairs of sequences. These distributions are used to estimate the performance of different threshold choices for partitioning sequences into clonotypes. These performance estimates are validated using simulated and experimental datasets. With this method, clonotypes can be identified from AIRR-seq data with sensitivity and specificity profiles that are user-defined based on the overall goals of the study. This new procedure has been implemented under the “findThreshold” function as part of the SHazaM R package in the Immcantation framework: www.immcantation.com.

Funding:

Title: Representing Epitope Specific Antibody and T cell Receptor Sequences in the Immune Epitope Database

Authors: Randi Vita[1], Swapnil Mahajan[1], Deborah Shackelford[1], Jerome Lane[1], Veronique Schulten[1], Alessandro Sette[1], Bjoern Peters[1]
[1]La Jolla Institute for Allergy and Immunology, La Jolla, California, USA

Abstract: The Immune Epitope Database (IEDB) is a free public resource presenting experimental immune epitope data. The IEDB contains experiments describing over 380,000 epitopes derived from more than 18,800 publications. This information is captured in up to 400 data fields per experiment. We recently extended our ability to capture epitope specific antibodies and T cell receptors, which are increasingly becoming available from receptor sequencing experiments. This required adding to previously existing fields that capture full-length receptor sequences to now also store CDR sequences, which are frequently the only data available. We also capture VDJ gene usage. Our data includes both author provided CDR and gene usage information as well as calculating these fields directly from the full-length sequence to maintain uniformity of CDR sequences using IMGT numbering and VDJ gene nomenclature format. We are grouping data from receptors of the same type with identical CDR3 sequences together to reduce reports of redundant results, while retaining the ability to inspect the outcome of individual experiments. As of November 2017, we have catalogued sequence information for more than 20,000 receptors, shown to bind to more than 2000 specific epitopes. These data are accessible as full exports and through a new dedicated query interface that combines the search for specific receptor characteristics with the existing IEDB query interface to allow search by, for example, epitopes from a specific infectious agent, recognized by a specific host, or recognized by a certain type of immune response.

Funding: NIH Contract no. HHSN272201200010C

Title: Measurement of B-cell clone size distributions in bulk populations

Authors: Aaron Rosenfeld[1]*, Dora Chen[2]*, Wenzhao Meng[2], Bochao Zhang[1], Tomer Granot[3], Donna Farber[3], Uri Hershberg[1,4] and Eline T. Luning Prak[2]
[1]School of Biomedical Engineering, Science and Health Systems, Drexel University, Philadelphia, PA, USA
[2]Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA
[3]Columbia Center for Translational Immunology, Columbia University, New York, NY, USA
[4]Department of Microbiology and Immunology, Drexel College of Medicine, Drexel University, Philadelphia, PA, USA
*co-first authors

Abstract: Expansion and contraction of B-cell clones occur during adaptive immune responses. B-cell clones can also persist or grow uncontrollably in lymphoproliferative and malignant disorders. To monitor and track B-cell clones of interest typically requires large-scale sampling of bulk populations, and is often accomplished by next generation sequencing. Knowing whether a specific clone is likely to be present or absent requires estimating the size of the clone and powering the analysis to detect clones of that size in the population. Here we describe different methods for estimating clone size and compare clonal size and diversity across spleen samples from 9 different human organ donors. We chose the spleen because it contains a mixture of small and large B-cell clones. Using genomic DNA of mononuclear cells and normalizing for the B-cell fraction (CD19+CD20+CD3-lymphocytes), we amplified and sequenced antibody heavy chain V region rearrangements and grouped related sequences into clones using ImmuneDB. We compared the rankings of different clone size measures including copy number fraction, unique sequence fraction and number of instances (number of times that members of the same clone are found in independently amplified sequencing libraries from the same sample). While for most very large clones these measures correlated, there were clones (for example, large un-mutated clones) where copy number or instance-based measures exhibited different ranks than unique sequence-based ranks. Of note, we also found differences in the frequency distributions of clone sizes in different organ donors, suggesting that B-cell clone sizes achieve different set-points in individuals. Inter-individual differences in B-cell clone size and clone representation also varied across the clone size continuum, underscoring the need to visualize the repertoire landscape with metrics that adequately describe clones in different scales of measurement.

Funding: NIH P01 AI106697, NIH P30 CA016520

Title: immuneACCESS™: A Public Repository for TCR and BCR Sequencing Data Using the immunoSEQ® Analyzer

Authors: Ian Kaplan[1], Jeff Darga[1],
[1]Adaptive Biotechnologies

Abstract: Adaptive Biotechnologies launched the immuneACCESS™ web portal to help researchers view, analyze, and most importantly, share immune repertoire data. immuneACCESS[1] is the world's largest public data repository for T- and B-cell receptor sequences; this ever-growing collection of immunoSEQ Data includes over 5,500 human samples and more than 460 mouse samples, that combine to represent 545,213,786 nucleotide sequences. These datasets come from research projects in basic immunology, cancer immunotherapy, autoimmune disease, infectious disease, and more. One of the key datasets is a cohort of 666 healthy adults with known cytomegalovirus serostatus[2]; and one of the most unique is the 37 million unique B-cell receptor sequences from three healthy adult donors—which is many fold deeper than any existing resource[3].

These freely available datasets are accessible in the immunoSEQ Analyzer as a software service designed specifically to support the analysis needs of researchers working with large immunosequencing datasets. immuneACCESS Data can be used to create new projects within the Analyzer, or exported for further analysis on other software platforms. The goal of immuneACCESS is to connect researchers throughout the immunosequencing community and accelerate discoveries by providing a platform to simplify data sharing and analysis.

References

1. www.immuneaccess.com
2. https://clients.adaptivebiotech.com/pub/emerson-2017-natgen
3. https://clients.adaptivebiotech.com/pub/robins-bcell-2016

Funding: N/A

Title: T- and B-cell receptor repertoire sequencing at the Academic Medical Center

Authors: Barbera D. C. van Schaik[1,*], Paul L. Klarenbeek[2,3], Marieke E. Doorenspleet[2,3], Sabrina Pollastro[2], Anne Musters[2], Giulia Balzaretti[2,4], Rebecca E. E. Esveldt[2], Frank Baas[5], Niek de Vries[2,4], Antoine H. C. van Kampen[1,6]
[1]Bioinformatics Laboratory, Clinical Epidemiology, Biostatistics and Bioinformatics, Academic Medical Center, Amsterdam, the Netherlands. [2]Amsterdam Rheumatology and Immunology Center, Department of Clinical Immunology & Rheumatology, Academic Medical Center, Amsterdam, the Netherlands. [3]Dept. of Genome Analysis, Academic Medical Center, Amsterdam, The Netherlands. [4]Dept. of Experimental Immunology, Academic Medical Center, Amsterdam, the Netherlands. [5]Department of Clinical Genetics, Leiden University Medical Center, Leiden, The Netherlands. [6]Biosystems Data Analysis, Swammerdam Institute for Life Sciences, University of Amsterdam, the Netherlands.
* b.d.vanschaik@amc.uva.nl

Disease studies

We use repertoire sequencing to identify T- and B-cell clones that play a role in auto-immune diseases [1], or to find clones that can be used for the identification of individuals that are at risk for rheumatoid arthritis [2]. In addition, we used repertoire sequencing to understand anti-viral responses [3], for gaining insight into immune responses in healthy donors [4], and for identifying in-variant T-cells [5]. Laboratory protocols were developed to study the T- and B-cell repertoire in human donors and mouse models.

Clone identification and quality assessment

All repertoire sequencing data is analysed with RESEDA [6], which is an in-house developed collection of methods. It contains methods for quality assessment including the detection of sample contamination based on overlapping clones between non-related samples sequenced in a single run. For each clone, RESEDA identifies the CDR3 region and assigns the V and J genes. Finally, it allows to select specific subsets of samples and clones.

Downstream analysis

After assignment of the V, CDR3 and J regions we usually examine the overlap of (dominant) clones between case and control samples, paired samples, or time series. Furthermore, we analyze somatic mutations in B-cell receptors, and are currently developing a method for reconstructing B-cell lineages from repertoire sequence data.

Research data management

In the past ten years we have performed ~90 NGS experiments with 100-200 samples per run. The data throughput requires automation of the data analysis and a system for long term storage and retrieval of samples. The data analysis is performed on a community cloud using a pilot job framework. Parameters for the data analysis are supplied via a sample sheet and records are kept about software versions. We participate in a local research data management project to store all the raw sequence data on an iRods system. Descriptions can be attached to directories and files in this system and these are searchable. This will enable reuse of these samples in future projects.

Share experiences

We would like to use this poster session to discuss best practices for analysis of immune repertoire sequencing experiments. We will present our current approach and studies and are interested in your experiences.

[1] Klarenbeek *et al.*, Ann Rheum Dis, 2012
[2] Tak *et al.*, Ann Rheum Dis, 2017
[3] Klarenbeek *et al.*, PLoS Pathog, 2012
[4] Klarenbeek *et al.*, PLoS One, 2015
[5] van Schaik *et al.*, J Immunol, 2014
[6] https://bitbucket.org/barbera/reseda

Title: Using genotype abundance to improve phylogenetic inference

Authors: William DeWitt III[1], Luka Mesin[2], Gabriel Victora[2], Vladimir Minin[3], Frederick Matsen IV[1]
[1]Fred Hutchinson Cancer Research Center; [2]Rockefeller University; [3]UC Irvine

Abstract: Modern biological techniques enable very dense genetic sampling of unfolding evolutionary histories, and thus frequently sample some genotypes multiple times. This motivates strategies to incorporate genotype abundance information in phylogenetic inference. In this paper, we synthesize a stochastic process model with standard sequence-based phylogenetic optimality, and show that tree estimation is substantially improved by doing so. Our method is validated with extensive simulations and an experimental single-cell lineage tracing study of germinal center B cell receptor affinity maturation.

Funding: NIH R01 GM113246, NSF CISE-1564137 and NSF CISE-1561334, and a Faculty Scholar grant from the Howard Hughes Medical Institute and the Simons Foundation.

Title: Screening for (auto)-antigen specific $CD8^+$ T cell receptor chains via TCRα NGS: a novel approach to assess TCR sequence relevance to T1 Diabetes

Authors: Yannick Fuchs[1], Anne Eugster[1], Sevina Dietz[1], Denise Kühn[1], Carmen Wilhelm[1], Annett Lindner[1], Anita Gavrisan[4], Anette-G. Ziegler[3,4] and Ezio Bonifacio[1,2]
[1]Center for Regenerative Therapies Dresden (CRTD), Dresden, Germany; [2] Paul Langerhans Institute Dresden, Germany; [3] Forschergruppe Diabetes e.V., Ingolstädter Landstraße 1, 85764 Neuherberg, Germany; [4] Institute of Diabetes Research, Helmholtz Zentrum München, and Forschergruppe Diabetes, Klinikum rechts der Isar, Technische Universität München, Neuherberg, Germany

Abstract: Autoreactive $CD8^+$ T cells are key mediators of the pancreatic beta cell destruction finally leading to type 1 diabetes (T1D). T cell receptor (TCR)-mediated recognition of (auto-) antigenic peptides presented on MHC class I molecules is a prerequisite for $CD8^+$ T cell mediated target cell destruction. The identification of TCRs directed against islet-autoantigens is a potential basis to follow up pathogenic T cells in a molecular fashion. Here we used MHC class I multimers to identify, isolate and characterize $CD8^+$ T cells directed against the beta cell antigen IGRP. Single cell TCR sequencing revealed $IGRP_{265-273}$ specific TCRs and identified TCRα chains of dominant clonotypes that were shared between T1D patients. Expanded $IGRP_{265-273}$ specific $CD8^+$ T cell clones were shown to kill peptide loaded target cells in an antigen-specific manner and were further charactererized via multiparameter single cell gene expression analysis. Applying a novel TCRα next generation sequencing (NGS) approach to naïve and memory $CD8^+$ T cells populations of HLA-A*0201 positive donors we assessed the abundance of these identified TCRα chains at varied stages before and after onset of autoimmunity and T1D.

Funding: The work was supported by grants from the Kompetenznetz Diabetes Mellitus (Competence Network for Diabetes Mellitus) funded by the Federal Ministry of Education and Research (FKZ 01GI0805) and the Juvenile Diabetes Research Foundation (JDRF; JDRF-No 17-2012-16), and by grants from the German Federal Ministry of Education and Research (BMBF) to the German Center for Diabetes Research (DZD e.V.).

Title: Initiation of antiretroviral therapy in the setting of HIV infected individuals increases public T cell responses

Authors: Andrea M.H. Towlerton [1], David Coffey [1,2], Scott Christley[3], Jared Ostmeyer[3], Inimary Toby[3], Lindsay G. Cowell[3], Edus H. Warren [1,2]

[1]Fred Hutchinson Cancer Research Center, Seattle, Washington, USA. [2]University of Washington, Seattle, Washington, USA. [3]University of Texas Southwestern Medical Center, Department of Clinical Sciences, Dallas, Texas

Abstract:

Background: An estimated 36.7 million people worldwide are infected with Human immunodeficiency virus (HIV). Although advancements in the prevention, treatment and management of HIV have been made, the restoration of the immune system following antiretroviral therapy (ART) remains ambiguous. We hypothesized that initiation of ART will alter the peripheral blood T cell repertoire through expansion and diversification of CD4+ T cells.

Methods: Pre- and post-ART peripheral blood mononuclear cells (PBMCs) samples were serially collected from 30 HIV+ individuals over 8-10 years through the Center for AIDS Research (CFAR) at the University of Washington, Seattle. Per patient, one to 4 time points were collected pre-ART and 2-6 collected post ART. For comparison, 16 PBMC samples were collected from heathy HIV-seronegative controls. High-throughput, survey level T-cell receptor β chain (*TRB*) sequencing using the ImmunoSeq assay (Adaptive Biotechnologies) was performed on all samples. Analyses were conducted using the LymphoSeq R package (http://bioconductor.org/packages/LymphoSeq). T cell repertoire diversity was estimated using the clonality score, a derivative of the Shannon entropy index. Public T cell clones with known antigen specificity to viral and human epitopes were identified from a curated database of 9,606 published *TRB* sequences built into the LymphoSeq package.

Results: Patients were stratified as responders or non-responders based on the detection of HIV RNA 100 days after the start of ART. Prior to ART, the *TRB* repertoire appeared significantly more clonal than the healthy controls. Following initiation of ART, we found a significant increase in T-cell diversity which accompanied the improved CD4 count and was particularly apparent among the treatment responders. To investigate the antigen specificity of a subset of these expanding T cells, we identified “public” T cell clones with known antigenic specificity. Thus, we identified a significant increase of “public” T cell clones, partially among those with specificity for HIV epitopes.

Funding: Biologic Determinants of the Natural History of AIDS-Defining Cancers in Uganda
NIH/NCI - U54CA190146
9/18/2014-8/31/2019

Title: BraCeR: A computational tool for reconstruction of B-cell receptor sequences and inference of clonality from single-cell RNA-seq data

Authors: Ida Lindeman[1,2], Guy Emerton[2], Ludvig M. Sollid[1,3], Sarah A. Teichmann[2,4], Michael J.T. Stubbington[2] [1]Centre for Immune Regulation (CIR) and Department of Immunology, University of Oslo and Oslo University Hospital, Oslo, Norway; [2]Wellcome Trust Sanger Institute, Cambridge, UK; [3]KG Jebsen Coeliac Disease Research Centre, University of Oslo, Oslo, Norway; [4]Theory of Condensed Matter, Cavendish Laboratory, Cambridge, UK.

Abstract: Reconstruction of antigen receptor sequences from single-cell RNA-sequencing (scRNA-seq) data allows the linking of antigen receptor usage to the full transcriptomic identity of individual B lymphocytes, without having to perform additional targeted repertoire sequencing. We developed BraCeR (freely available at https://github.com/teichlab/bracer/), an extension of TraCeR (Stubbington *et al.*, *Nat Methods*, 2016) for reconstruction of paired full-length antigen receptor sequences and inference of clonality from scRNA-seq data. BraCeR takes into account the inherent differences between B-cell receptor (BCR) and T-cell receptor (TCR) sequences, particularly the presence of kappa or lambda light chains, isotype switching and the process of somatic hypermutation (SHM).

BraCeR can be applied to scRNA-seq data created by any protocol that sequences full-length mRNA. We tested BraCeR against experimental human and mouse scRNA-seq data with various SHM rates and repertoire diversities. The reconstruction efficiency was similar to or higher than that of BASIC, a previously reported tool for BCR reconstruction (Canzar *et al.*, *Bioinformatics*, 2017), with similar reconstruction accuracies. Importantly, BraCeR allows for reconstruction of multiple chains present in a cell, and it can identify non-productively rearranged chains. BraCeR also reports and removes potential cell multiplets from downstream analyses.

BraCeR provides a complete pipeline for clonal inference and lineage tracing of B cells; raw scRNA-seq reads can be processed all the way to clonal networks and lineage trees. We applied BraCeR to published scRNA-seq datasets and demonstrated its ability to infer clonality from reconstructed sequences. Our approach establishes clonal relationships based on paired heavy and light chain sequences, and allows linkage of antigen specificity with the full transcriptomic profile of each B cell. BraCeR also creates tab-delimited database files compatible with the Immcantation portal tool suite, thus facilitating further analyses of the reconstructed BCR sequences.

Funding: IL was funded by a grant from the University of Oslo to LMS. LMS was further supported by grants from the Research Council of Norway through its Centre of Excellence funding scheme (project number 179573/V40) and from Stiftelsen KG Jebsen (SKGH-MED-017). MJTS, GE and SAT were supported by the Wellcome Trust Grant 206194.

Title: Kymouse™ + IntelliSelect™ = a winning combination for antibody discovery

Authors: Jacob D. Galson[1], Simon Watson[1], Nick England[1], Thomas Gallagher[1], Ian Fairman[1], Liam Smith[1], David Melvin[1], Qi Liang[1], Allan Bradley[1]
[1]Kymab Ltd, Cambridge, UK

Abstract: Kymab's transgenic Kymouse platform, which contains an extensive repertoire of human immunoglobulin genes, provides an advanced in vivo system for the rapid discovery of fully human therapeutic monoclonal antibodies. When immunized with target antigens, the mice mount normal immune responses, and produce affinity-matured antibodies from these human building blocks. To interrogate these immune responses and find the best antibody drug candidates, we have now developed IntelliSelect which we present here.
IntelliSelect combines an automated laboratory and bioinformatic pipeline for data generation and a user interface for exploration of these data. Antigen-specific single cell sorting and sequencing allow rapid generation of paired VH:VL sequence data from 1,000s of antibodies at a time. In parallel, direct transfection and expression of these sequences into mammalian cells allows functional assays to be conducted. Next, deep sequencing of the entire B cell repertoire is performed and computationally mapped onto the functional data. This combination of the antigen-specific functional data with information from the total repertoire allows us to deeply mine the entire antibody space of the mouse. When promising leads are identified, more advanced antibodies can then be discovered by network analysis and tested through synthesis. Modeling the processes of affinity maturation within the mice also provides an opportunity to finely tune the properties of the lead antibodies to find those with potent therapeutic properties, and optimal biophysical attributes.

The Kymouse combined with the IntelliSelect antibody discovery platform allows efficient discovery of human antibodies with no lead optimization required. This approach has successfully been used to rapidly discover functional antibodies against a range of therapeutic targets. The large datasets generated can now be put to further use for informing novel algorithms for more advanced repertoire mining approaches.

Funding:

Title: T-cell receptor-β V and J usage, in combination with particular HLA class I alleles, correlates with cancer survival patterns

Authors: Blake M. Callahan[1], John M. Yavorski[1], Yaping N. Tu[1], Wei Lue Tong[1], Jacob C. Kinskey[1], Kendall R. Clark[1], Timothy J. Fawcett[2], and George Blanck[1,3,*]

[1]Department of Molecular Medicine, Morsani College of Medicine; and [2]Research Computing, University of South Florida; and [3]Immunology Program, H. Lee Moffitt Cancer and Research Institute, Tampa, Florida

Abstract: Class I and class II HLA proteins, respectively, have been associated with subsets of V(D)J usage resulting from recombination of the T-cell receptor (TcR) genes. Also, particular HLA alleles, in combination with dominant TcR V(D)J recombinations, have been associated with several autoimmune diseases. The recovery of TcR recombination reads from tumor specimen exome files has allowed relatively rapid assessments of V(D)J usage, likely for cancer resident T-cells, across large cancer datasets. The results from this process have permitted an extensive alignment of the indicated TcR-β VDJ usage and HLA class I alleles, also available from the cancer exome files. Results indicate the correlation of both better and worse cancer survival rates with particular TcR-β, V or J usage/HLA class I allele combinations, with differences in median survival times ranging from 7-35 months, depending on the cancer type and the TcR-β V or J usage/HLA class I allele combination.

Funding: N/A

Title: Defining the germline IGHV repertoire in South African individuals

Authors: Cathrine Scheepers[1,2], Alaine Marsden[1,2], Martin Corcoran[4], Bronwen Lambson[1,2], Penny Moore[1,2,3], Gunilla Karlsson-Hedestam[4], Lynn Morris[1,2,3]
[1]Centre for HIV and STIs, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa; [2]School of Pathology, University of the Witwatersrand, Johannesburg, South Africa, [3]Centre for the AIDS Programme of Research in South Africa, KwaZulu-Natal, South Africa, [4]Department of Microbiology, Tumor and Cell Biology, Karolinska Institutet, Sweden

Abstract: The human immunoglobulin heavy chain variable region (IGHV) genes are the most diverse of all immunoglobulin genes and form a major part of the antigen binding portion of functional antibodies. Defining IGHV diversity has major benefits for biomedical research involving antibody therapeutics and vaccine design focused on engaging specific antibody lineages such as broadly neutralizing antibodies against HIV. Using MiSeq technology we have previously identified novel germline IGHV alleles as well as alleles only seen in rearranged VDJ sequences, many of which are expressed in the antigen naive IgM repertoire. Thus far, through cloning and Sanger sequencing of full-length IGHV1 and IGHV2 V-genes (including the L-part 1, the intron and V-exon) we have validated 5 of these alleles. It is our aim to sequence the remaining 5 IGHV subgroups and validate the most frequently used novel alleles. In addition, we have sequenced the full-length V-genes of some truncated IMGT sequences. These full-length V-gene sequences have highlighted additional diversity within the L-part 1, V-heptamer, V-spacer and V-nonamer regions, where the same allele (as defined by the V-exon) has different recombination signals, which may impact the overall VDJ repertoire. This data validates some of the novel alleles identified through NGS data and suggests additional variation in the IGHV region that has not been previously studied. Thus, improvements in technology and antibody repertoire sequencing continue to identify and address gaps in our understanding of antibody diversity.

Funding: This work was supported by the National Institute of Allergy and Infectious Diseases, H3Africa Grant (1U01AI136677-01)

Title: Quantitative Analysis of Repertoire Scale Immunoglobulin properties in Vaccine-Induced B-cell Responses

Authors: Sidhartha Chaudhury[1], Ilja V. Khavrutskii[1], Sabrina M. Stronsky[2] , Donald W. Lee[1] , Jacqueline G. Benko[2] , Anders Wallqvist[1] , Sina Bavari[2] and Christopher L. Cooper[2]
1. Biotechnology High Performance Computing Software Applications Institute (BHSAI), Telemedicine and Advanced Technology Research Center, United States Army Medical Research and Materiel Command, Fort Detrick, MD, USA; 2. Molecular and Translational Sciences, United States Army Medical Research Institute of Infectious Diseases (USAMRIID), Fort Detrick, MD, USA

Abstract: Recent advances in the next-generation sequencing of B-cell receptors (BCRs) enable the characterization of humoral responses at a repertoire-wide scale and provide the capability for identifying unique features of immune repertoires in response to disease, vaccination, or infection. Immunosequencing now readily generates 103 –105 sequences per sample; however, statistical analysis of these repertoires is challenging because of the high genetic diversity of BCRs and the elaborate clonal relationships among them. To date, most immunosequencing analyses have focused on reporting qualitative trends in immunoglobulin (Ig) properties, such as usage or somatic hypermutation (SHM) percentage of the Ig heavy chain variable (IGHV) gene segment family, and on reducing complex Ig property distributions to simple summary statistics. However, because Ig properties are typically not normally distributed, any approach that fails to assess the distribution as a whole may be inadequate in (1) properly assessing the statistical significance of repertoire differences, (2) identifying how two repertoires differ, and (3) determining appropriate confidence intervals for assessing the size of the differences and their potential biological relevance. To address these issues, we have developed a technique that uses Wilcox' robust statistics toolbox to identify statistically significant vaccine-specific differences between Ig repertoire properties. The advantage of this technique is that it can determine not only whether but also where the distributions differ, even when the Ig repertoire properties are non-normally distributed. We used this technique to characterize murine germinal center (GC) B-cell repertoires in response to a complex Ebola virus-like particle (eVLP) vaccine candidate with known protective efficacy. The eVLP-mediated GC B-cell responses were highly diverse, consisting of thousands of clonotypes. Despite this staggering diversity, we identified statistically significant differences between non-immunized, vaccine only, and vaccine-plus-adjuvant groups in terms of Ig properties, including IGHV-family usage, SHM percentage, and characteristics of the BCR complementarity-determining region. Most notably, our analyses identified a robust eVLP-specific feature—enhanced IGHV8-family usage in B-cell repertoires. These findings demonstrate the utility of our technique in identifying statistically significant BCR repertoire differences following vaccination. More generally, our approach is potentially applicable to a wide range of studies in infection, vaccination, auto-immunity, and cancer.

Funding:

Title: BRILIA: sequence annotation and lineage reconstruction tool for B cell repertoire and vaccine studies

Authors: Donald W. Lee[1], Ilja Khavrutskii[1], Anders Wallqvist[1], Sina Bavari[2], Christopher L. Cooper[2], and Sidhartha Chaudhury[1*]
[1]Biotechnology HPC Software Applications Institute (BHSAI), Telemedicine and Advanced Technology Research Center, US Army Medical Research and Materiel Command, Fort Detrick, Frederick, MD, USA
[2]Molecular and Translational Sciences, US Army Medical Research Institute of Infectious Diseases, Fort Detrick, Frederick, MD, USA
*Corresponding author

Abstract: Sequencing a repertoire of B cells, particular their B cell receptors (BCRs) that are later secreted as antibodies, can reveal how vaccines affect humoral immunity. B cells produce unique BCRs through VDJ recombination and somatic hypermutation (SHM) during B cell development and affinity maturation. Tracking the evolution and selection of BCRs involves identifying mutations of the V, D, and J genes that encode the antigen-binding CDR3 region. However, most tools for VDJ annotation do not deploy repertoire information to improve SHM identification, and separate tools are needed for annotation and lineage reconstruction. Here, we present BRILIA, an algorithm that leverages repertoire-wide sequencing data to globally improve the VDJ annotation coverage, lineage tree assembly, and SHM identification. On benchmark tests against simulated human and mouse BCR sequences, BRILIA annotated germline (unmutated) and clonally related (mutated) sequences with 94% and 70% accuracy, respectively. When applied to B cell repertoires collected from C57BL/6 mice, BRILIA returned SHM patterns that are consistent across the VDJ junction and with known biological mechanisms of SHM. We are now using BRILIA to study the mouse B cell response to vaccines when adjuvants are used. Our initial analysis shows that adjuvants increase the chance that different mouse subjects will select for B cells with similar CDR3 regions, and these “convergent” B cell clones have highly complex lineage trees. Our results support the hypothesis that adjuvants enhance the affinity maturation of B cells.

Funding:

Title: Signatures of acute EBV infection in peripheral B and T cell compartments

Authors: Katherine JL Jackson[1,2], Shilpa A Joshi[2], Daniel Bernstein[3], Carlos O Esquivel[4], Sheri M Krams[4], Olivia Martinez[4], Kenneth I Weinberg[3], Scott D Boyd[2]
[1] Immunology Division, The Garvan Institute of Medical Research, Sydney, NSW, AU; Departments of [2] Pathology, [3] Pediatrics, [4] Surgery, Stanford University School of Medicine, Stanford, CA, USA

Abstract: EBV infection is common; more than 90% of adults are seropositive for the virus. In the context of organ transplantation, EBV-naïve transplant recipients are at increased risk of post-transplant complications in the form of EBV-driven post-transplantation lymphoproliferative disorder (PTLD). The B cell clonal dynamics during acute infections that lead to PTLD, compared to those that lead only to chronic asymptomatic infection in healthy subjects, are not well understood.

To study the B cell responses to acute EBV infection in healthy subjects, we analysed rearranged immunoglobulin heavy chain (IGH) gene transcripts in 5 young adults diagnosed with acute primary mononucleosis. All heavy chain isotypes and subclasses were amplified from cDNA derived from peripheral B cells isolated at the time of acute infection and longitudinally at 6 week intervals for up to 6 months post-infection. T cell receptor beta (TCRB) chains were also amplified from the same samples. Longitudinal immune receptor repertoires were deep sequenced using Illumina MiSeq 2x300 pair-end reads. Additionally, phage libraries of non-native paired single chain Fvs derived from the IGH at early, mid and late time points from each subject were panned against key EBV antigens.
IGH repertoires for each subject revealed signatures of EBV infection on peripheral B cells, with significant alteration of a subset of IGH expressed as IgD. IGH inferred as EBV-specific by phage display were used to annotate each subject's repertoire to distinguish between clonal expansions of B cells that express antibodies specific for EBV antigens, from those B cells of other specificities that may be expanded by EBV infection of the B cell itself. In contrast to the dynamic B cell responses over the time course, the TCRB repertoire was found to be relatively stable, and was found to include public clonotypes with inferred EBV-specificity.

Funding:

Title: Reconstructing antibody repertoires from error-prone immunosequencing reads

Authors: Alexander Shlemov*, Sergey Bankevich*, Andrey Bzikadze,
Maria A. Turchaninova †, Yana Safonova*,‡ and Pavel A. Pevzner*,§

*Center for Algorithmic Biotechnology, St. Petersburg University, Russia
†Institute of Bioorganic Chemistry, Russian Academy of Sciences, Moscow, Russia
‡Information Theory and Applications Center, University of California, San Diego, USA
§Department of Computer Science and Engineering, University of California, San Diego, USA

Abstract: Transforming error-prone immunosequencing datasets into antibody repertoires is a fundamental problem in immunogenomics, and a prerequisite for studies of immune responses. Although various repertoire reconstruction algorithms were released in the last three years, it remains unclear how to benchmark them and how to assess the accuracy of the reconstructed repertoires. We describe an accurate IGREC algorithm for constructing antibody repertoires from high-throughput immunosequencing datasets and a new framework for assessing the quality of reconstructed repertoires. Surprisingly, antibody repertoires constructed by IGREC from barcoded immunosequencing datasets in the blind mode (without using information about unique molecular identifiers) improved upon the repertoires constructed by the state-of-the-art tools that use barcoding. This finding suggests that IGREC may alleviate the need to generate repertoires using the barcoding technology (the workhorse of current immunogenomics efforts) because our computational approach to error correction of immunosequencing data is nearly as powerful as the experimental approach based on barcoding.

Funding:

Title: BALDR – A Computational Pipeline for Paired Heavy and Light Chain Immunoglobulin Reconstruction in Single-Cell RNA-Seq Data

Authors: Amit A. Upadhyay[1], Robert C. Kauffman[2], Amber N. Wolabaugh[1], Alice Cho[2], Nirav B. Patel[3], Samantha M. Reiss[4,5], Colin Havenar-Daughton[4,5], Reem A. Dawoud[1], Gregory K. Tharp[3], Ignacio Sanz[2,6], Shane Crotty[4,5,7], F. Eun-Hyung Lee[2,8], Jens Wrammert[2] and Steven E. Bosinger[1,3,9]
[1] Division of Microbiology & Immunology, Yerkes National Primate Research Center.
[2] Department of Pediatrics, School of Medicine, Emory University.
[3] Yerkes NHP Genomics Core Laboratory, Yerkes National Primate Research Center.
[4] Division of Vaccine Discovery, La Jolla Institute for Allergy and Immunology, La Jolla, CA, USA.
[5] Scripps Center for HIV/AIDS Vaccine Immunology and Immunogen Discovery (CHAVI-ID), La Jolla, CA, USA.
[6] Division of Rheumatology, School of Medicine, Emory University.
[7] Division of Infectious Diseases, Department of Medicine, University of California, San Diego, La Jolla, CA, USA.
[8] Divisions of Pulmonary, Allergy, & Critical Care Medicine, Emory University, Atlanta, GA
[9] Department of Pathology & Laboratory Medicine, School of Medicine, Emory University.

Abstract: B cells play a critical role in the immune response by producing antibodies (referred to the B cell receptor – BCR, or immunoglobulins - Ig) that bind to pathogens and foreign bodies. The remarkable diversity of antibodies is conferred by genomic DNA rearrangements of V, D and J gene-segments. Next generation sequencing technology has made it possible to characterize BCR sequence diversity of unpaired, single BCR chains, and had a significant impact on the fields of immunology, vaccine design and cancer diagnostics. Here, we describe a bioinformatics pipeline, (BALDR BCR Assignment ofLineage using De novo Reconstruction) capable of reconstructing the paired heavy and light chain gene sequences of BCR from Illumina single-cell RNA-seq (scRNA-Seq) data.

Funding:

Title: Y-Tools, a toolkit for analysis of adaptive immune repertoires using immunosequencing data

Author: Yana Safonova, postdoc, Center for Information Theory and Applications, Computer Science and Engineering Department, University of California at San Diego, USA

Abstract: Rapid development of sequencing technologies opened new avenues for analyzing immune system through deep interrogation of adaptive immune repertoire repertoires. Emergence of Rep-seq data promoted development of immunoinformatics methods. We present Y-Tools, a novel toolkit for construction and analysis of adaptive immune repertoires using Rep-seq data. It includes:

- **IgReC**, an algorithm for antibody repertoire construction and immunoproteogenomics analysis, and **IgQUAST**, a quality assessment method for repertoires constructed from Rep-seq data;
- **IgDiversityAnalyzer**, a tool for analysis of diversity properties of adaptive immune repertoires;
- **AntEvolo**, an algorithm for construction of clonal trees for antibody repertoires;
- **IgPopulation**, an algorithm for inference of novel V, D, J segments from Rep-seq data;
- **IgSimulator**, a versatile simulator of antibody repertoire.

Y-Tools are publicly available at the Github: http://yana-safonova.github.io/ig_repertoire_constructor/. Surprisingly, advanced computational analysis of Rep-seq data revealed many poorly understood properties of adaptive immune system. For example, using Rep-seq of naive B-cells, we detected many novel germline segments that are not presented in the IMGT database, but widely distributed across the human population. Understanding of germline segment diversity helped us to improve clonal reconstruction of antibody repertoires. Ability to recognize extensive variations of Ig loci allows us to detect unique somatic hypermutations and accurately reconstruct development of clonal lineages in antibody repertoires.

Funding:

Title: sciReptor "*T. rex*": A toolkit for analysis of single-cell level T cell receptor repertoires

Authors: Ilka Bartsch[1], Julia Ludwig[1], Hedda Wardemann[1] and Christian E. Busse[1]
[1] Division of B Cell Immunology, German Cancer Research Center, Heidelberg, Germany

Abstract: The sequencing of T cell receptor (TCR) transcripts from individual T cells yields essential information about TCR alpha:beta chain pairing, which is lost in conventional bulk sequencing experiments. In addition, single-cell techniques allow the assignment complementary data types (e.g. cell surface marker expression), which are important for biological interpretation. However, many of the currently available computational tools are not designed to handle single-cell data and do not provide integral solutions for linking of sequence information to other biological data types.

Here we introduce the substantially extended "*T. rex*" version of our previously published sciReptor toolkit, which features the capability to process and analyze single-cell TCR repertoire sequencing data. The software is compatible to the previously established relational database backend, which stores and links raw data and analysis results. sciReptor furthermore supports the attribution of additional data categories such as cell surface marker expression or immunological metadata. Additionally, it comprises a quality control module, customizable reference databases and a set of basic repertoire visualization tools.

In summary, sciReptor "*T. rex*" is a flexible framework for standardized sequence analysis of both TCR and immunoglobulin (Ig) repertoires on the single-cell level.

Funding:

Title: Per-sample immunoglobulin germline inference with partis

Authors: Duncan Ralph[1], Frederick A Matsen IV[1] [1]Fred Hutch

Abstract: The collection of immunoglobulin germline genes, which give rise to B cell receptors via recombination, is known to vary significantly across individuals. In humans, for example, while there are several hundred known V alleles, less than 100 are present in each individual. Furthermore, this set of known V alleles is both incomplete (particularly for non-European samples), and contains a significant number of spurious alleles. The resulting uncertainty as to which immunoglobulin alleles are present in any given sample results in inaccurate annotations, and in particular inaccurate inferred naive ancestors. Such incorrect naive sequences could lead to incorrect conclusions about the mature antibody's antecedents when they are synthesized and tested in the lab. In this paper we first show that the currently widespread practice of aligning each sequence to its closest match in the full IMGT set results in an extremely inaccurate per-sample germline set. We then describe a new method for inferring each individual's germline set from deep sequencing data, and make a detailed comparison to existing methods on a variety of simulated and real data samples. This new method is part of the public partis repository at http://github.com/psathyrella/partis, and is run by default before annotation and partitioning (it does not increase run times).

Funding: NIH

Title: Phylogenetic substitution models for antibody lineages

Author: Kenneth Hoehn, Postdoctoral Associate, Yale University, Department of Pathology

Abstract: Phylogenetic methods have shown promise in understanding the development of broadly neutralizing antibody lineages (bNAbs). However, the mutational process that generates these lineages – somatic hypermutation – is biased by hotspot motifs, which violates important assumptions in most phylogenetic substitution models. Here, we develop a modified GY94-type substitution model that partially accounts for this context-dependency while preserving independence of sites during calculation. This model shows a substantially better fit to three well-characterized bNAb lineages than the standard GY94 model. We also demonstrate how our model can be used to test hypotheses concerning the roles of different hotspot and coldspot motifs in the evolution of B-cell lineages. Further, we explore the consequences of the idea that the number of hotspot motifs – and perhaps the mutation rate in general – is expected to decay over time in individual bNAb lineages.

Funding:

Title: Updates on ImmuneDB: a system for the analysis and exploration of high-throughput adaptive immune receptor sequencing data

Authors: Aaron M. Rosenfeld[1], Wenzhao Meng[2], Bochao Zhang[1], Dora Chen[2], Eline T. Luning Prak[2], and Uri Hershberg[1,3]
[1]School of Biomedical Engineering, Science and Health Systems, Drexel University, Philadelphia, PA, USA
[2]Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA
[3]Department of Microbiology and Immunology, Drexel College of Medicine, Drexel University, Philadelphia, PA, USA

Abstract:
As high-throughput sequencing of B cells becomes more common, the need for tools to analyze the large quantity of data also increases. We here present an updated version of ImmuneDB, a system for analyzing and visualizing heavy and light chain variable region sequences from high-throughput sequencing experiments. It can take as input raw FASTQ or processed FASTA files, identify genes, determine clones, construct clonal lineages, as well as provide information such as selection pressure and mutation analysis. It uses an industry-leading database, MySQL, to provide rapid analysis and avoid the complexities of using error-prone flat files.

Since its introduction in 2016, a number of new features have been added to ImmuneDB. The process for looking at non-human datasets and T cells has been streamlined. Sequences can be locally aligned to identify insertions or deletions, and can be assigned as a "sub-clone" of a larger clone. In addition to its existing similarity method, clones can now also be assigned by constructing lineages of all sequences and then separating branches based on common mutations. Further, users can import clones rather than use an ImmuneDB method. For interoperability with other AIRR tools, sequences and clones can also now be exported in VDJtools, pRESTO, and GenBank formats.

Funding: NIH P01 AI106697, NIH P30 CA016520

Title: CEDAR Technologies for AIRR Submissions

Authors: John Graybeal [1], Martin J. O'Connor[1], Syed Ahmad Chan Bukhari[2], Marcos Martínez-Romero[1], Attila L. Egyedi[1], Debra Willrett[1], Steven H. Kleinstein [2], Kei-Hoi Cheung[3,4], Mark A. Musen[1]

[1] Stanford Center for Biomedical Informatics Research, Stanford University, Stanford, CA, USA,
[2] Department of Pathology, Yale School of Medicine, New Haven, CT, USA,
[3] Department of Emergency Medicine, [4] Yale Center for Medical Informatics, Yale University School of Medicine, New Haven, CT, USA.

Abstract: The Center for Expanded Data Annotation and Retrieval (CEDAR) has released the CEDAR Workbench, a set of web services enabling easy model creation and submission of biomedical metadata. The software provides user interfaces and REST APIs to create, share, and manage high-quality metadata resources, and to submit those resources to selected repositories. CEDAR is partnering with members of the AIRR community from Yale University to develop an AIRR-specific set of templates and submission capabilities, and will be conducting formal evaluations of this submission pipeline in early 2018. The system's semantic and metadata capabilities include several features that will be particularly applicable to AIRR submissions, as well as many general user capabilities that are optimized for any domain.

The CEDAR Workbench provides a versatile environment for authoring metadata that are coupled with terms from ontologies, obtained from Stanford's BioPortal ontology repository. Template authors can precisely define the fields in the template, controlling how they appear and how they can be filled out by metadata providers. Metadata entries can be updated over time until they are ready for submission. Once the AIRR templates and features are validated and approved, the metadata can be submitted with the corresponding data to the NCBI SRA repository. The resulting metadata are also available via the API as JSON, JSON-LD, or RDF for easy integration in scientific applications and reusability on the Web. In addition to filling out and managing forms via the user interface, users can also leverage our APIs directly to validate metadata, share them, and submit them to external repositories.

The CEDAR Workbench is freely available to use, and the software is available as open source on GitHub in the *metadatacenter* project. Other information on CEDAR can be found at its web site, https://metadatacenter.org.

Funding:

Title: CAIRR: A pipeline to submit AIRR-seq data to the NCBI through the CEDAR-workbench

Authors: Syed Ahmad Chan Bukhari Martin J. O'Connor , Marcos Martínez-Romero , Attila L. Egyedi[2] , Debra Willrett[2] , John Graybeal [2] , Mark A. Musen[2] , Florian Rubelt[3] , Steven H. Kleinstein [1] , Kei-Hoi Cheung [4,5]

[1] Department of Pathology, Yale School of Medicine, New Haven, CT, USA,

[2] Stanford Center for Biomedical Informatics Research, Stanford University, Stanford, CA, USA,

[3] Department of Microbiology and Immunology and Institute for Immunity, Transplantation and Infection, Stanford University School of Medicine, Stanford, CA, USA,

[4] Department of Emergency Medicine,

[5] Yale Center for Medical Informatics, Yale University School of Medicine, New Haven, CT, USA.

Abstract: As part of the AIRR (Adaptive Immune Receptor Repertoire) Community Minimal Standards Working Group (http://airr-community.org/working_groups/minimal_standards), we have developed an initial set of metadata standards (https://github.com/airr-community/airr-standards) for publishing AIRR sequencing studies and are incorporating these recommendations into the National Center for Biotechnology Information (NCBI) repositories. Submissions of AIRR-seq data to the NCBI repositories typically use a combination of web-based and flat-file templates and include only a minimal amount of term validation. As a result, metadata at the NCBI are often described using inconsistent terminologies, limiting scientists' ability to access, find, interoperate, and reuse the data sets and to understand how the experiments were performed. The Center for Expanded Data Annotation and Retrieval (CEDAR, https://metadatacenter.org/) develops technologies involving the use of data standards and ontologies to improve metadata quality. In order to improve metadata quality and ease the AIRR-Seq study submission process, we have developed an AIRR-seq data submission pipeline named CEDAR-AIRR (CAIRR). CAIRR leverages CEDAR's technologies to i) create web-based templates whose entries are controlled by ontology terms, and ii) submit the ontology-linked metadata and sequence files (FASTQ) to the NCBI BioProject, BioSample, and Sequence Read Archive (SRA) databases. CAIRR provides a web-based metadata submission interface that is compliant with AIRR Community Minimal Standards Working Group recommendations. The interface enables ontology-based validation for several data elements, including: organism, disease, cell type and subtype, and tissue. This pipeline will facilitate the NCBI submission process and improve the metadata quality of AIRR-seq studies.

Funding: CEDAR is supported by the National Institutes of Health through an NIH Big Data to Knowledge program under grant U54AI117925.

Title: The Immcantation Framework for analysis of AIRR-seq data: Containerization and AIRR data standards.

Authors: Jason A. Vander Heiden[1], Namita Gupta[2], Gur Yaari[3], Mohamed Uduman[4], Daniel Gadala-Maria[2], Susanna Marquez[4], Julian Zhou[2], Nima Nouri[4], Ruoyi Jiang[5], Steven H. Kleinstein[2,4,5].

1. Department of Neurology, Yale School of Medicine.
2. Interdepartmental Program in Computational Biology & Bioinformatics, Yale University.
3. Bioengineering program, Bar-Ilan University.
4. Department of Pathology, Yale School of Medicine.
5. Department of Immunobiology, Yale School of Medicine.

Abstract: The field of high-throughput adaptive immune receptor repertoire sequencing (AIRR-seq) has experienced significant growth in recent years, but this growth has come with considerable complexity and variety in experimental design. These complexities, combined with the high germline and somatic diversity of immunoglobulin repertoires, present analytical challenges requiring specialized methodologies. We have developed the Immcantation Framework for AIRR-seq data analysis which covers a broad range of quality control, read processing, and analysis tasks.

We have recently updated the Immcantation Framework to support MiAIRR, the new AIRR Community data standards, added tools to facilitate MiAIRR-compliant GenBank submissions, and included parsers to transform IMGT/HighV-QUEST and IgBLAST output into an AIRR complaint format.

To facilitate use on computing clusters and promote reproducibility, we have encapsulated the entire Immcantation Framework, along with a set of accessory scripts, germline databases, and applicable third party tools into a set of meta-versioned Docker containers that are fully compatible with Singularity. Standalone software and detailed documentation are available online http://immcantation.readthedocs.io. Containers are available through docker hub at https://hub.docker.com/r/kleinstein/immcantation.

Funding: National Library of Medicine grant T15 LM07056; National Institutes of Health grant RO1AI104739; United States-Israel Binational Science Foundation grant 2009046.

Title: *iReceptor*: A Platform for Storing and Sharing AIRR-Seq Data

Authors: Felix Breden[1,2], Nishanth Marthandan[1,3], Bojan Zimonja[1], Jerome Jaglale[1], Nicole Knoetze[1], Emily Barr[1], Frances Breden[1], Yang Zhou[1], Richard Bruskiewich[1], Jamie K. Scott[1,3,4], and Brian Corrie[2]
[1]The IRAMCS Centre, [2]Dept. of Biological Sciences, [3]Dept. of Molecular Biology and Biochemistry, and [4]Faculty of Health Sciences, Simon Fraser University, Burnaby, British Columbia, V5A 1S6, Canada

Abstract: Recent advances in next generation sequencing (NGS) technologies have made it possible to sample the human antibody (Ab)/B-cell (BcR) and T-cell receptor (TcR) repertoires in tremendous detail. A typical human has ~10^9 B cells at any given time, with a very large percentage of these cells changing very rapidly over time. It is this rapid adaptation of our immune system over time that makes it so effective at fighting pathogens. NGS technologies allow investigators to sequence the immune receptor genes in essentially all the B and T cells in an individual, allowing researchers to explore how the immune system adapts. Adaptive Immune Receptor Repertoire sequencing (AIRR-Seq) has enormous promise for understanding the immune repertoire dynamics in vaccinology, infectious disease, autoimmunity, and cancer biology.

The iReceptor system is a platform to integrate and analyze these immense data sets by combining: 1) an international network of AIRR-Seq data repositories; 2) the ability to federate AIRR-Seq data across these distributed repositories; 3) advanced analytical tools unique to AIRR-Seq data; and 4) a scientific gateway that hides the complexity of performing research queries and advanced analyses across these federated data. iReceptor enables data federation across many distributed data sets by defining a web based API that implements the emerging AIRR Community (airr-community.org) data standard. In this tools demonstration, we will demonstrate the iReceptor Scientific Gateway and its ability to federate and explore data across multiple, distributed, AIRR compliant repositories.

Funding: Current: Canada Foundation for Innovation – Challenge 1
Past: CANARIE Network Enabled Platforms Program

Title: Updates on ImmuneDB: a system for the analysis and exploration of high-throughput adaptive immune receptor sequencing data

Authors: Aaron M. Rosenfeld[1], Wenzhao Meng[2], Bochao Zhang[1], Dora Chen[2], Eline T. Luning Prak[2], and Uri Hershberg[1,3]
[1]School of Biomedical Engineering, Science and Health Systems, Drexel University, Philadelphia, PA, USA
[2]Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA
[3]Department of Microbiology and Immunology, Drexel College of Medicine, Drexel University, Philadelphia, PA, USA

Abstract: As high-throughput sequencing of B cells becomes more common, the need for tools to analyze the large quantity of data also increases. We here present an updated version of ImmuneDB, a system for analyzing and visualizing heavy and light chain variable region sequences from high-throughput sequencing experiments. It can take as input raw FASTQ or processed FASTA files, identify genes, determine clones, construct clonal lineages, as well as provide information such as selection pressure and mutation analysis. It uses an industry-leading database, MySQL, to provide rapid analysis and avoid the complexities of using error-prone flat files.

Since its introduction in 2016, a number of new features have been added to ImmuneDB. The process for looking at non-human datasets and T cells has been streamlined. Sequences can be locally aligned to identify insertions or deletions, and can be assigned as a "sub-clone" of a larger clone. In addition to its existing similarity method, clones can now also be assigned by constructing lineages of all sequences and then separating branches based on common mutations. Further, users can import clones rather than use an ImmuneDB method. For interoperability with other AIRR tools, sequences and clones can also now be exported in VDJtools, pRESTO, and GenBank formats.

Funding: NIH P01 AI106697, NIH P30 CA016520

Title: immuneACCESS™: A Public Repository for TCR and BCR Sequencing Data Using the immunoSEQ® Analyzer

Authors: Ian Kaplan[1], Jeff Darga[1], [1]Adaptive Biotechnologies

Abstract: Adaptive Biotechnologies launched the immuneACCESS™ web portal to help researchers view, analyze, and most importantly, share immune repertoire data. immuneACCESS[1] is the world's largest public data repository for T- and B-cell receptor sequences; this ever-growing collection of immunoSEQ Data includes over 5,500 human samples and more than 460 mouse samples, that combine to represent 545,213,786 nucleotide sequences. These datasets come from research projects in basic immunology, cancer immunotherapy, autoimmune disease, infectious disease, and more. One of the key datasets is a cohort of 666 healthy adults with known cytomegalovirus serostatus[2]; and one of the most unique is the 37 million unique B-cell receptor sequences from three healthy adult donors—which is many fold deeper than any existing resource[3].

These freely available datasets are accessible in the immunoSEQ Analyzer as a software service designed specifically to support the analysis needs of researchers working with large immunosequencing datasets. immuneACCESS Data can be used to create new projects within the Analyzer, or exported for further analysis on other software platforms. The goal of immuneACCESS is to connect researchers throughout the immunosequencing community and accelerate discoveries by providing a platform to simplify data sharing and analysis.

References

1. www.immuneaccess.com
2. https://clients.adaptivebiotech.com/pub/emerson-2017-natgen
3. https://clients.adaptivebiotech.com/pub/robins-bcell-2016

Funding: N/A

Title: Querying for Epitope Specific Immune Receptors in the Immune Epitope Database (IEDB)

Authors: Randi Vita[1], Swapnil Mahajan[1], Deborah Shackelford[1], Jerome Lane[1], Veronique Schulten[1], Alessandro Sette[1], Bjoern Peters[1]
[1]La Jolla Institute for Allergy and Immunology, La Jolla, California, USA

Abstract: The Immune Epitope Database (IEDB) contains over 385,000 peptidic and non peptidic epitopes curated from over 18,000 journal articles, and provides a targeted query interface to retrieve epitopes based on their defining characteristics. These characteristics include the molecular composition of the epitope, such as peptide sequence or chemical structure, the organism source which naturally contains the epitope, such as viral, bacterial or eukaryotic species, or the type of immune response mounted against the epitope, such as antibodies capable of neutralizing infectious agents or T cells producing specific cytokines. We here present an update to the IEDB query and reporting infrastructure that 1) extends the outputs generated in response to an IEDB query to include the sequence and if available the 3 dimensional structure of specific T cell (TCR) and/or B cell (BCR) immune receptors known to recognize the retrieved epitopes and 2) expands the query interface to facilitate a search for particular characteristics of immune receptors, such as TCRs with specific CDR3 beta chain sequences or BCRs with known 3D structures. These query results can be downloaded as a comma separated file. As the IEDB continues to capture epitopes reported in the literature and with increasing availability of information on the immune receptors recognizing them, this new interface will grow in power to advance the understanding of epitope:receptor interactions

Funding: NIH Contract no.HHSN27220120001

Title: Partis: B cell receptor annotation, partitioning, simulation, and per-sample germline inference

Authors: Duncan Ralph[1], Frederick A Matsen IV[1]
[1]Fred Hutchinson Cancer Research Center

Abstract: partis is a tool for accurate, efficient annotation and clonal family inference on deep sequencing B cell receptor data. It also contains a detailed simulation engine. We have recently incorporated a new per-sample germline inference method (see poster). 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 https://github.com/psathyrella/partis/blob/master/manual.md), or at least get as far as you can in the installation.

Funding: NIH

Title: IgReC: a toolkit for construction and analysis of adaptive immune repertoire using Rep-seq data

Authors: Yana Safonova[1,2], Alexander Shlemov[2], Sergey Bankevich[2], Andrey Bzikadze[2,3], Pavel A. Pevzner[2,4]
[1]Center for Information Theory and Applications, University of California, San Diego, USA;
[2]Center for Algorithmic Biotechnology, Saint Petersburg State University, Russia;
[3]Bioinformatic and Systems Biology Department, University of California, San Diego, USA;
[4]Computer Science and Engineering Department, University of California, San Diego, USA;

Abstract: Transforming error-prone immunosequencing datasets into antibody or TCR repertoires is a fundamental problem in immunogenomics, and a prerequisite for studies of immune responses. Although various repertoire reconstruction algorithms were released in the last 3 years, it remains unclear how to benchmark them and how to assess the accuracy of the reconstructed repertoires. We present an accurate IgReC algorithm for constructing AIRR from high-throughput immunosequencing datasets and a new framework for assessing the quality of reconstructed repertoires. Surprisingly, repertoires constructed by IgReC from barcoded immunosequencing datasets in the blind mode (without using information about unique molecular identifiers) improved upon the repertoires constructed by the state-of-the-art tools that use barcoding. This finding suggests that IgReC may alleviate the need to generate repertoires using the barcoding technology (the workhorse of current immunogenomics efforts) because our computational approach to error correction of immunosequencing data is nearly as powerful as the experimental approach based on barcoding. IgReC package also includes barcodedIgReC, a repertoire construction tool for barcoded Rep-seq data; IgQUAST, a quality assessment tool for AIRR; IgDiversityAnalyzer, a tool for diversity analysis of antibody repertoires. IgReC package is publicly available at http://yana-safonova.github.io/ig_repertoire_constructor/.

Funding: This work was supported by the Russian Science Foundation (Grant 14-50-00069) and the National Institutes of Health (Grant 2-P41-GM103484).

Title: BRILIA demo: B cell sequence annotation and lineage reconstruction tool for repertoire studies

Authors: Donald Lee[1], Ilja Khavrutskii[1], Anders Wallqvist[1], Sina Bavari[2], Christopher Cooper[2], and Sidhartha Chaudhury[1*]
[1]Biotechnology HPC Software Applications Institute (BHSAI), Telemedicine and Advanced Technology Research Center, US Army Medical Research and Materiel Command, Fort Detrick, Frederick, MD, USA;
[2]Molecular and Translational Sciences, US Army Medical Research Institute of Infectious Diseases, Fort Detrick, Frederick, MD, USA

Abstract: We will show how and when to use BRILIA, which annotates VDJ and/or VJ junctions based on the inferred lineages of B cells from the same repertoire. BRILIA streamlines data analysis by taking sequence files from high-throughput sequencing, and returning annotation results, lineage trees, somatic hypermutation statistics, and high-resolution plots of various analyses. It runs in Linux and Windows, uses up to 12 cores in parallel, and comes in command-line and GUI versions. BRILIA is written in MATLAB and requires a one-time installation of the freely available MATLAB Runtime Library.

BRILIA: B-cell Repertoire Inductive Lineage and Immunosequence Annotator
Website: https://github.com/BHSAI/BRILIA (source codes + binary files)

Funding: Support for this research was provided by the Military Infectious Diseases Research Program of the United States (US) Army Medical Research and Materiel Command and the US Department of Defense (DoD) High-Performance Computing Modernization Program. This research was in-part funded by grants provided to USAMRIID by the US Department of Defense's Defense Threat Reduction Agency (DTRA).

**The opinions and assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the US Army or the US DoD.*

Title: Phylogenetic tree building and somatic hypermutation modeling with IgPhyML

Authors: Kenneth B Hoehn[1] Gerton Lunter[2], Oliver G Pybus[3]
[1]Department of Pathology, Yale University; [2]Wellcome Trust Centre for Human Genetics, University of Oxford; [3]Department of Zoology, University of Oxford;

Abstract: Estimating phylogenetic tree topology is an important part of many B cell receptor sequence analyses. Many programs have been previously developed to create phylogenetic trees in a variety of organisms using models of nucleotide substitution rates. These substitution models almost always assume that single nucleotide sites mutate independently of each other, and that mutation operates in a reversable manner (e.g. the mutation rate from A to T is the same as T to A). Unfortunately, somatic hypermutation (SHM) during B cell affinity maturation is biased by nucleotide context, which violates both of these assumptions. Recently, we developed a phylogenetic substitution model which relaxes these assumptions, and allows for phylogenetic topology estimation while simultaneously estimating rates of biased mutation in SHM hot- and coldspots. This model has been implemented in the program IgPhyML. Here, I demonstrate this core functionality.

Funding: This work was funded by the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement no. 614725-PATHPHYLODYN, and by Wellcome Trust grant 090532/Z/09/Z. KBH was also supported by a Marshall scholarship.

Title: Demonstrating Use of Recon for Obtaining Robust Estimates of Overall Immune-Repertoire Diversity from High-Throughput Measurements on Samples

Authors: Joseph Kaplinsky PhD[1], Rohit Arora PhD[2], and Ramy Arnaout, MD, DPhil[2].
[1]Technical University of Denmark Department of Micro- and Nanotechnology, Lyngby, Denmark.
[2]Department of Pathology and [3]Division of Clinical Informatics, Department of Medicine, Beth Israel Deaconess Medical Center, Boston, MA, USA.

Abstract: The diversity of an organism's B- and T-cell repertoires is both clinically important and a key measure of immunological complexity. However, diversity is hard to estimate by current methods, because of inherent uncertainty in the number of B- and T-cell clones that will be missing from a blood or tissue sample by chance (the missing-species problem), inevitable sampling bias, and experimental noise. We will present Recon, a modified maximum-likelihood method that outputs the overall diversity of a repertoire from measurements on a sample. Recon outputs accurate, robust estimates by any of a vast set of complementary diversity measures, including species richness and entropy, at fractional repertoire coverage. It also outputs error bars and power tables, allowing robust comparisons of diversity between individuals and over time. We apply Recon to *in silico* and experimental immune-repertoire sequencing data sets as proof of principle for measuring diversity in large, complex systems. Recon is freely available on Github for non-commercial use.

Funding: This work was supported by AHA 15GPSPG23830004 and NIH K08 AI11495801.

Title: sciReptor "*T. rex*": A toolkit for analysis of single-cell level T cell receptor repertoires

Authors: Ilka Bartsch[1], Julia Ludwig[1], Hedda Wardemann[1] and Christian E. Busse[1]

[1] Division of B Cell Immunology, German Cancer Research Center, Heidelberg, Germany

Abstract: The sequencing of T cell receptor (TCR) transcripts from individual T cells yields essential information about TCR alpha:beta chain pairing, which is lost in conventional bulk sequencing experiments. In addition, single-cell techniques allow the assignment complementary data types (e.g. cell surface marker expression), which are important for biological interpretation. However, many of the currently available computational tools are not designed to handle single-cell data and do not provide integral solutions for linking of sequence information to other biological data types.

Here we introduce the substantially extended "*T. rex*" version of our previously published sciReptor toolkit, which features the capability to process and analyze single-cell TCR repertoire sequencing data. The software is compatible to the previously established relational database backend, which stores and links raw data and analysis results. sciReptor furthermore supports the attribution of additional data categories such as cell surface marker expression or immunological metadata. Additionally, it comprises a quality control module, customizable reference databases and a set of basic repertoire visualization tools.

In summary, sciReptor "*T. rex*" is a flexible framework for standardized sequence analysis of both TCR and immunoglobulin (Ig) repertoires on the single-cell level.

Funding:

Title: Single cell transcriptomics for dissecting the tumor microenvironment

Authors: Paul Ryvkin, Zachary Bent, Stéphane Boutet, Sreenath Krishnan, Sofia Kyriazopoulou-Panagiotopoulou, Valeria Giangarra, Josephine Lee, Patrick Marks, Samuel Marrs, Elliott Meer, Jeff Mellen, Luz Montesclaros, Daniel Riordan, Joe Shuga, Jessica Terry, Kevin J. Wu, Grace X.Y. Zheng, Tarjei Mikkelsen and Deanna M. Church.

Abstract: The advent of high throughput, droplet based systems for assaying transcriptomes at single cell resolution has revolutionized our approach to studying complex biological systems. We recently described a fully-integrated, droplet based approach, the Chromium™ single cell system, that enables unbiased expression profiling of single cells coupled with single cell resolution of the immune receptor pairs for B cells and T cells, from a single workflow. High efficiency cell capture coupled with a low doublet rate (<1% per 1000 cells), facilitates the profiling of precious and rare cell populations.

Using this system we have profiled three distinct tumors and their microenvironments. We demonstrate the ability to distinguish tumor cells from surrounding cells, including immune cells. We are able to identify an expanded T cell clonotype in one tumor type and a rarer B cell clonotype expansion in a different tumor.

To support this work we provide an open source analysis pipeline, Cell Ranger, and a data visualization tool (Loupe™ Cell Browser)All of the scRNA-seq solutions described here include open source software packages (Cell Ranger™ and Loupe™ Cell Browser) for data analysis and visualization. New versions of Loupe Cell Browser and Loupe™ V(D)J Browser allow researchers to concurrently explore gene expression and the immune repertoire of the same cells, yielding new details about the tumor microenvironment.The Chromium™ Single Cell V(D)J Software Suite includes additions to the Cell Ranger pipeline, adding algorithms for *de novo* assembly of V(D)J sequences from paired-end UMI-tagged scRNA-seq data, annotation of V(D)J germline segments via Smith Waterman, and the Loupe™ V(D)J Browser for interactive investigation of clonotypes and read verification.

We believe that integrated droplet-based systems will enable widespread adoption of high throughput single cell mRNA analysis and accelerate the characterization of diverse developmental systems as well as tumor samples.

Funding:

Title: Profiling immunoglobulin repertoires across multiple human tissues by RNA Sequencing

Authors: Serghei Mangul*#[1,2], Igor Mandric*[3], Harry Taegyun Yang[1], Nicolas Strauli[4], Dennis Montoya[5], Jeremy Rotman[1], Will Van Der Wey[1], Jiem R. Ronas[6], Benjamin Statz[1], Douglas Yao[5], Alex Zelikovsky[3], Roberto Spreafico[2], Sagiv Shifman**[7], Noah Zaitlen**[8], Maura Rossetti**[9], K. Mark Ansel**[10], Eleazar Eskin**#[1,11]

[1]Department of Computer Science, University of California Los Angeles, Los Angeles, CA, USA
[2]Institute for Quantitative and Computational Biosciences, University of California Los Angeles, Los Angeles, CA, USA
[3]Department of Computer Science, Georgia State University, Atlanta, USA
[4]Biomedical Sciences Graduate Program, University of California, San Francisco, CA, USA
[5]Department of Molecular, Cell, and Developmental Biology, University of California, Los Angeles, Los Angeles, CA, USA
[6]Department of Microbiology, Immunology, and Molecular Genetics, University of California, Los Angeles, Los Angeles, CA 90095, USA
[7]Department of Genetics, The Institute of Life Sciences, The Hebrew University of Jerusalem, Jerusalem, Israel
[8]Department of Medicine, University of California, San Francisco, CA, USA
[9]Immunogenetics Center, Department of Pathology and Laboratory Medicine, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA, USA
[10]Department of Microbiology and Immunology, Sandler Asthma Basic Research Center, University of California, San Francisco, San Francisco, CA, USA
[11]Department of Human Genetics, University of California Los Angeles, Los Angeles, USA

Correspondence to: Serghei Mangul smangul@ucla.edu and Eleazar Eskin eeskin@cs.ucla.edu

*Equal contribution **Equal contribution

Abstract: Assay-based approaches provide a detailed view of the adaptive immune system by profiling immunoglobulin (Ig) receptor repertoires. However, these methods carry a high cost and lack the scale of standard RNA sequencing (RNA-Seq). Here we report the development of ImReP, a novel computational method for rapid and accurate profiling of the immunoglobulin repertoire from regular RNA-Seq data. ImReP can also accurately assemble the complementary determining regions 3 (CDR3s), the most variable regions of Ig receptors. We applied our novel method to 8,555 samples across 53 tissues from 544 individuals in the Genotype-Tissue Expression (GTEx v6) project. ImReP is able to efficiently extract Ig-derived reads from RNA-Seq data. Using ImReP, we have created a systematic atlas of 3.6 million Ig sequences across a broad range of tissue types, most of which have not been studied for Ig receptor repertoires. We also compared the GTEx tissues to track the flow of Ig clonotypes across immune-related tissues, including secondary lymphoid organs and organs encompassing mucosal, exocrine, and endocrine sites, and we examined the compositional similarities of clonal populations between these tissues. The Atlas of Immune repertoires (The AIR), is freely available at https://smangul1.github.io/TheAIR/ , is one of the largest collection of CDR3 sequences and tissue types. We anticipate this recourse will enhance future immunology studies and advance development of ther

apies for human diseases. ImReP is freely available at https://mandricigor.github.io/

Funding:

LAST NAME	FIRST NAME	COMPANY	EMAIL ADDRESS
Aniba	Mohamed	Amgen, Inc.	maniba@amgen.com
Arnaout	Ramy	BIDMC	rarnaout@gmail.com
Ashby	Meredith	Pacific Biosciences	mashby@pacificbiosciences.com
Atkins	Ngan	Atreca Inc.	nnguyen@atreca.com
Bagnara	Davide	University of Genova	davide.bagnara@edu.unige.it
Blanck	George	University of South Florida	gblanck@health.usf.edu
Bosinger	Steven	Emory University/Yerkes National Primate Research Center	steven.bosinger@emory.edu
Bostick	Magnolia	Takara Bio USA, Inc. (formerly Clontech Laboratories, Inc.)	magnolia_bostick@takarabio.com
Boyd	Scott	Stanford University	sboyd1@stanford.edu
Breden	Felix	Simon Fraser University	breden@sfu.ca
Briney	Bryan	The Scripps Research Institute	briney@scripps.edu
Bubela	Tania	Simon Fraser University	tbubela@sfu.ca
Bukhari	Syed Ahmad Chan	Yale University, School of Medicine	ahmad.chan@yale.edu
Bürckert	Jean-Philippe	Luxembourg Institute of Health	jean.buerckert@gmail.com
Busby	Ben	NCBI/NLM	busbybr@mail.nih.gov
Busse	Christian	German Cancer Research Center	christian.busse@dkfz-heidelberg.de
Chaudhury	Sidhartha	BHSAI/U.S. Army Medical Research and Materiel Command	sidhartha.chaudhury.civ@mail.mil
Chen	Noel	Novogene Corporation Inc	noel.chen@novogene.com
Chen	Dora	Perelman School of Medicine	dorayoungchen13@gmail.com
Chen	Wei	MacroGenics, Inc	chenw@macrogenics.com
Christley	Scott	UT Southwestern Medical Center	scott.christley@utsouthwestern.edu
Coelho	Camila	LMIV - NIAID/NIH	camila.coelho@nih.gov
Collins	Andrew	University of New South Wales	a.collins@unsw.edu.au
Corcoran	Martin	Karolinska Institute	martin.corcoran@ki.se
Corrie	Brian	iReceptor	bdcorrie@gmail.com
Cowell	Lindsay	UT Southwestern Medical Center	lindsay.cowell@utsouthwestern.edu
Crowe	James	Vanderbilt University Medical Center	james.crowe@vanderbilt.edu
Darga	Jeff	Adaptive Biotechnologies	jdarga@adaptivebiotech.com

Darko	*Sam*	*NIH/NIAID/VRC*	*sam.darko@nih.gov*
DeKosky	*Brandon*	*The University of Kansas*	*dekosky@ku.edu*
Douek	*Danny*	*NIH/NIAID/VRC*	*ddouek@mail.nih.gov*
Dubay	*Christopher*	*Providence Cancer Center*	*christopher.dubay@providence.org*
Durrett	*Russell*	*The University of Texas at Austin*	*russdurrett@utexas.edu*
Eugster	*Anne*	*CRTD TU Dresden*	*anne.eugster@tu-dresden.de*
Faison	*Jamie*	*Duke Human Vaccine Institute*	*jamie.faison@duke.edu*
Feeney	*Ann*	*The Scripps Research Institute*	*feeney@scripps.edu*
Feng	*Hanping*	*University of Maryland at Baltimore*	*hfeng@umaryland.edu*
Ferguson	*Stacy*	*NIAID/DAIT/BIB*	*fergusonst@niaid.nih.gov*
Fritz	*Brian*	*10x Genomics*	*brian.fritz@10xgenomics.com*
Frost	*Simon*	*University of Cambridge*	*sdf22@cam.ac.uk*
Fucile	*Christopher*	*University of Alabama at Birmingham*	*cfucile@uab.edu*
Gabriel	*Richard*	*DKFZ Heidelberg*	*richard.gabriel@nct-heidelberg.de*
Galson	*Jacob*	*Kymab Ltd*	*kerry.williamson@kymab.com*
Georgiou	*George*	*The University of Texas at Austin*	*gg@che.utexas.edu*
Geretz	*Aviva*	*US Military HIV Research Program/WRAIR*	*ageretz@hivresearch.org*
Gildersleeve	*Jeff*	*National Cancer Institute*	*gildersj@mail.nih.gov*
Graybeal	*John*	*Stanford University*	*jgraybeal@stanford.edu*
Greiff	*Victor*	*University of Oslo*	*victor.greiff@medisin.uio.no*
Guo	*Yicheng*	*Columbia university*	*yg2521@cumc.columbia.edu*
Hackett	*Charles*	*NIH*	*chackett@niaid.nih.gov*
Hershberg	*Uri*	*Drexel University*	*uh25@drexel.edu*
Hoehn	*Kenneth*	*Yale University*	*kenneth.hoehn@gmail.com*
Ippolito	*Gregory*	*The University of Texas at Austin*	*gregory.ippolito@gmail.com*
Jackson	*Katherine*	*Garvan Institute of Medical Research*	*k.jackson@garvan.org.au*
Kaplan	*Ian*	*Adaptive Biotechnologies*	*ikaplan@adaptivebiotech.com*
Karlsson	*Gunilla*	*Karolinska Institutet*	*Gunilla.Karlsson.Hedestam@ki.se*
Kelil	*Abdellali*	*University of Toronto*	*abdellali.kelil@utoronto.ca*
Kim	*Kangchon*	*Department of Ecology and Evolution,*	*kckim@uchicago.edu*
Klatzmann	*David*	*Sorbonne Université - UPMC*	*david.klatzmann@upmc.fr*
Kleinstein	*Steven*	*Yale School of Medicine*	*steven.kleinstein@yale.edu*

Kong	*Rui*	*NIH*	*rui.kong@nih.gov*
Lapham	*Cheryl*	*NIAID, NIH*	*clapham@niaid.nih.gov*
Laserson	*Uri*	*Icahn School of Medicine at Mount Sinai*	*uri@lasersonlab.org*
Lee	*Donald*	*BHSAI*	*dlee@bhsai.org*
Lees	*William*	*Birkbeck College, University of London*	*william@lees.org.uk*
Lefranc	*Marie-Paule*	*IMGT®, University of Montpellier, CNRS*	*MariePaule.Lefranc@igh.cnrs.fr*
Lindeman	*Ida*	*University of Oslo*	*il5@sanger.ac.uk*
Luning Prak	*Nina*	*University of Pennsylvania*	*luning@mail.med.upenn.edu*
Mangul	*Serghei*	*UCLA*	*smangul@ucla.edu*
Marasco	*Wayne*	*Dana-Farber Cancer Institute*	*wayne_marasco@dfci.harvard.edu*
Mariotti-	*Encarnita*	*Sorbonne Université - UPMC*	*encarnita.mariotti@upmc.fr*
Martins	*Andrew*	*National Institute of Allergy and Infectious*	*martinsaj@niaid.nih.gov*
Matsen	*Frederick ("Erick")*	*Fred Hutchinson Cancer Research Center*	*ematsen@gmail.com*
Mehr	*Ramit*	*Bar-Ilan University*	*ramit.mehr@biu.ac.il*
Meier	*Stephen*	*Stony Brook University*	*stephen.meier@stonybrook.edu*
Meng	*Wenzhao*	*University of Pennsylvania*	*wmeng@pennmedicine.upenn.edu*
Naito	*Ikue*	*Miraca Research Institute G.K.*	*ikue.ito@miraca.com*
Nouri	*Nima*	*Yale University School of Medicine*	*Nima.nouri@yale.edu*
Ohlin	*Mats*	*Lund University*	*mats.ohlin@immun.lth.se*
Olsen	*Christian*	*Pacific Biosciences*	*colsen@pacificbiosciences.com*
Olson	*Branden*	*Fred Hutchinson Cancer Research Center*	*branden.olson@gmail.com*
Olweus	*Johanna*	*Oslo University Hospital Radiumhospitalet*	*johanna.olweus@medisin.uio.no*
O'Malley	*John*	*Dana-Farber/Brigham and Women's*	*jtomalley@bwh.harvard.edu*
Or-Guil	*Michal*	*Humboldt University Berlin*	*michal.orguil@gmail.com*
Peters	*Bjoern*	*La Jolla Institute for Allergy and*	*bpeters@lji.org*
Petrova	*Velislava*	*Wellcome Trust Sanger Institute*	*vp5@sanger.ac.uk*
Pevzner	*Pavel*	*Computer Science Dept., UCSD*	*ppevzner@ucsd.edu*
Pierce	*Brian*	*University of Maryland*	*pierce@umd.edu*
Quach	*Tam*	*The Feinstein Institute for Medical*	*tquach@nshs.edu*
Ralph	*Duncan*	*Fred Hutch*	*dkralph@gmail.com*
Ramesh	*Akshaya*	*UCSF*	*Akshaya.Ramesh@ucsf.edu*
Reichert	*Janice*	*The Antibody Society*	*janice.reichert@antibodysociety.org*

Rele	*Shyam*	*NIH*	*shyam.rele@nih.gov*
Robins	*Harlan*	*Adaptive Biotechnologies*	*hrobins@adaptivebiotech.com*
Rosenberg	*Alexander*	*University of Alabama at Birmingham*	*afr@uab.edu*
Rosenfeld	*Aaron*	*Drexel University*	*aaron.rosenfeld@drexel.edu*
Rotrosen	*Daniel*	*NIH/NIAID/DAIT*	*drotrosen@niaid.nih.gov*
Rubelt	*Florian*	*Stanford University*	*frubelt@stanford.edu*
Ryder	*Dan*	*BISC Global*	*dan.ryder@biscglobal.com*
Safonova	*Yana*	*UCSD*	*isafonova@eng.ucsd.edu*
Sanz	*Ignacio*	*Emory School Of Medicine*	*isanz@emory.edu*
Scheepers	*Cathrine*	*National Institute for Communicable*	*cathrinem@nicd.ac.za*
Scott	*Jamie*	*Simon Fraser University*	*jkscott@sfu.ca*
Setliff	*Marion (Ian)*	*Vanderbilt University*	*ian.setliff@vanderbilt.edu*
Shepherd	*Adrian*	*Birkbeck, University of London*	*a.shepherd@mail.cryst.bbk.ac.uk*
Sherkow	*Jacob*	*New York Law School*	*jacob.sherkow@nyls.edu*
Singh	*Kumud*	*NIAID/NIH*	*kumud.singh@nih.gov*
Singh	*Anjali*	*NIAID/NIH*	*anjalisingh@niaid.nih.gov*
Singleton	*Kentner*	*NIH / NIAID*	*singletonkl@niaid.nih.gov*
Sollid	*Ludvig M*	*University of Oslo*	*l.m.sollid@medisin.uio.no*
Soto	*Cinque*	*Vanderbilt Vaccine Center*	*Cinque.Soto@Vanderbilt.Edu*
Sun	*Ying*	*University of Maryland, Baltimore*	*yisun@umaryland.edu*
Thomas	*Rasmi*	*US Military HIV Research Program,*	*rthomas@hivresearch.org*
Tipton	*Christopher*	*Emory University*	*chris.tipton@emory.edu*
Toby	*Inimary*	*UT Southwestern Medical Center*	*inimary.toby@utsouthwestern.edu*
Towlerton	*Andrea*	*Fred Hutchinson Cancer Research Center*	*atowlert@fhcrc.org*
Trück	*Johannes*	*University Children's Hospital Zurich*	*johannes.trueck@kispi.uzh.ch*
Tsai	*Shang-Pu*	*Dana Farber Cancer Institute*	*ShangPu_Tsai@dfci.harvard.edu*
Tsang	*John*	*Laboratory of Systems Biology and Center*	*john.tsang@nih.gov*
Upadhyay	*Amit*	*Emory University*	*aupadh4@emory.edu*
Van Schaik	*Barbera*	*Academic Medical Center*	*b.d.vanschaik@amc.uva.nl*
Vander Heiden	*Jason*	*Yale*	*jason.vanderheiden@yale.edu*
Varthakavi	*Vasundhara*	*NIAID*	*varthakaviv@niaid.nih.gov*
Vieira	*Marcos*	*University of Chicago*	*mvieira@uchicago.edu*

Vijh	*Sujata*	*NIAID/DAIDS*	*sujata.vijh@nih.gov*
Watson	*Corey*	*University of Louisville School of Medicine*	*corey.watson@louisville.edu*
Watson	*Simon*	*Kymab*	*simon.watson@kymab.com*
Wine	*Yariv*	*Tel Aviv University*	*yarivwine@tauex.tau.ac.il*
Ye	*Jian*	*National Institutes of Health*	*jianye@ncbi.nlm.nih.gov*
Yeh	*Johannes*	*CSHL*	*yehjth@gmail.com*
Yen	*Amanda*	*Y's Lab*	*kosmos.bio@gmail.com*
Yin	*Yiyuan*	*NGM Biopharmaceuticals*	*yyin@ngmbio.com*
Yu	*Hua*	*University of Maryland Baltimore*	*hyu@umaryland.edu*
Zhang	*Yongrong*	*University of Maryland Baltimore*	*yzhang1@umaryland.edu*
Zhou	*Quan (Julian)*	*Yale University*	*julian.zhou@yale.edu*

The AIRR Community gratefully acknowledges our sponsors

ANTIBODY SOCIETY

10x GENOMICS®

Adaptive biotechnologies®

Genentech
A Member of the Roche Group

GRIFOLS