Breakout Session Topic: Disease & Diversity: Autoimmune and infectious disease

Date/Time: Tuesday, 10 June, 4:25 - 5:25 SGT

Breakout Leaders: Ponpan Choopong and Jong-Eun Park

Scribes: TBC

Onsite Room: Auditorium

Zoom Room: Afternoon Webinar, live stream

Breakout Description

Autoimmune and infectious diseases are driven by complex interactions between host genetics, environmental exposures, and tissue-specific immune regulation. Critical cellular processes in these contexts include dysregulation of immune pathways, inflammatory signaling, cell-cell communication, and pathogen interactions with host cells. This breakout session will explore strategies to integrate diverse data types, spanning single-cell, spatial, and multi-omics technologies, to unravel the mechanisms driving disease heterogeneity. The session will emphasize how diversity, across populations, genetic backgrounds, and environments, shapes immune responses and disease progression.

Key Questions to discuss and explore

Disease of interest

What are the autoimmune and infectious diseases of interest in the HCA community, across regions?

How can we leverage single-cell and spatial technologies to map cellular and molecular heterogeneity in autoimmune and infectious diseases?

Disease mechanism and diversity

What are the key drivers of disease heterogeneity, and how do they differ between autoimmune and infectious contexts?

How should we integrate genetic, environmental (including pathogens and microbiome), and geographic diversity into study designs for autoimmune and infectious disease atlases?

How does HLA and immune repertoire diversity shape susceptibility or resistance and responses to vaccination and therapeutic modalities in autoimmune and infectious diseases?

Technology

What computational models or machine learning approaches can best capture repertoire diversity and its functional consequences?

What are the technical challenges in integrating immune repertoire data with other single-cell and spatial modalities in disease contexts?

Scribe notes and takeaways

Summary

This breakout session focussed on autoimmune and infectious diseases and included discussions on diseases of interest, tissue samples (collection, types, availability and quality), metadata, protocol harmonization and applying multiomics to study of diseases. Dengue was suggested as a focus of a cross-regional collaborative project because there are cohorts across Latin America and Asia. TB and post-TB Bronchitis, common autoimmune diseases (sclerosis, lupus. rheumatoid arthritis, psoriasis) and Covid-19 were also mentioned. There was interest in understanding why people respond differently to diseases, vaccines and treatments and interest in pathogen interactions. It was noted that environment/location, ethnicity and severity of disease would be useful metadata to capture. There was interest in integrating microbiome studies and engagement with immunologists, microbiologists and Bionetwork experts. Examples of relevant spatial transcriptomics work were shared.

Challenges identified included technical preparation of samples (not always possible in rural areas), methodological aspects in correcting batch effects, and low access to some samples. The importance of quality samples was raised, and whether PBMCs provide adequate disease information. Concern was raised over data privacy, but it was noted that cell atlases could be used to contextualize and project data onto allowing full control of sample data.

Actions

- Deciding which tissue/cells is important considering that the protocol and metadata will slightly differ depending on the chosen tissues/cells – No consensus on this.
- Deciding the diseases – No consensus on this. Suggestion to start creating flagship diseases atlas project for certain diseases and approach immune cell atlas for guidance.
- Harmonizing the protocol of data collection and data processing.
- The importance of collecting and standardized metadata. Working together with the diversity task force (for the diversity aspect) and immunologist or clinician (for the diseases of interest specific information).
- Engagement: socializing protocols and metadata standardization and harmonisation to people in the data collection site and clinicians to minimize the variability. This could include introducing single cell sequencing to clinicians who routinely collect samples.
- Inviting microbiologists or immunologists to join the flagship project.
- Strike balance between the effort in scaling up (size, diversity, etc) and the technical difficulties (practical)

Discussion

The audience was invited to share ideas about infectious disease and autoimmune diseases that they care about.

- Ponpan Choopong: Tropical infectious diseases, Dengue common along the equator, not just in Asia but also Africa and Latin America
    - Patricia Savarino: Cohorts in Latin America and Asia great opportunities to work across regions.
    - Jared de larosa Philippine:, TB - establishing first single facilities in Philippines and trying to introduce cohorts post-TB Bronchitis, working with blood, PBMCs
    - Ponpan Matangkasombut Choopong: Collaborators in Thailand also have TB
    - Covid-19 how it relates to other viral infections
- Ponpan Choopong: Different populations responded to the virus differently, still important to integrate global data to investigate why?
- Jong-Eun Park: Potential to focus, on bacteria,or viruses or parasites
- Ponpan Choopong: Anyone work with Vaccine responses?
    - Common auto-immune disease, rheumatoid arthritis - interested in things that respond together, or respond to similar treatments
    - Vaccines, systems biology approach, influenza vaccine
- Jong-Eun Park: Anyone interested in Genetic diversity in autoimmune diseases?
    - Sclerosis, more prevalent in some areas, unclear why.
- Jong-Eun Park: What metadata needs to be collected?
    - Want to capture genetic diversity, should be a component to consider, but also the depth of immunological profiles, scalable assays and then deep dive into deeper profiling targets to conditions of interest. Diseases of local importance should be the focus.

    - Problem of benchmarking genetic diversity of different locales - protocols that will allow us to benchmark data across data sets + capturing ethnicities in different environments.
- Ponpan Choopong: experience from AIDA: Harmonising protocol in sample collection of data collection is critical and single reference distributed to every country.
    - Need a set of metadata related to the disease of interest. Good idea to keep plasma, serum, extra tissue and blood for add-on study.
- Jong-Eun Park: Think about whether PBMC is the most representative sample for disease, freezing cell, causes loss in plasma e.g. should they rely on fresh samples? Thinking about the type of tissues, e.g. with blood what type of cells matter? Should be targeted.
    - Identifying niches of interest, leveraging single cell resolution for spatial technologies. Metadata as part of building immune cells (spent time thinking about Tier 2) but could be relevant across different atlases. Relevant for cross-tissue atlases to study immune responses.
- Ponpan Choopong: Inclusion and exclusion criteria and categorise timepoints of disease in a clinical setting
    - Autoimmune diseases, familial metadata should be recorded.
    - Ponpan Matangkasombut Choopong: Define what is healthy
    - What is truly healthy, the next core of HCA is clinically applied, how to include clinical information about a patient or donor, e.g. previous infections.
    - Quite an extensive metadata schema, but disease-wide targeted more towards covid, half samples come from Covid,
        - Metadata fields for Covid:
        - Days after infection was the sample taken
        - Multiple samples? If so, how far?
        - Estimation of disease severity
        - Diversity related factors e.g. sex, age different organs
    - With Covid your getting data from different centres, not practical to retrospectively ask for metadata, we can be ambitious but also practical and realistic about metadata that we can routinely collect from different centres
- Ponpan Choopong: Does anyone routinely have clinical labs in samples?
    - How important it is to have **quality samples,** can biobank yield good enough data for single cell sequencing?
    - Heart atlas, the biggest difficulty is that people have to rewrite metadata, someone with medical background knowledge to write down information. Tier 2 needs to be sent to people early on. It takes time.
    - Sample handling - collecting metadata not just for clinical relevance but technical preparation and sample handling (because it can have effects) Tier 1 applicable across all organs, methodological aspects - main challenge methods to remove batch effects, unclear if they are doing a good job. Disease: Important considering disease variation when building these atlases, in cases where we map diseases to the atlas

    - Problems with HCA data so far, won't be an issue for data collected for disease atlases, because metadata recorded by clinicians. So hopeful that this is unlikely to be an issue. Also, now a lot more awareness on diversity. The biggest issue is going to deciding which disease to focus on.
- Jong-Eun Park: We should design this prospectively rather than retrospectively. AIDA is a good example to just start something and collect PBMC in a uniform manner. Important to connect with experts working on existing atlas
    - Robust batch 2 single cell repertoire. Data availability not a lot of B cell sequenced human data, bone marrow data out there but few have B cells, low hanging fruit
- Jong-Eun Park: How to coordinate?
    - Before studying samples we need a shared protocol - Practical important step for cross-country collaboration
    - Having a baseline of lymphocytes would be another step, before talking about disease, having prior immune responses.
- Jong-Eun Park: What can be collected from samples? Good to have them all listed
    - Treatment paradigm? Could there be a focus on responses to treatment. Diseases: not only heterogeneous but also dynamic.
    - Jong-Eun Park: Longitudinal sampling - Can you see the Bioinformatic difficulties you foresee? How to apply spatial Omics? Inflammatory disease?
    - Cystic Fibrosis studies using Visium, challenge to bacteria samples low, Would the plan to be spatial transcriptomics 3D model showing pathogen interaction?
    - Lupus, plan to do spatial analysis on lupus patients, difficulties: kidney biopsies, capture can be low (you want early disease dynamics) early biopsies, and different sections of the organ, treatment agnostic lupus patients
- Jong-Eun Park: Difficult getting samples.
    - Work in pediatrics - autoimmune diseases are rare - what do people think about privacy, if that disease is rare, could they be identifiable?
    - Use cases of these kind of atlases, having representation of immune disease, one idea is if you have your own data, you should be able to use atlas to augment/contextualize your own data practical solution and means full control over own sample

From virtual audience:

- Metadata related to treatment received by the subjects/ tissue donors, will be very crucial in the interpretation and understanding of the scSeq data at the later stage, especially in translation for clinical application in future.Clinical scoring, lifestyle and other demographic data or important to be recorded.

- How about non-auto immune disease, but related to the Systemic low grade inflammation or disease related to it? Eg. Osteoarthritis?