**Sharing MS repository data with the IEDB**

As a possible first step towards linking and/or the establishment of interoperability of our database with SWATHAtlas, we propose an interim plan whereby the IEDB handles the uploading of relevant data identified from the SWATHAtlas web page. This proposal is aimed at achieving the most immediate functionality with the least amount of effort on the part of ISB, in Seattle.

At the bottom of the SWATHAtlas web page, for example, there is the "Submit" button. We propose that the addition of a check box for the user to click that would read something like, "**Click here if you are submitting MHC ligand data (HLA peptide libraries) and would like your data to be considered for submission to the Immune Epitope Database (IEDB) in addition to submission to SWATHAtlas.**"

This link could then send an email to our group and we can quickly review the data files, follow up if necessary with the authors and prepare the submission here in house. Because we already have a system in place to accommodate the direct submission of large datasets this may be the quickest way to start sharing data.

**Incorporating unpublished data not yet within repository from the Human Immunopeptidome Project (HIPP) group members**In order for MHC ligand data to be included in the IEDB, certain inclusion criteria must be met. In the interest of efficiency, I have included below an outline of fields captured by the IEDB for MHC ligand experiments [termed ‘MHC Ligand Elution (MHCLE)’ in our system]. This list includes required fields (in yellow), as well as others that are not required, but relevant to many MHC ligand studies. This is provided here to give the HIPP members an idea of the level of detail we need and/or can accommodate for consideration of their unpublished data. The programmatic scope of the IEDB as established by NIAID includes four main disease categories, infectious disease, allergy, autoimmunity, transplantation. For the purpose of MHC ligand data, we also include cancer. Going forward, we can make use of our Excel-type submission templates for ease of uploading these data. The current MHCLEs templates will accommodate the data field entries nicely (for in-house or future user-direct submissions).

**Tracking/Linking MHC ligand dataset submissions**

In order to provide searchable identifiers for each dataset (since not all will be associated with a PubMed ID) we can generate External Submission identifiers which could then be provided/housed on SWATHAtlas within the associated dataset file for *link-in* to the IEDB (fastest: create link to www.IEDB.org and copy paste ID into identifier search). This would require accommodation of this ID within each of the SWATHAtlas HLA libraries [perhaps next to the Publication annotation (‘IEDB Submission ID’)]. For any MHC ligand dataset submission incorporated into the IEDB that is associated with a curated published reference, the submission ID could be linked to the PubMed ID. Then within the IEDB on the Submission Reference Details page, a *link-out* to www.SWATHAtlas.org could be provided (either next to ‘Submission’ in the Reference Type field or perhaps in the Abstract free text field).

**Data Field Details for MHC ligand elution assays**

**Reference details**

Title, Author(s) name, abstract/description, author affiliation, PubMed ID (if applicable), date

**Epitope details**

Epitope Name (not required)

Epitope Linear Sequence (required)

Modified Residues (only if applicable)

Modification (only if applicable)

Source Antigen ID (required; GI, SwisProt, Uniprot)

Residues positions (if given; not required)

Epitope Source Organism Taxonomy ID (required; NCBI taxonomy)

**Assay Details**

Host Organism Tax ID (required; NCBI taxonomy. In the case of MHC ligand elution: organism from which the cell line was derived)

Host Organism Sex (not required)

Host Organism Age (not required)

Host Geolocation ID (only if known; not required)

*In Vivo* or *In vitro* Process (required)

This set of fields establishes whether the peptides are eluted from APCs in the context of 1) naturally-occurring disease (infection, allergy, AI), 2) in the context of an *in vivo* Administration (e.g. vaccine), 3) following an *in vitro* Ag stimulation, or 4) from established cell lines, with no in vivo or in vitro Ag.

Disease ID (only if known; not required)

Disease Stage (only if known; not required)

Immunogen (required if infectious disease, allergy or in vivo/in vitro administered antigen)

Immunogen Accession ID (required if infectious disease, allergy or in vivo/in vitro administered antigen)

Immunogen Taxonomy ID (required if infectious disease, allergy or in vivo/in vitro administered ag)

Adjuvants, Route, Dose, Schedule (only if applicable and/or known)

Immunization Comments (not required, but recommended for method details)

Assay Type (required) - Cellular MHC/mass spec, secreted MHC/mass spec, Edman degradation, etc.

Qualitative Measurement (required) – always positive in this case

Number of Subjects Tested/Responded (Response Frequency; only if known)

APC Tissue Type (required) – e.g. Blood, lymph node, spleen, etc.

APC Cell Type (required) – e.g. lymphocyte, splenocyte, dendritic, B cell, etc.

APC Culture Conditions (required) – e.g. cell lines, B-LCLs, direct ex vivo, etc.

MHC Allele/serotype/supertype (required) – e.g. B35, A\*01:01

Assay Comments (not required) – describes additional details of methods and findings.

Fields that MS experts may be interested in adding to address:

1. Isolation techniques
2. MS platforms
3. MS analysis tools/software