**Draft Project Proposal**

**Impact of HLA disparity between patient and donors on forming post-transplant antibodies.**

**Aims:**

1. Measure HLA disparity between patients and donors in terms of antigens, alleles, epitopes, amino acid and electrostatic mismatch (EMS) scores.
2. Measure correlations of these with post-transplant antibody formation.
   1. DSA to antigen / allele mismatch
   2. DSA to epitope mismatches
   3. ‘Third party’ antibody formation
3. Compare the different parameters to see which allows the best model of prediction.

**Methods:**

**Data sets.**

We have Gideon’s pregnancy dataset to work with.

We also have our dataset of transplanted patients along with post-transplant antibody data (n=599). We can select a subset from this. ( eg all first transplant candidates).

**Converting 1 field genotypes to 2 field haplotypes (for our 1 field dataset)**

For our dataset we will take the 1 field genotypes and perform maximum likelihood estimations of haplotypes using genomic analysis software. We will then use 2 field haplotype frequency data sets in the UK from [allelefrequencies.net](http://www.allelefrequencies.net) (AFND) to convert the 1 field haplotypes into the most likely 2 field haplotypes.

(I have an application that does this ;) )

To validate our approach we can compare the haplotype frequencies of our cohort to other populations in our region available in AFND to see how the frequencies of each haplotype compares.

We will further validate our approach by using the inferred 2 field haplotypes of recipients and their SAB data to see the frequency of allele-specific antibodies to ‘self’. (eg how many recipients inferred as A\*02:01 react to A\*02:01?).

**Generating HLA disparities**

Once we are satisfied with the converted 2 field haplotypes, we will convert HLA-types to epitopes and AA’s and generate mismatches and EMS scores for each recipient-donor pair. (I have an application that does this ;) )

**Results**

**Measuring correlations with post-transplant de-novo antibody formation.**

We will randomly split the data set into a test and validation set. A model of prediction will be built on the test set and then see how well our model predicts the outcome in the remaining cohort.

**What we will measure.**

1. The frequency of DSA to mismatched alleles and epitopes.
2. Correlations of each parameter (allele/epitope/AA/EMS) to DSA/antibody formation to inform thresholds that predict reactivity.

Find what proportions of recipient-donor pairs have low, medium, high HLA disparity (as measured by each of the parameters) in:

* Patients who remain unsensitised?
* Patients not previously sensitised who become sensitised?
* Patients sensitised before and after transplant, comparing those with static antibody profiles to those with *de-novo* antibodies?

**Outcomes**

* Estimate the immunogenicity of epitopes as well as their impact upon on a patient’s sensitisation profile from the epitope frequency itself.
* See which parameter correlates the best with DSA/antibody formation and what thresholds are significant.
* How do results compare to our validation subset and also Gideon’s results of the pregnancy dataset?