1. Genotyping was performed using the Illumina HumanCoreExome-24-v1-0 (**Batch 1**) or Illumina InfiniumCoreExome-24-v1-1 (**Batch 2**) according to the manufacturer’s SOP.
2. Raw idats from the Illumina iScan instrument were imported into GenomeStudio. Samples < 90 % call rate were excluded. Data was exported to PLINK PED/MAP format on the forward strand.
3. Data was converted from PED/MAP to BED/BIM/FAM using PLINK v1.07.
4. HumanCoreExome-24v1-0\_A\_PopulationReport\_MAF\_022015.txt or InfiniumCoreExome-24v1-1\_A\_PopulationReport\_MAF.txt was used to obtain a list of all variants on the array with a MAF > 0.005. This list was used to extract the variants from the genotype file.
5. Variants with a GenomeStudio Cluster Separation < 0.3 were excluded.
6. Variants with < 98% call rate excluded.
7. Individuals with < 98% call rate were excluded.
8. Palindromic SNPs (AT/CG) were excluded and the file split per chromosome.
9. Variants IDs were updated to match Haplotype Reference Consortium v1.1 using GenotypeHarmonizer
10. SNPs in common between the two arrays were extracted from each dataset and the files combined (**TACERA\_combined**).
11. W.Rayner’s script HRC-1000G-check-bim-4.23.pl was used to align SNPs to HRC v1.1 panel.
12. PLINK v1.9 was used to convert the PLINK files to VCF.
13. Sanger imputation server was used to impute the data to the HRC v1.1 panel using SHAPEIT2 for phasing and PBWT for imputation.

**HLA imputation**

1. See 1-7, above
2. PLINK was used to extract 28Mb-34Mb on chromosome 6.
3. SNP2HLA was used for imputation with the Type 1 Diabetes Genetics Consortium (T1DGC) Panel.
4. Custom bash and STATA scripts were used to extract the HLA haplotypes from the \*bgl.phased file.

Description of files:

/data/home/mpx225/tacera\_genotypes/snp2hla

SNP2HLA output files and HLA\_haplotypes.txt constructed from the bgl.phased file

/data/home/mpx225/tacera\_genotypes/imputed\_fromSanger

VCF files direct from Sanger imputation server + tabix index. One file per chromosome (22). Log files are also included.

/data/home/mpx225/tacera\_genotypes/genotypes\_before\_imputation

PLINK files of genotypes from batch1 and batch2 detailed above. Combined file also included. BED/BIM/FAM files for each dataset (total = 9)