

IHIW Registry Diversity Project

Carlheinz Müller, ZKRD

Martin Maiers, NMDP

Steven GE Marsh, Anthony Nolan

Goals

- Comparison, validation and improvement of **Bioinformatics** methods and tools for HLA haplotype frequency analysis
- Specifically designed to address issues of blood stem cell **registry** datasets:
 - Very large samples
 - Heterogeneous typing resolution and methods
 - Missing data for one or more loci
 - Ethnic diversity (self-identified and unidentified)
 - Population sub-structure determination (clustering)
 - Hardy-Weinberg deviation (non-random mating)

Products

- New methods and tools for analysis of registry populations
- Results from meta-analysis of Bone Marrow Donors Worldwide (>18M)

Previous Work

14th IHIWS - Melbourne

- Task 00: “Validation of software for haplotype frequency estimation on registry data”.
 - In this project, five donor registries agreed to compare their computer programs for haplotype frequency estimation (HFE) in order to validate them and to establish the degree of quality and consensus reached. Discrepancies between methods were informative and led to improvement but also highlighted the need for simulated data.



15th IHIWS - Buzios

- Task 01: “Accuracy of haplotype frequency estimations on 100 simulated data sets with complete typing”.
 - The goal of this task is to measure estimation error introduced by the EM algorithm and to consider it in isolation from sampling error.
- Task 02: “Accuracy of haplotype frequency estimations on simulated data sets with missing DR typing”.
 - Datasets with an increasing rate of missing HLA-DR typings were generated. The HLA-DR types in these sets were revealed in different proportions after undergoing simulated biased patient driven typing or random selection. The experience of the group shows that by refining the individual programs a broad consensus on the best practice will be achieved. In particular, we have developed recommendations for the implementation and application of the EM for HLA and related systems.

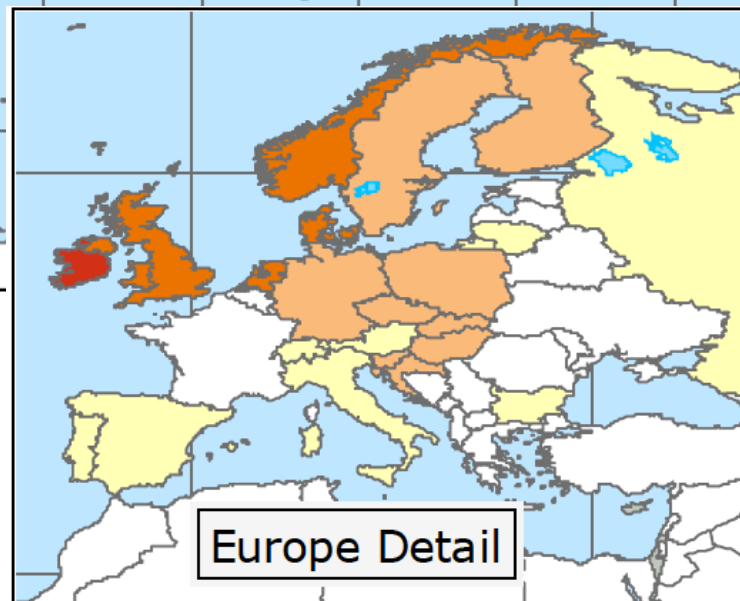
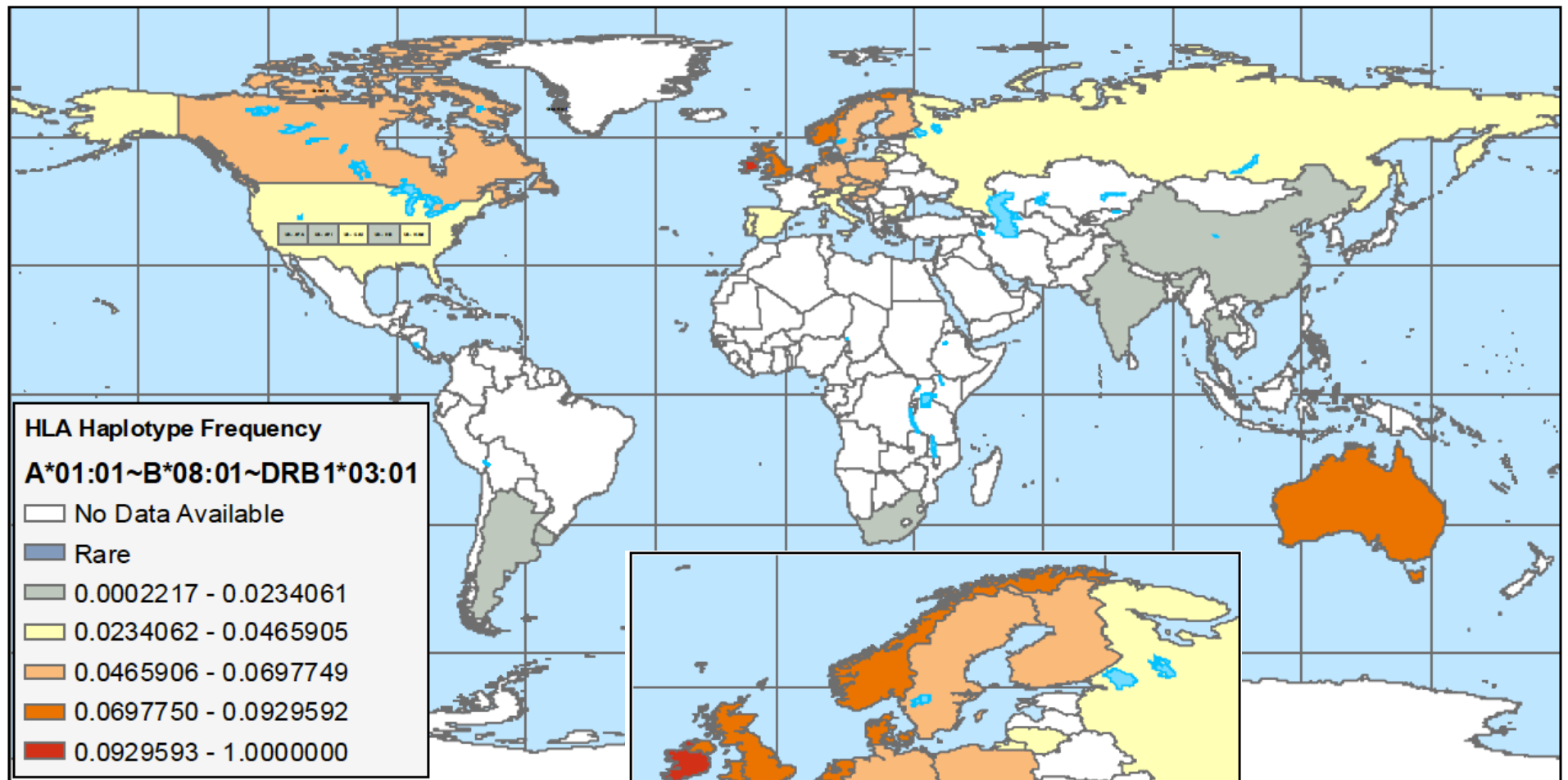


16th IHIWS – Liverpool

- **Task 03: “Accuracy of haplotype frequency estimations on simulated data sets with deviation from Hardy Weinberg”.**
 - Datasets will be distributed to participants and the results compared in order to validation implementations and confirm assumptions. One particular topic to address in this task is the accuracy of methods of identifying and even correcting for specific sources of HWE deviation such as cryptic population structure (the Wahlund effect).
- **Task 04: “Accuracy of methods of estimating high-resolution haplotypes from large datasets of mixed resolution”.**
 - All HLA typing methods include some level of ambiguity. DNA methods are typically reported in terms of allelic ambiguity but there are other forms of ambiguity: genotypic (inability to set phase), historical (typing to a fixed snapshot of a growing allele list) and ambiguity due to incomplete definition of reference alleles (missing sequence). Datasets will be distributed with a variety of forms and levels of ambiguity. Results will be compared between implementations and to the sample and source populations to quantify the error introduced by typing ambiguity.
- **Task 05: “Worldwide donor registry analysis: high-resolution HLA A-B-DRB1 haplotype frequencies for BMDW registries”**
 - Haplotype frequencies will be generated for all registries of the BMDW. This data will be applied to tools to project registry growth and diversity and the effects on matching.
 - Genetic distance analysis will be performed to determine the relatedness of donor registries.
 - Another product of this analysis will be a dictionary of haplotypes for low frequency or recently-described alleles for inclusion in the HLA DNA dictionary.
 - This data has direct practical clinical use for donor search and informing strategies for donor recruitment



Global HLA Haplotype Map



Participants

- Carlheinz Müller, Zentrales Knochenmarkspender-Register (ZKRD), Germany
- Hans-Peter Eberhard, Zentrales Knochenmarkspender-Register (ZKRD), Germany
- Martin Maiers, National Marrow Donor Program (NMDP), USA
- Loren Gragert, National Marrow Donor Program (NMDP), USA
- David Steiner, Czech Stem Cells Registry (CSCR), Czech Republic
- Steven GE Marsh, Anthony Nolan Research Institute (ANRI), United Kingdom
- Pierre-Antoine Gourraud, representing France Greffe de Moelle (FGM), France
- Marie-Lorraine Balere, France Greffe de Moelle (FGM), France
- Machteld Oudshoorn, Bone Marrow Donors Worldwide (BMDW) The Netherlands
- Henk van der Zanden, Bone Marrow Donors Worldwide (BMDW) The Netherlands
- Alexander Schmidt, Deutsche Knochenmarkspenderdatei (DKMS), Germany
- Paul Dunn, New Zealand Bone Marrow Donor Registry, New Zealand

