Department of Genetics, Cell- and Immunobiology
Semmelweis University, Faculty of Medicine

Haplotyping Hungarian paediatric ALL HLA-Class II genes using NGS

14th East-West Immunogenetics Conference, Budapest, 2020.03.07

Lívia Lázár, Dr. Nóra Kutszegi, Prof. Csaba Szalai, Dr. Dániel Erdélyi, Ágnes Félné Dr. Semsei, Prof. András Falus

Presented by Lívia Lázár, Omixon

Introduction – 1/3

- Ongoing effort between 1990 & 2012:
  - Collection of peripheral blood samples of patients with acute lymphoblastic leukemia (ALL)
  - A total of 576 specimen collected at 9 Hungarian paediatric hematology centers
  - 2017: first round of NGS analysis

- Hungarian population data on HLA:
  - No public information available(fig.1)
  - Neighboring populations show very similar motifs to our current study

Fig. 1: Central-European map of allelefrequencies.net displaying available population data on HLA loci (accessed: 2020.02.29)

Introduction – 2/3

- **Original goal of the precursor study:**
  - uncover allelic & haplotypic association of asparaginase hypersensitivity
  - 2 field resolution typing of HLA-DRB1 & HLA-DQB1
- **Published article:** HLA-DRB1*07:01–HLA-DQA1*02:01–HLADQB1*02:02 haplotype is associated with a high risk of asparaginase hypersensitivity in acute lymphoblastic leukemia
  (Kutszegi et. al., **Haematologica** 2017, Volume 102(9):1578-1586) – See Fig.2

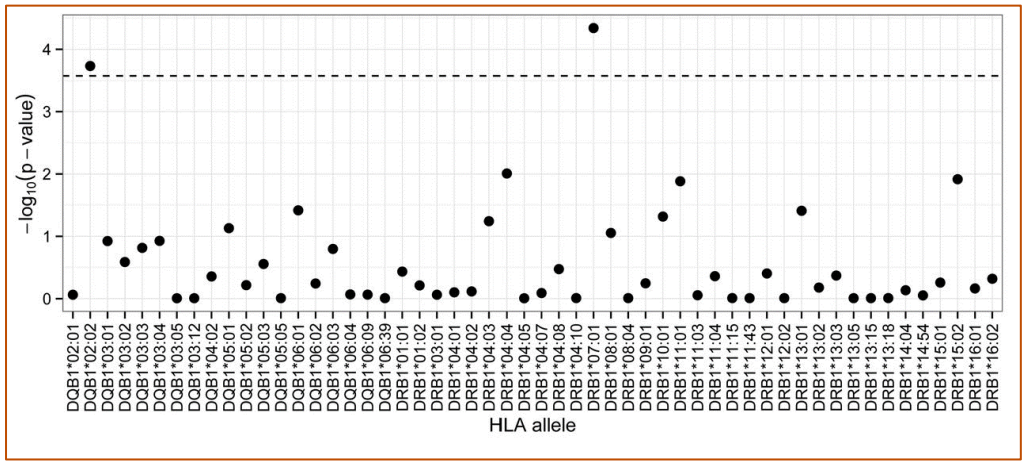


Fig. 2: Correlation coefficient calculated for loci of patients with asparaginase hypersensitivity on 2 fields. (Kutszegi et. al)

Introduction – 3/3

- **Changes in the current study:**
  - Targeted amplification of HLA-DQA1 & -DRB3, -DRB4 & -DRB5
  - 4 field resolution typing → Increased variability (Table 1)
  - Cohort size reduced (359 → 245) due to DNA quantity

Locus	Unique alleles on 2 fields	Unique alleles on 4 fields
HLA-DQA1	19	62
HLA-DQB1	21	37
HLA-DRB1	117	205

Table 1: Comparison of unique allele call counts on 2- and 4 fields

Methods & Materials -1/2

- **Amplification & sequencing:** Omixon Holotype™ HLA – Protocol v 3.0.1
  - Singleplex targeted amplification on HLA-DQA1, -DQB1, -DRB1, -DRB3, -DRB4 & -DRB5
  - 2x150 cycle sequencing on Illumina MiSeq
- **HLA typing**: Omixon HLA Twin™ v. 4.0.1 RUO
- **In-house tools**: extraction of GL strings, quality control tools
- **3rd party tools:**
  - https://www.allelefrequencies.net for haplotype accession (as of 2020.03.01)
  - GNU awk/bash for data formatting
  - Google Colab & Python 3.8 (with Pandas 1.0.1) – Haplotype matching

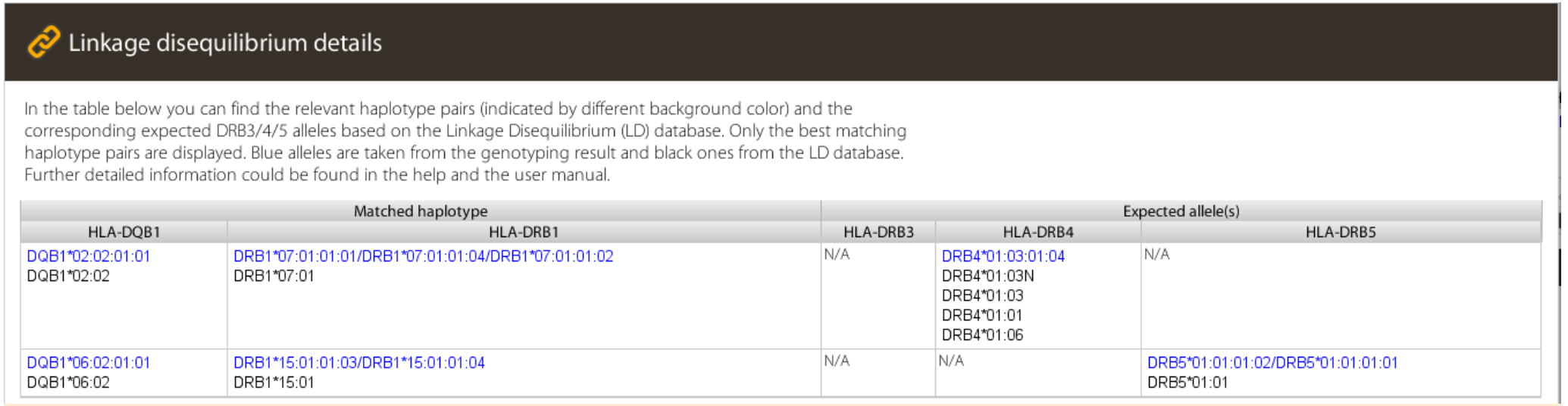
Methods & Materials – 2/2

- Extracted data from AFND (Table 2)

DQB1 & DRB1	54 149
DQA1 & DRB1	37
DQA1 & DRB1	581
DQA1, DQB1 & DRB1	6938

Table 2: Frequency of available loci for queried haplotypes

- Linkage Disequilibrium data from Twin (Fig. 3)



Linkage disequilibrium details

In the table below you can find the relevant haplotype pairs (indicated by different background color) and the corresponding expected DRB3/4/5 alleles based on the Linkage Disequilibrium (LD) database. Only the best matching haplotype pairs are displayed. Blue alleles are taken from the genotyping result and black ones from the LD database. Further detailed information could be found in the help and the user manual.

Matched haplotype		Expected allele(s)		
HLA-DQB1	HLA-DRB1	HLA-DRB3	HLA-DRB4	HLA-DRB5
DQB1*02:02:01:01 DQB1*02:02	DRB1*07:01:01:01/DRB1*07:01:01:04/DRB1*07:01:01:02 DRB1*07:01	N/A	DRB4*01:03:01:04 DRB4*01:03N DRB4*01:03 DRB4*01:01 DRB4*01:06	N/A
DQB1*06:02:01:01 DQB1*06:02	DRB1*15:01:01:03/DRB1*15:01:01:04 DRB1*15:01	N/A	N/A	DRB5*01:01:01:02/DRB5*01:01:01:01 DRB5*01:01

Fig. 3: Comparison of LD data (as described by Gragert et. al, Six-locus high resolution HLA haplotype frequencies derived from mixed-resolution DNA typing for the entire US donor registry. Human Immunology, 74(10), 1313–132)

Results & observations – 1/2

- All possible subsets of alleles were calculated on 2, 3 and 4 fields resulting in roughly 4000, 5000 and 8000 unique allele combinations respectively
- These were evaluated against known populations – previous estimations were excluded for the sake of “fair comparison”
  - The total set of available alleles and haplotypes matched up to over 800 000 hypothetical combinations
  - The most frequent haplotype proved to be the DQ5 on 3 fields (below) appeared in 261 populations and in 9 of the 245 samples, accompanied by DRB1*01:02:01

HLA-DQA1*01:01:02	HLA-DQB1*05:01:01	HLA-DRB1*01:02:01

Results & observations – 2/2

- In the examined population, the following set of alleles seemed to be the most abundant, making up 6.5% of the total allele calls

DQA1*01:01:01:01+DQA1*02:01:01:02	DQB1*02:02:01:01+DQB1*05:01:01:03	DRB1*01:01:01:01+DRB1*07:01:01:01| DRB1*01:01:01:01+DRB1*07:01:01:02| DRB1*01:01:01:01+DRB1*07:01:01:04| DRB1*01:01:01:02+DRB1*07:01:01:01| DRB1*01:01:01:02+DRB1*07:01:01:02| DRB1*01:01:01:02+DRB1*07:01:01:04| DRB1*01:100+DRB1*07:01:01:01| DRB1*01:100+DRB1*07:01:01:02| DRB1*01:100+DRB1*07:01:01:04	DRB4*01:03:01:04+DRB4*01:03:01:04

- DRB4 had some variation: 01:01:01:01 and 01:02:01:02N both appeared:
  - In combination with DQB1*02:02 and 03:03
  - Similarly as in
    The distribution of the DRB4*01:03:01:02N null allele in HLA-DRB1~DQB1 haplotypes in the Croatian population (Grubic Z. et al, **HLA,** 2018, Volume 91(1):23-28)

Next steps – 1/2

- With the inclusion of these results, further testing will be executed to measure allelic correlation of 3-4 field resolution data and asparaginase hypersensitivity, which means
  - Previously unavailable data from non-key exons
  - Possible correlation with intronic SNPs (e.g. Fig. 4)

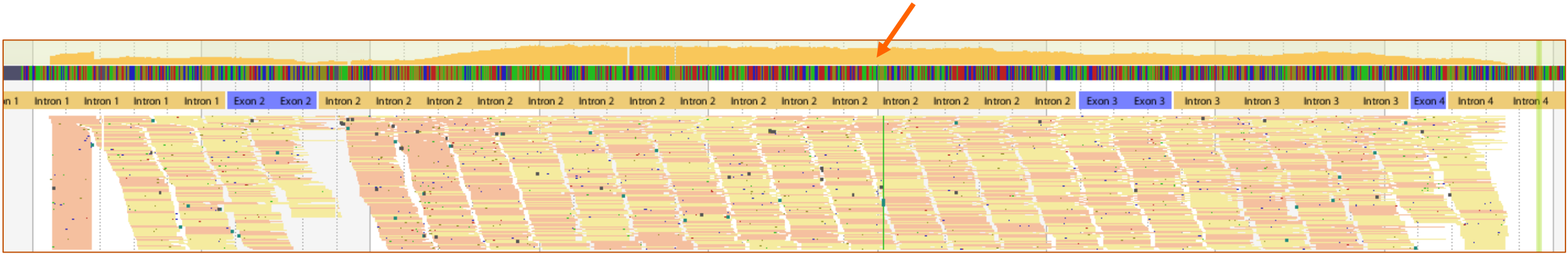


Fig. 4: Reference alignment of a best matching allele in comparison to the consensus with a well defined novel SNP in In2 on DRB1

Next steps – 1/2

- With the inclusion of these results, further testing will be executed to measure allelic correlation of 3-4 field resolution data and asparaginase hypersensitivity, which means
  - Previously unavailable data from non-key exons
  - Possible correlation with intronic SNPs

- Haplotype matching comparison
  - Previously HLA-DQA1 genotype was unavailable, was inferred based on the DQB1-DRB1 haplotype → actual result will be compared to estimation
  - Possible improvements of association precision can be expected

Acknowledgements

- **SOTE-GSI:**
  Dr. Nóra Kutszegi, Prof. Csaba Szalai, Dr. Dániel Erdélyi, Ágnes Félné Dr. Semsei, Prof. András Falus
- **Omixon:**
  Krisztina Rigó, Dr. Gergely Tölgyesi, Vivien Hársfalvi, Noémi Tóth
- Péter Horváth & Tünde Vágó

Thank you for your attention!

Any questions?

support@omixon.com