Belgian
Red Cross
Flanders

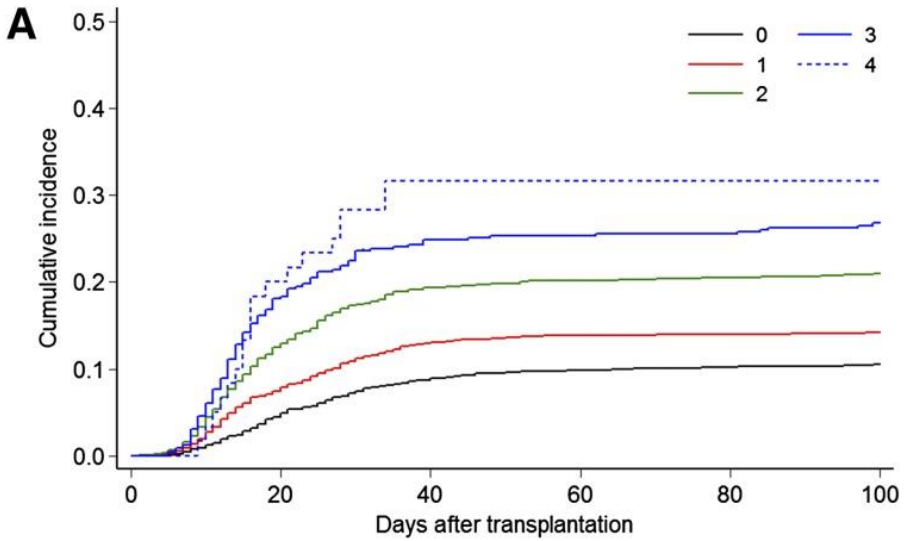
HLA and new technologies.

Vicky Van Sandt

Life-threatning malignant and non malignant blood disorders can be cured by hematopoetic stem cell transplantation (HSCT).

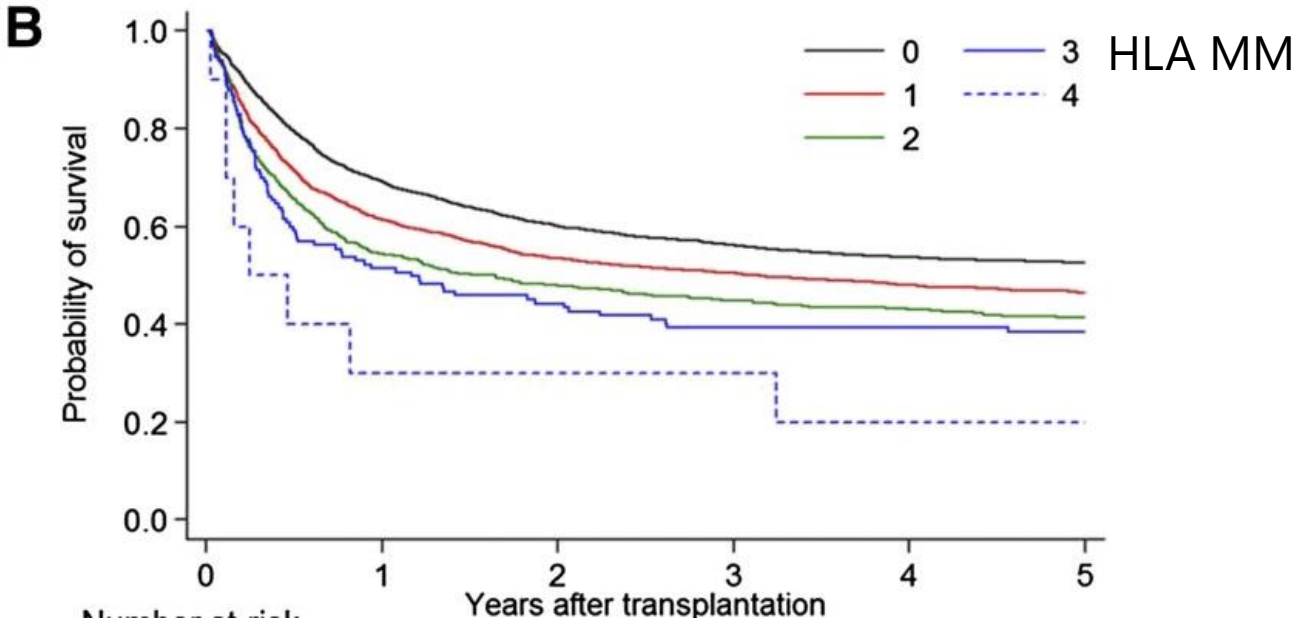
GVHD is the 2nd most prevalent cause of mortality after HSCT.

HLA mismatching is the strongest risk factor for GVHD.



Cumulative incidence of grade III-IV acute GVHD by the mismatch number of HLA-A, -B, -C, -DRB1_DQB1, and -DPB1 at the allele level in the GVH direction.

Morishima et al., 2015, Blood



Morishima et al., Blood, 2015

→ Overcome the HLA barrier

HLA

Where are we today? Registry - Laboratory

What does Next Generation Sequencing of HLA has to offer?

High resolution vs allelic resolution.

Future Perspectives

Human **L**eukocyte **A**ntigens

Recognize self from non self
Maximal response against foreign pathogens
Protect self
Survive

Transplantation

Mimic self: HLA matching or compatible donor
Avoid response against donor
Protect self and the non self donor
Survive

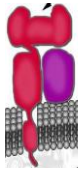
HLA is very diverse

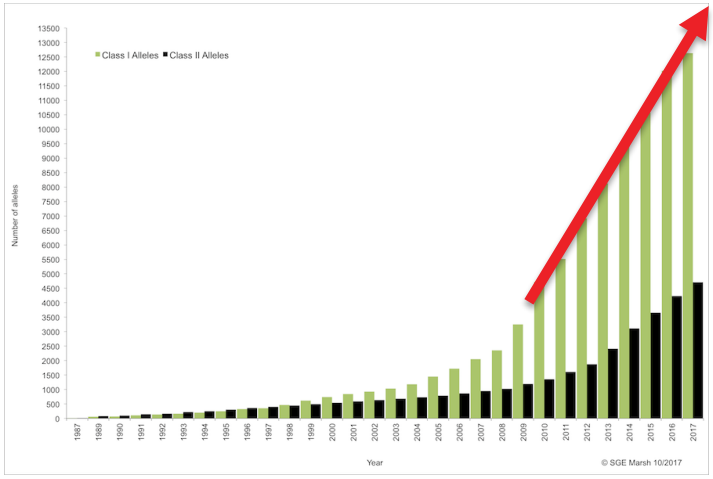
HLA in the human genome: IMGT HLA Database NOV 2017:

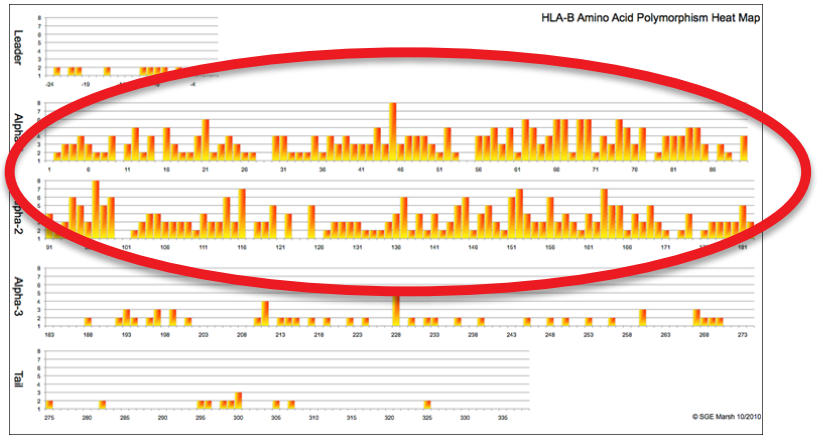
HLA Class I Alleles	12631
HLA Class II Alleles	4700
TOTAL	**17331**

DNA

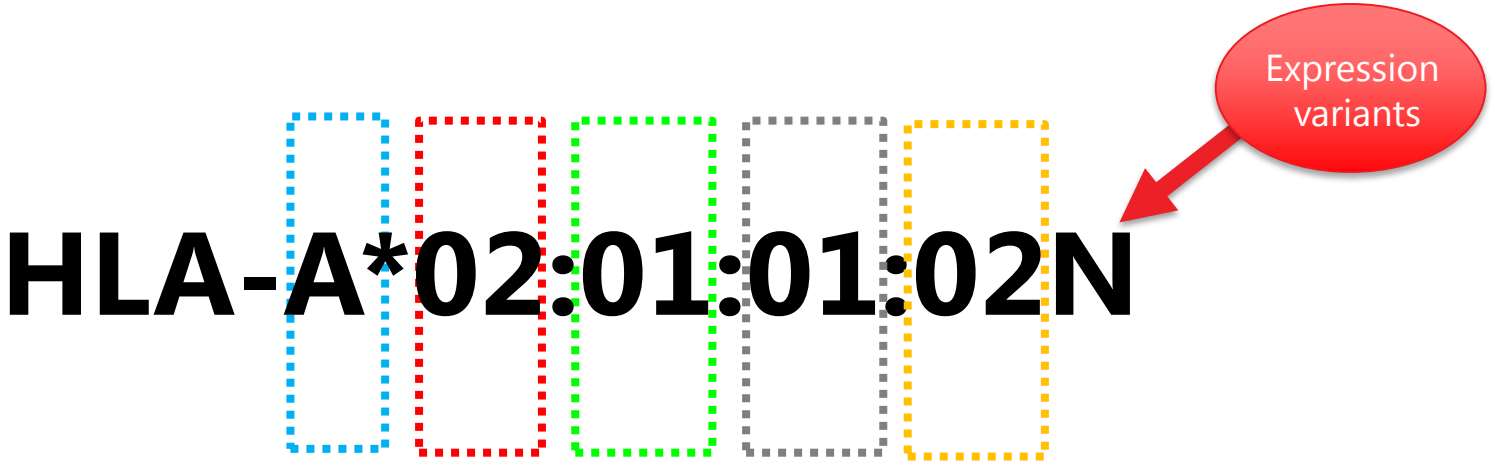
Antigen







HLA nomenclature



Locus

1st field: allelgroep, group of alleles with similar serological behaviour

2nd field, specification of a unique HLA protein → **High resolution:** identification of alleles that encode the same protein within the antigen binding site

3rd field synonym mutations

Allelic resolution: identification of a single allele

4th field variations in non coding regions.

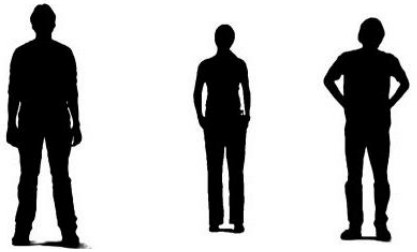
31 million chances to survive

You have no doubt already heard about stem cell donation. In the media, you can read about seriously ill people who are saved by stem-cell therapy or people who are urgently looking for a donor. The Belgian Bone Marrow Donor Registry's **31 million chances** campaign aims to better inform you about stem cell donation.

31 million chances represents the number of people registered in the international stem cell bank. **This means that sick patients have a staggering 31 million chances to survive.**

31 million chances is an awful lot. The Belgian Bone Marrow Donor Registry currently manages a database that stores the tissue types of people who are willing to donate stem cells. The registry also works very closely with foreign registries to give patients a better chance of finding a suitable donor. As such, we have access to stem cells from all over the world, significantly increasing the chances of finding a match. At the moment, some 31 million voluntary donors are registered worldwide.

DID YOU KNOW?

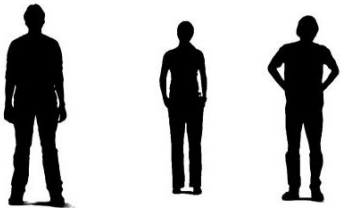


1 patient

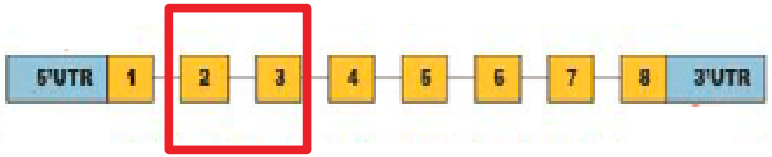
31 million potential donors

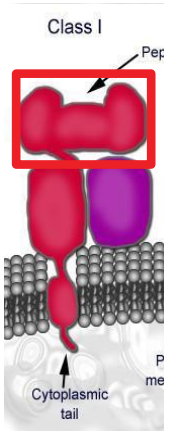
In most cases more than 1 HLA identical donor

Sanger sequencing

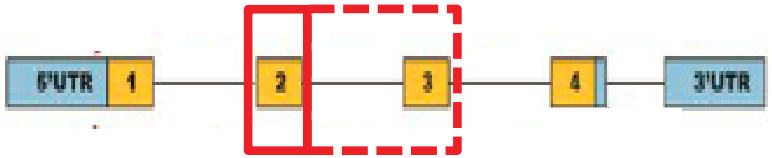


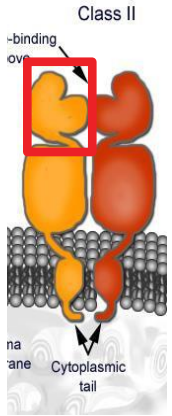
HLA-A, B, C





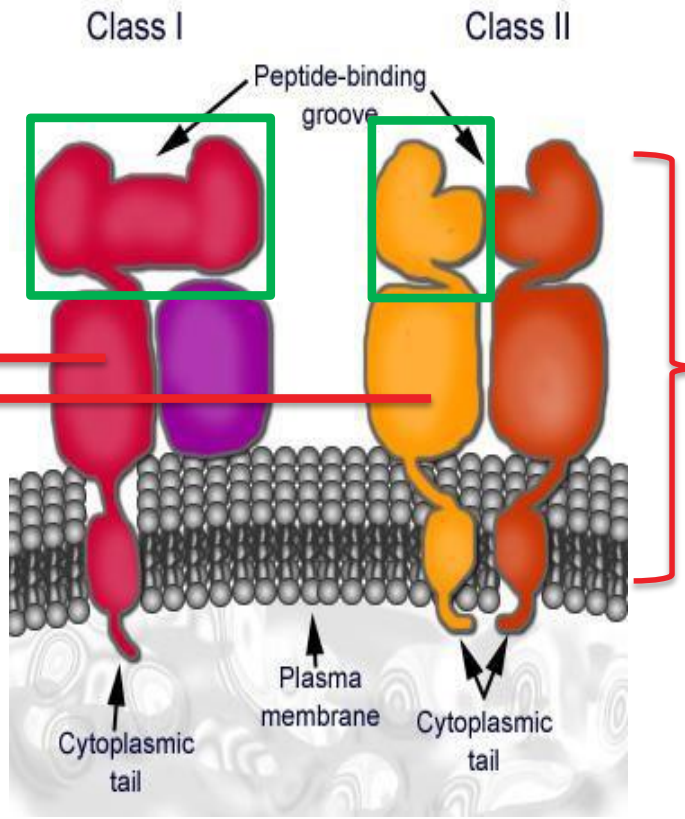
HLA-DRB1, DQB1, DPB1





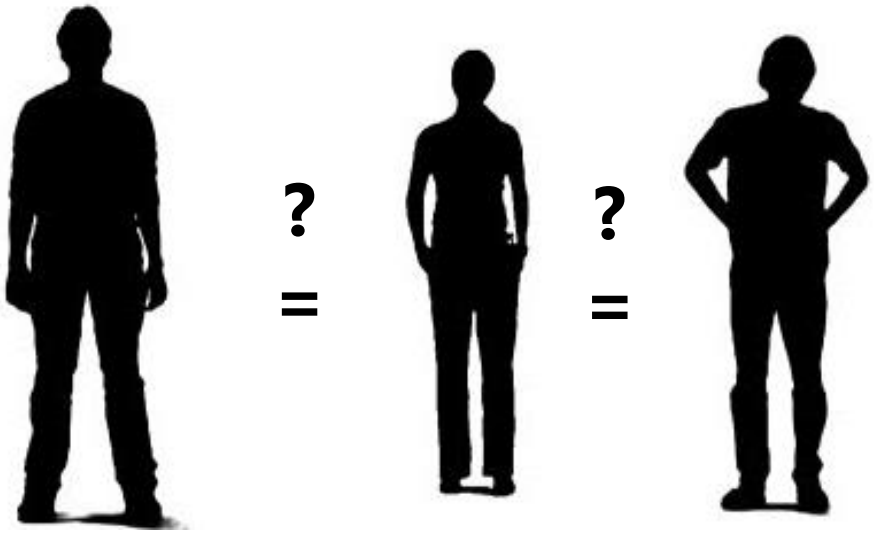
→ HLA genotyping limited to the antigen binding site

Other exon sequences ?



Class II α –chain? DQA1, DPA1?

- **Non coding regions are unknown: expression variants? Nulls?**
- **DRB345?**



Are these donors identical?
Which donor is best matched with our patient?

Techniques of tissue typing & HLA matching

Serology

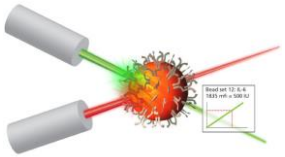
....

1994 Anasetti *et al,* Tranfus; Sci.:

Graf versus host disease and HSCT outcome correlates with the number of HLA-**ABDR** MM

Techniques of tissue typing & HLA matching

Serology



SSP, SSO

....

1994 Anasetti *et al,* Tranfus; Sci.:

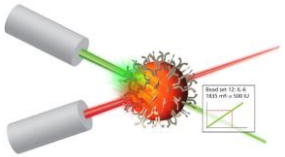
Graf versus host disease and HSCT outcome correlates with the number of HLA-**ABDR** MM

2001 Petersdorf *et al*, Blood:

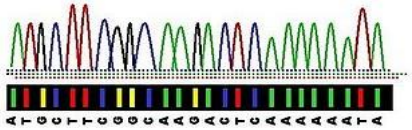
Risk of graft failure increased with incresaing number of **C** MM.

Techniques of tissue typing & HLA matching

Serology



SSP, SSO



Sanger

....

1994 Anasetti *et al,* Tranfus; Sci.:

Graf versus host disease and HSCT outcome correlates with the number of HLA-**ABDR** MM

2001 Petersdorf *et al*, Blood:

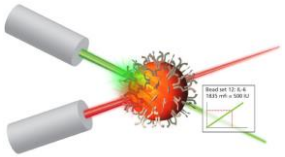
Risk of graft failure increased with incresaing number of **C** MM.

2007 Lee *et al.,* Blood:

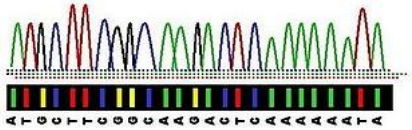
High resolution donor-recipient HLA matching contributes to the succes of unrelated HSCT.

Techniques of tissue typing & HLA matching

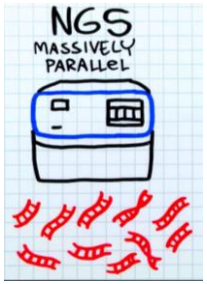
Serology



SSP, SSO



Sanger



....

1994 Anasetti *et al,* Tranfus; Sci.:

Graf versus host disease and HSCT outcome correlates with the number of HLA-**ABDR** MM

2001 Petersdorf *et al*, Blood:

Risk of graft failure increased with incresaing number of **C** MM.

2007 Lee *et al.,* Blood:

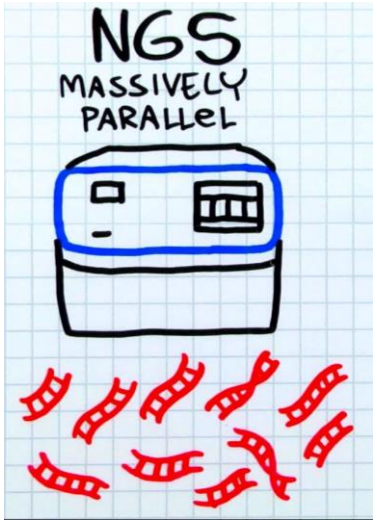
High resolution donor-recipient HLA matching contributes to the succes of unrelated HSCT.

2013 Fernandez-Viña *et al.,* Blood:

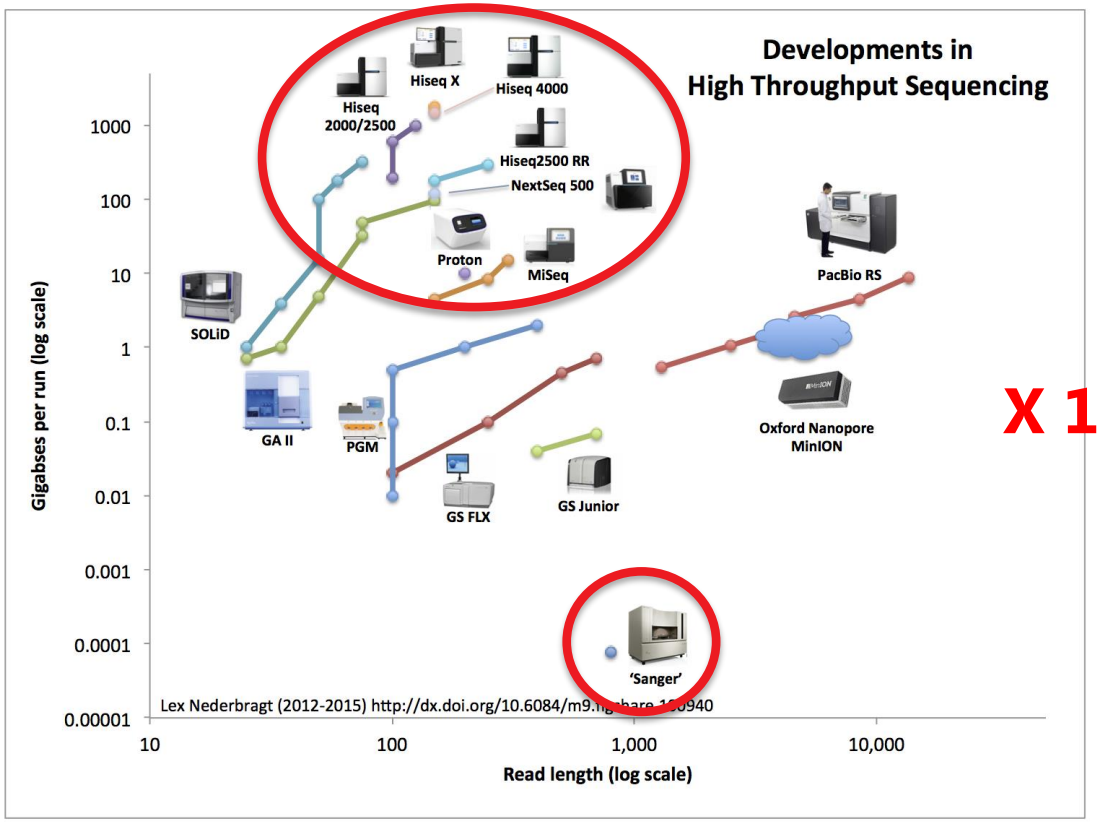
Multiple MM **DQB1, DPB1** and **DRB345** associate with adverse outcomes in HSCT.

New technologies have created opportunities to improve HLA matching and TX outcome

Next Generation Sequencing: opportunities



Throughput

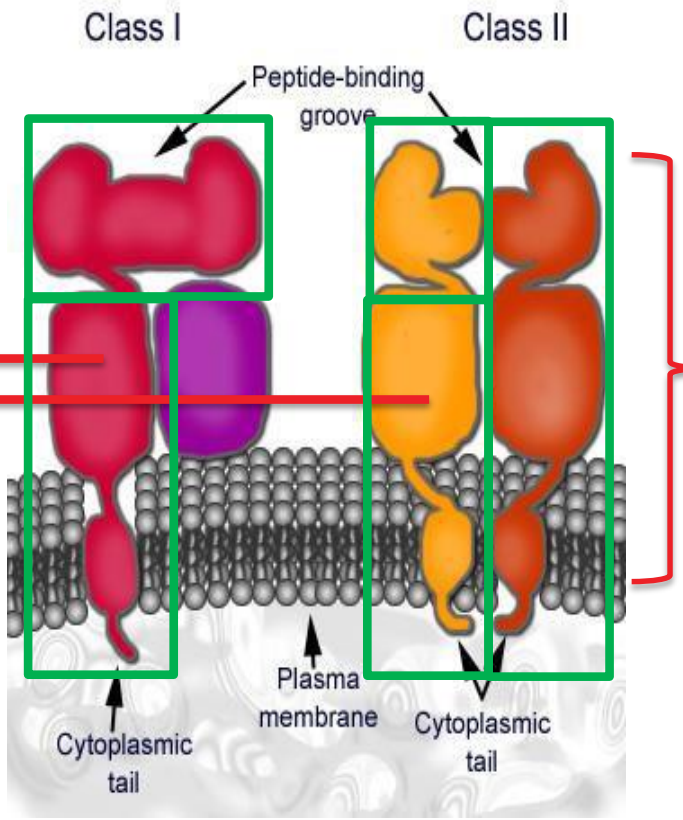


X 100 000

Next Generation Sequencing: opportunities

Other exon sequences ?

√



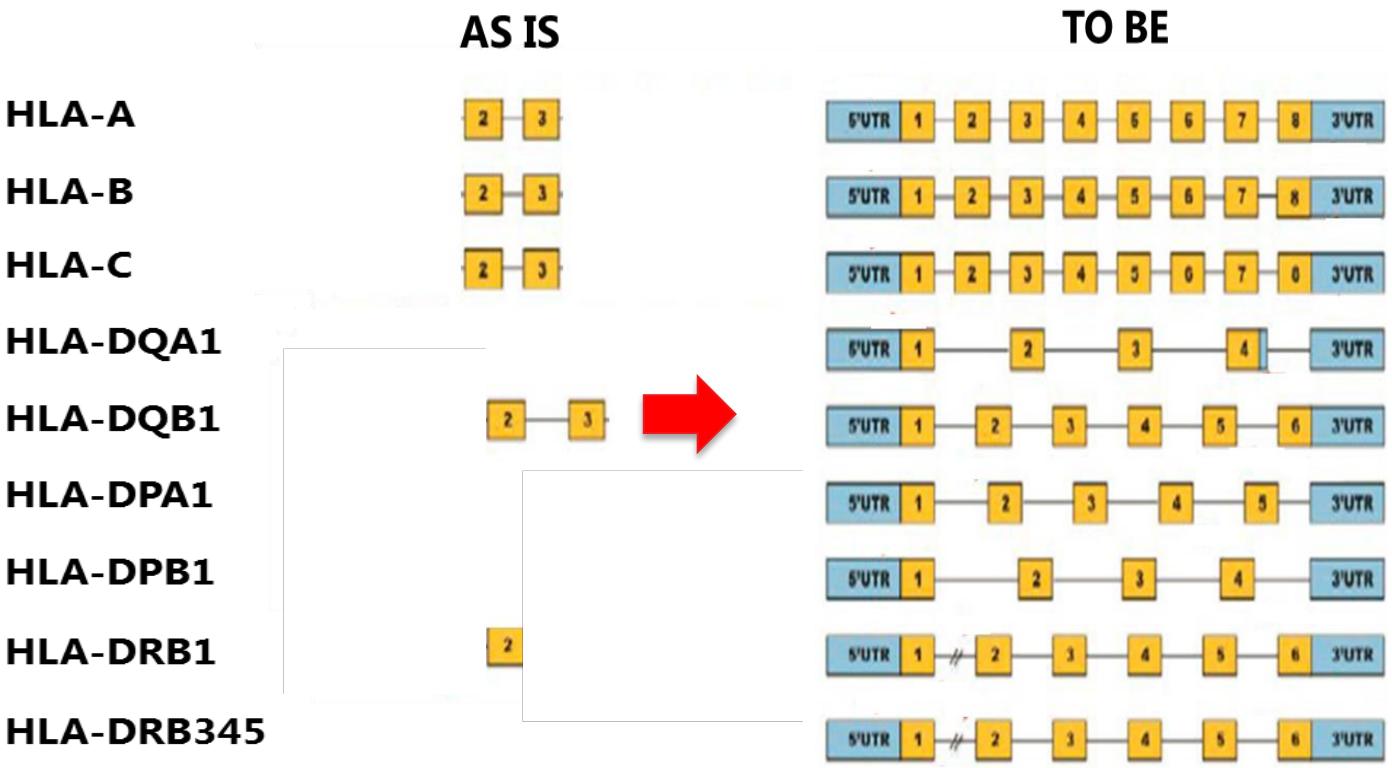
Class II
α –chain?
DQA1, DPA1?

√

- **Non coding regions are unknown: expression variants? Nulls?**
- **DRB345?**

√

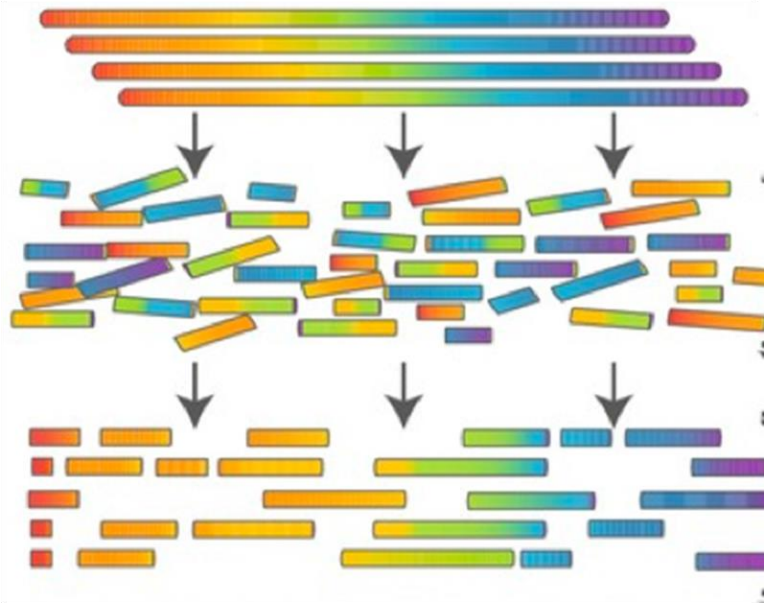
Next Generation Sequencing: opportunities



Short read NGS

Amplification of the complete gene

Short read NGS



ATGCGTTGACCGCTCACGTCGACGTTGCTGAACTCATACTGCGTC
atgcgctgacgtcgtacatcataggagggcctagtacaatgcgccgcactcatcgc

atgcgctgacgtcgtacatcCACGTCGACGTTGCTGAACTCATACTGCG
ATGCGTTGACCGCTtcataggagggcctagtacaatgcgccgcactcatcgct

Amplification of the full gene

Fragmentation

Sequening of small fragments

Computational assembly to full gene

= Challenging task for HLA

Key properties of an HLA NGS workflow

ACCURATE

ROBUST

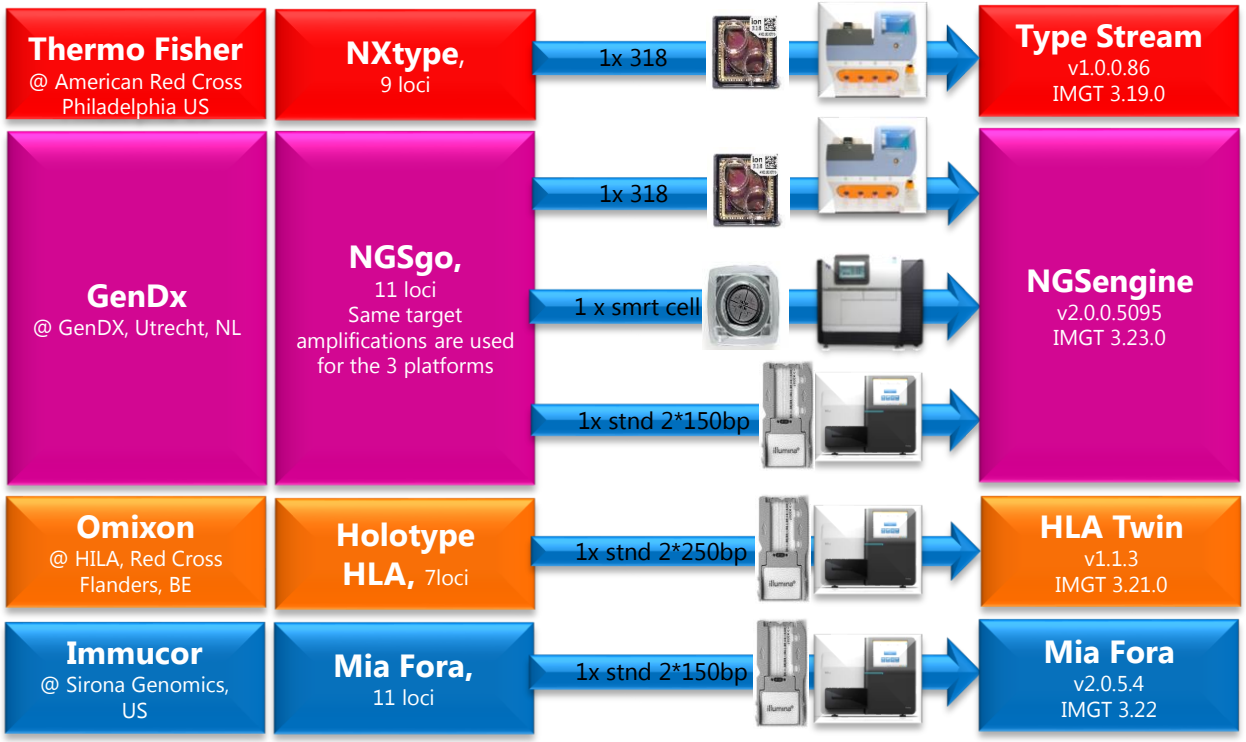
- Confidently identify alleles out of 17000 possible results
- Distinguishing poor amplifying alleles from the back ground
- Confidently call homozygous samples

Evaluation of the commercial NGS/TGS workflows for HLA typing 2016

24 samples

Sanger PRETYPE HLA-A, B, C, DRB1, DRB345, DQA1, DQB1, DPB1, SSO PRETYPE DPA1

Including:
- Variabel DNA concentration (6,5- 40ng/µl),
- Rare, null and new alleles
- Homozygous results



Belgian Red Cross Flanders

ME TO YOU
SAMEN STERK

Selection criteria

GenDx
@ GenDX, Utrecht, NL

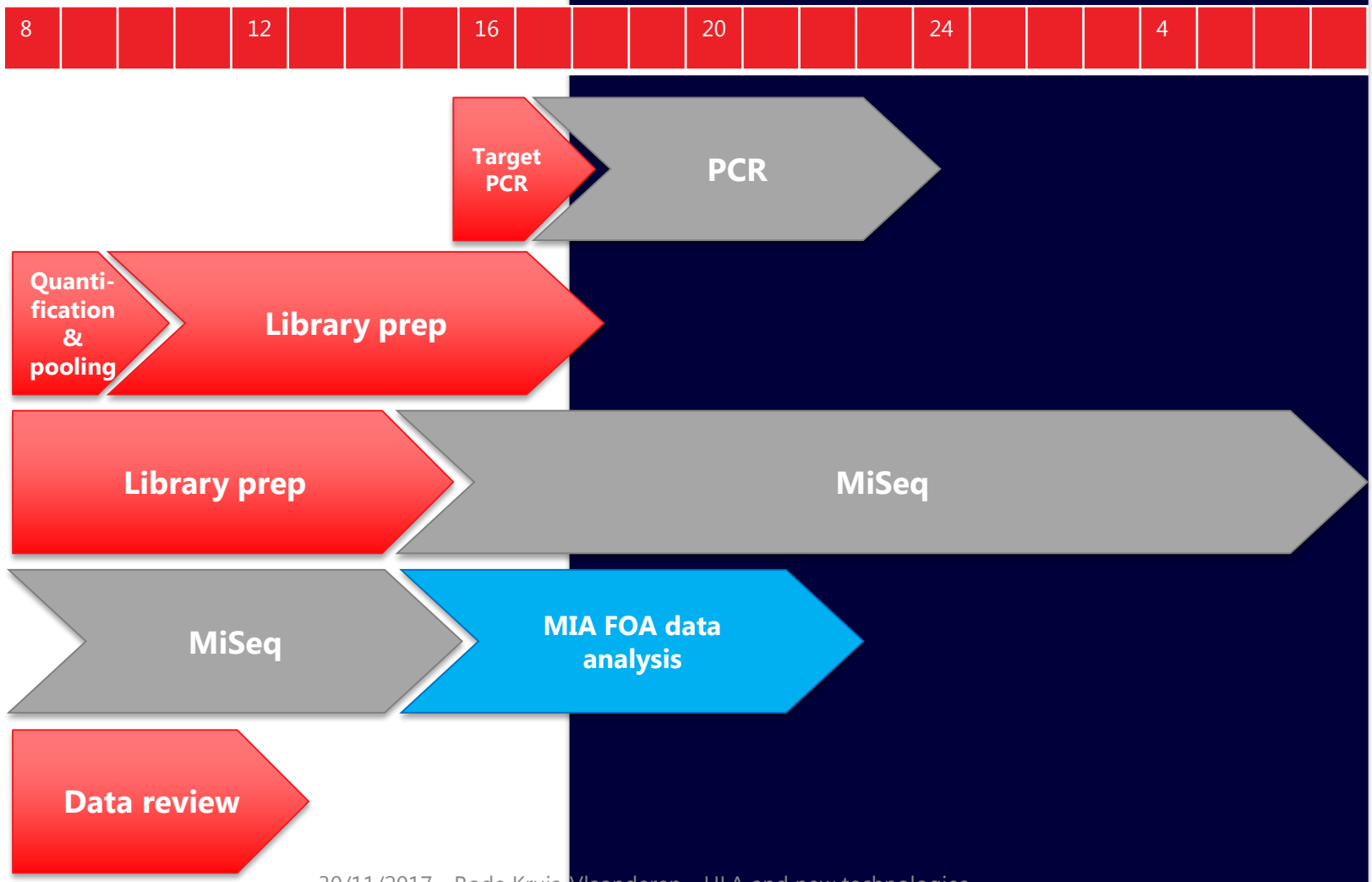
Immucor
@ Sirona Genomics, US

Omixon
@ HILA, Red Cross Flanders, BE

- Data concordancy
- Software reliability
- Data quality
- Robustness
- Cost
- Support
- TAT
- Ambiguities

Immucor
@ Sirona Genomics, US

Immucor:
MIA FORA workflow



Turn Around Time

- NGS workflow takes 6 working days.
  In HILA this workflow starts every Tuesday.

  This means

  - Samples that arrive on Monday will be processed the same week.
    Results will be available Tuesday the week after.
    Minimal TAT 6 working days.
  - Samples that arrive on Tuesday will be processed the next week.
    Maximal TAT 10 working days, excluding failures.

Time saved:

- **Ambiguity resolution (2nd step tests) no longer required.**
- **1st order HSCT (LR SSO) en 2nd order HSCT (HR NGS) can be processed in parallel.**

MIA FORA FLEX Automation

Hamilton Star

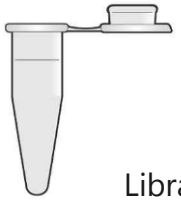
PRE PCR target amplification

Hamilton Starlet

Pico Green quantification

Dilution and pooling

Library prep unitil consolidation



Manual

Size selection PIPIN

Library amplification

Library quantification QUBIT

MiSeq prep

END-TO-END

What to expect from HLA NGS?

Sanger

	Allel 1	Allel 2	Ambiguiteiten
A*	02:01	68:01	A*02:01:01:02L
B*	35:03	44:02	B*44:02:01:02S
C*	04:01	05:01	**C*04:09N,** C*04:30 C*04:82
DRB1*	01:01	04:01	DRB1*01:50 DRB1*01:67
DQB1*	03:01	03:02	**DQB1*03:03:02,03:04:01** DQB1*03:10:01,03:184 DQB1*03:14:02,03:113 DQB1*03:74,03:138 DQB1*03:69,03:70
DPB1*	03:01	14:01	DPB1*104:01, DPB1*124:01 DPB1*351:01 DPB1*498:01
DRB345*	NT	NT	
DQA1*	NT	NT	
DPA1*	NT	NT	

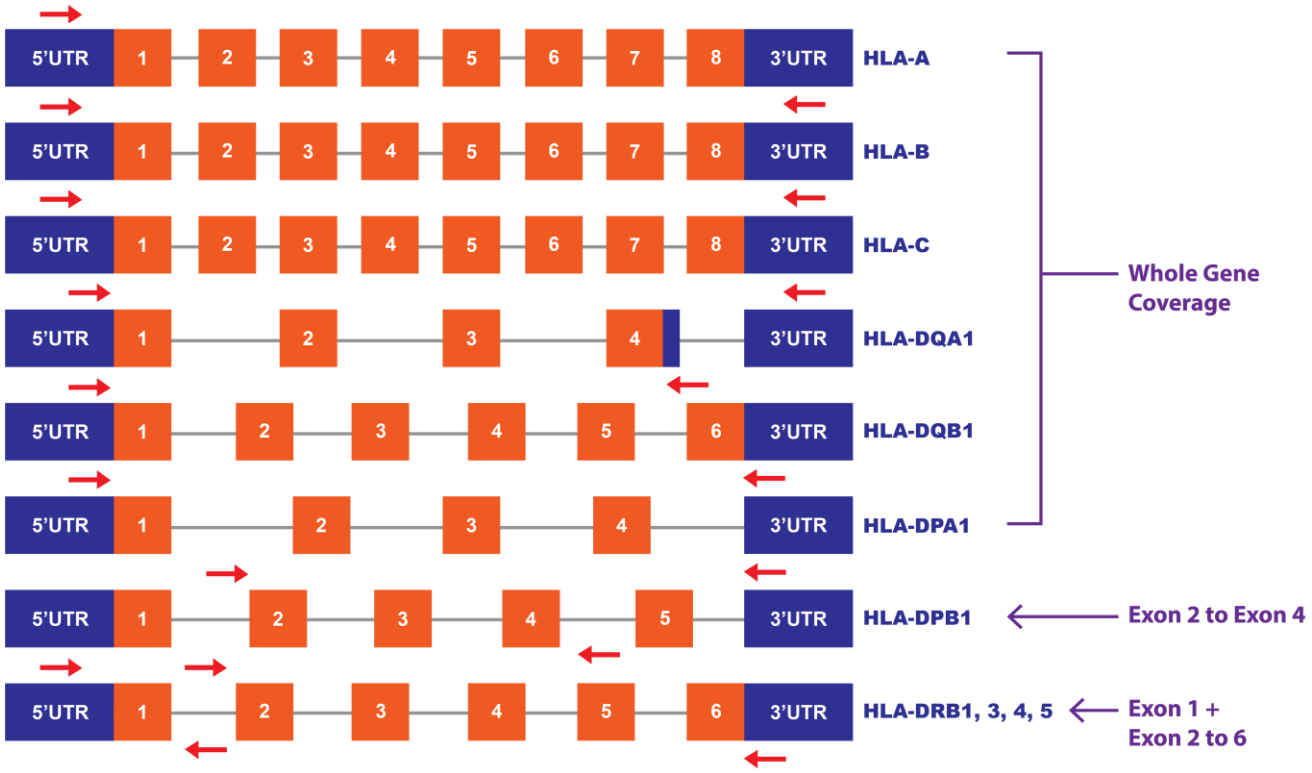
→ outside target ampfication:
Reflex SSP test

→ cis trans ambiguity of common alleles:
Reflex SSP test

What to expect from HLA NGS?

	Sanger			NGS		
	Allel 1	**Allel 2**	**Ambiguiteiten**	**Allel 1**	**Allel 2**	**Ambiguiteiten**
A*	02:01	68:01	A*02:01:01:02L	02:01:01:01	68:01:02	-
B*	35:03	44:02	B*44:02:01:02S	35:03:01:03	44:02:01:01	-
C*	04:01	05:01	**C*04:09N,** C*04:30 C*04:82	**04:01:01:01**	05:01:01:02	No extra test needed !
DRB1*	01:01	04:01	DRB1*01:50 DRB1*01:67	01:01:01	04:01:01:01	-
DQB1*	03:01	03:02	**DQB1*03:03:02,03:04:01** DQB1*03:10:01,03:184 DQB1*03:14:02,03:113 DQB1*03:74,03:138 DQB1*03:69,03:70	03:01:01:01	03:02:01:01	No extra test needed !
DPB1*	03:01	14:01	DPB1*104:01, DPB1*124:01 DPB1*351:01 DPB1*498:01	104:01	14:01	-
DRB345*	NT	NT		DRB4*01:03:01:01		-
DQA1*	NT	NT		02:01:01:01	05:05:01:02	-
DPA1*	NT	NT		01:03:01:01	01:03:01:03	-

Nearly ambiguity free!



→ Very broad target amplification !

Ambiguities are mostly limited to genotype ambiguities HLA-DPB1

What to expect from HLA NGS?

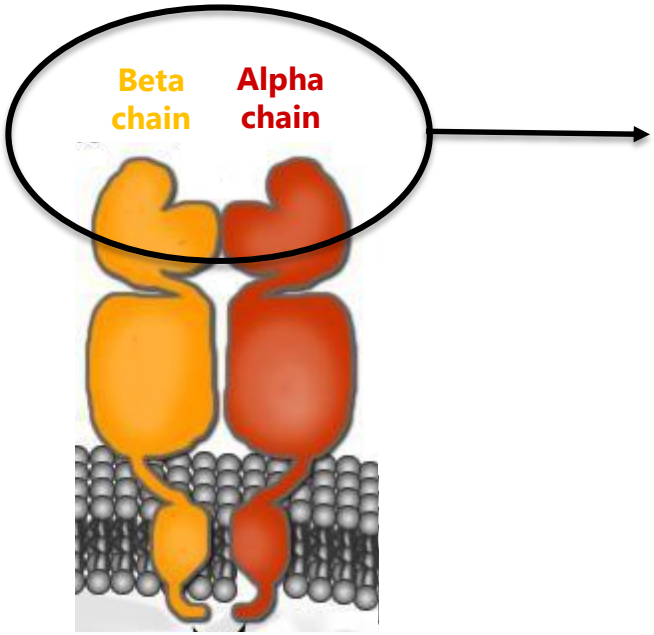
	Sanger			NGS		
	Allel 1	**Allel 2**	**Ambiguiteiten**	**Allel 1**	**Allel 2**	**Ambiguiteiten**
A*	02:01	68:01	A*02:01:01:02L	02:01:01:01	68:01:02	-
B*	35:03	44:02	B*44:02:01:02S	35:03:01:03	44:02:01:01	-
C*	04:01	05:01	**C*04:09N,** C*04:30 C*04:82	04:01:01:01	05:01:01:02	-
DRB1*	01:01	04:01	DRB1*01:50 DRB1*01:67	01:01:01	04:01:01:01	-
DQB1*	03:01	03:02	**DQB1*03:03:02,03:04:01** DQB1*03:10:01,03:184 DQB1*03:14:02,03:113 DQB1*03:74,03:138 DQB1*03:69,03:70	03:01:01:01	03:02:01:01	-
DPB1*	03:01	14:01	DPB1*104:01, DPB1*124:01 DPB1*351:01 DPB1*498:01	**104:01**	14:01	-
DRB345*	NT	NT		DRB4*01:03:01:01		-
DQA1*	NT	NT		02:01:01:01	05:05:01:02	-
DPA1*	NT	NT		01:03:01:01	01:03:01:03	-

What to expect from HLA NGS?

	Sanger			NGS		
	Allel 1	**Allel 2**	**Ambiguiteiten**	**Allel 1**	**Allel 2**	**Ambiguiteiten**
A*	02:01	68:01	A*02:01:01:02L	02:01:01:01	68:01:02	-
B*	35:03	44:02	B*44:02:01:02S	35:03:01:03	44:02:01:01	-
C*	04:01	05:01	**C*04:09N,** C*04:30 C*04:82	04:01:01:01	05:01:01:02	-
DRB1*	01:01	04:01	DRB1*01:50 DRB1*01:67	01:01:01	04:01:01:01	-
DQB1*	03:01	03:02	**DQB1*03:03:02,03:04:01** DQB1*03:10:01,03:184 DQB1*03:14:02,03:113 DQB1*03:74,03:138 DQB1*03:69,03:70	03:01:01:01	03:02:01:01	-
DPB1*	03:01	14:01	DPB1*104:01, DPB1*124:01 DPB1*351:01 DPB1*498:01	104:01	14:01	-
DRB345*	NT	NT		DRB4*01:03:01:01		-
DQA1*	NT	NT		02:01:01:01	05:05:01:02	-
DPA1*	NT	NT		01:03:01:01	01:03:01:03	-

Resolving the Class II alpha chains

Gene	***DRA***	***DRB***	***DQA1***	***DQB1***	***DPA1***	***DPB1***
Alleles	7	2,395	92	1,152	56	942



Sequencing the complete Class II antigen binding site

What to expect from HLA NGS?

	Sanger			NGS		
	Allel 1	**Allel 2**	**Ambiguiteiten**	**Allel 1**	**Allel 2**	**Ambiguiteiten**
A*	02:01	68:01	A*02:01:01:02L	**02:01**:01:01	**68:01**:02	-
B*	35:03	44:02	B*44:02:01:02S	**35:03**:01:03	**44:02**:01:01	-
C*	04:01	05:01	**C*04:09N,** C*04:30 C*04:82	**04:01**:01:01	**05:01**:01:02	-
DRB1*	01:01	04:01	DRB1*01:50 DRB1*01:67	**01:01**:01	**04:01**:01:01	-
DQB1*	03:01	03:02	**DQB1*03:03:02,03:04:01** DQB1*03:10:01,03:184 DQB1*03:69,03:70	**03:01**:01:01	**03:02**:01:01	-
DPB1*	03:01	14:01	DPB1*104:01, DPB1*124:01 DPB1*351:01 DPB1*498:01	**104:01**	**14:01**	-
DRB345*	NT	NT		**DRB4*01:03**:01:01		-
DQA1*	NT	NT		**02:01**:01:01	**05:05**:01:02	-
DPA1*	NT	NT		**01:03**:01:01	**01:03**:01:03	-

Results to the clinic will be limited to 2nd field

What to expect from HLA NGS?

HLA-A*02:01:01:02N

Locus

1st field: allelgroep, group of alleles with similar serological behaviour

2nd field, specification of a unique HLA protein

3rd field synonym mutations.

4th field variations in non coding regions.

Why 2nd field instead of allelic?

Avoiding:
- **most likely interpretation**
- **introducing systemic errors in HLA genotyping:**

Reason:

- Variation in homopolymers (TTTTTTTTTT)
- Short tandem repeat regions (GTGTGTGTGTGT)

Vb: DRB1*15:01:01:01/02/03
intron 2:

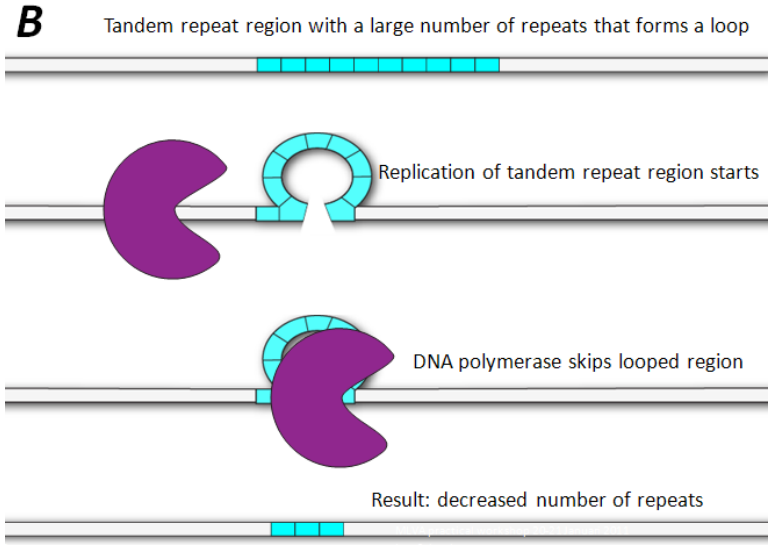
```
gDNA                      5700       5710       5720       5730       5740       5750       5760
DRB1*15:01:01:01     GAGTGTGTGT GTGTGTGTGT GTGTGTGTGT GTGTGTGTGT GTGAGAGAGA GACAGAGAGA GACAGAGAGA
DRB1*15:01:01:02     ---------- ---------- ---------- ---------- -XX------- ---------- ----------
DRB1*15:01:01:03     ---------- ---------- ---------- ---------X XXX------- ---------- ----------
```

Allelels differ in number of GT – AG repeats
Correct assingment of these STRs is hard!

Why 2nd field instead of allelic?

Causes of homopolymer and STR variation:

- **PCR artefacts** Baptiste and Eckert 2012, Environ Mut Gen



→ **Future: avoid PCR**

Why 2nd field instead of allelic?

Causes of STR variation:

- **NGS short read assembly error**

```
gDNA                        5700       5710       5720       5730       5740       5750       5760
DRB1*15:01:01:01      GAGTGTGTGT GTGTGTGTGT GTGTGTGTGT GTGTGTGTGT GTGAGAGAGA GACAGAGAGA GACAGAGAGA
DRB1*15:01:01:02      ---------- ---------- ---------- ---------- -XX------- ---------- ----------
DRB1*15:01:01:03      ---------- ---------- ---------- ---------X XXX------- ---------- ----------
```

GTGTGTGTGTGTGTGTGAGAGAG

GTGTGTGTGTGTGTGTGAGAGAG

GTGTGTGTGTGTGTGTGAGAGAG

Tollerance necessary to assemble short reads generates INDELs in STRs

→ Future: avoiding fragmentation & assembly: single molecule sequencing.

Thesis: A. Senev HILA

Why 2nd field instead of allelic?

1. Allele discrimination based on variation in short tandem repeat area's is a shear guess!

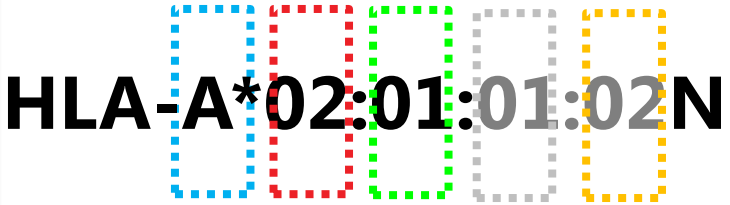
2. A lot of data is still unknown.
   Allele identification of is based on the closest related known genes and often includes mismatches.

3. HLA NGS analysis software's excludes some parts of the genes:
   - some intronic data
   - UTR: includes promotor: expression control.

What is HLA allelic resolution worth using NGS today?

What to expect from HLA NGS?

HILA chooses to offer an HLA 2nd field result without assumptions that is relevant and usable for the clinic.



Locus

1st field: allelgroep, group of alleles with similar serological behaviour

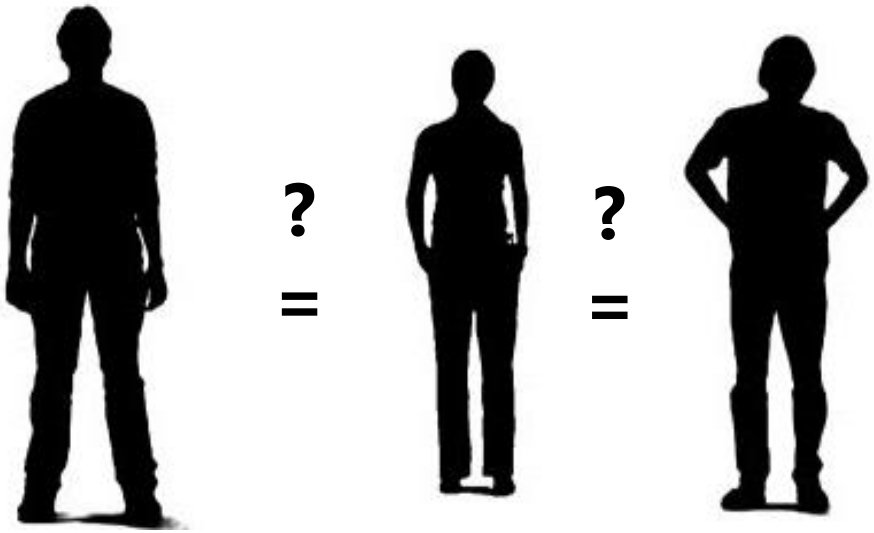
2nd field, specification of a unique HLA protein

3rd field synonym mutations.

4th field variations in non coding regions.

Added value of NGS:

- All exon ambiguities are resolved
- All published expression variants are excluded
- Other relevant loci are included.



NGS will allow us to further define which donor is the most suitable.

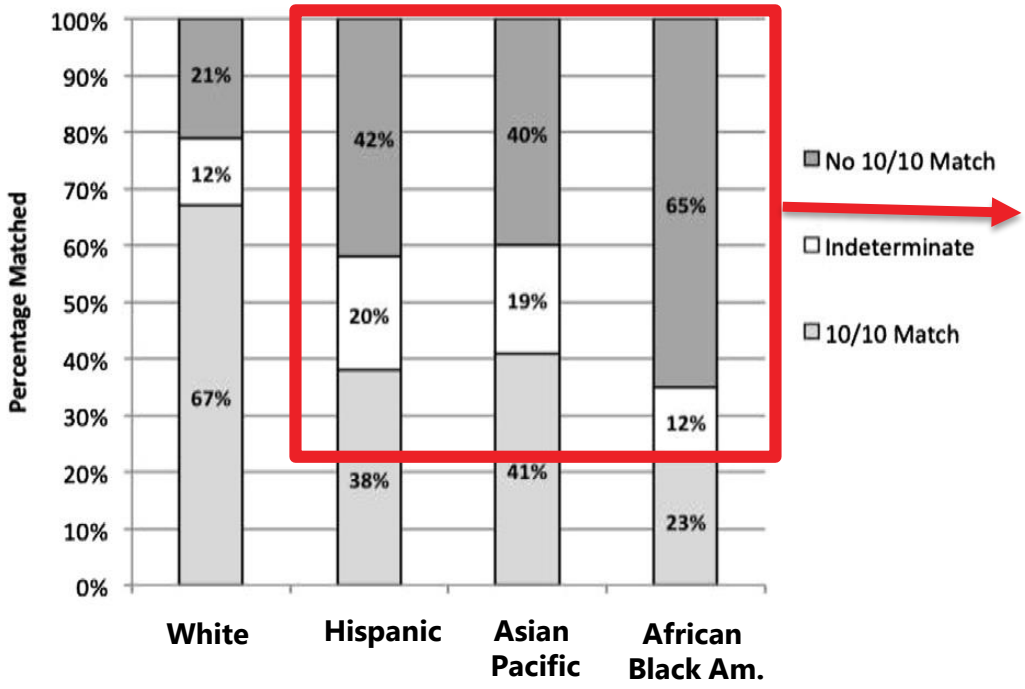
31 million chances to survive

DID YOU KNOW?

We are mainly still looking for young men of immigrant origin or with a mixed background.

Probability of finding an HLA-matched donor is dependent on the allele and haplotype frequencies present in de registries.

Chance to have a 10/10 match vary in race



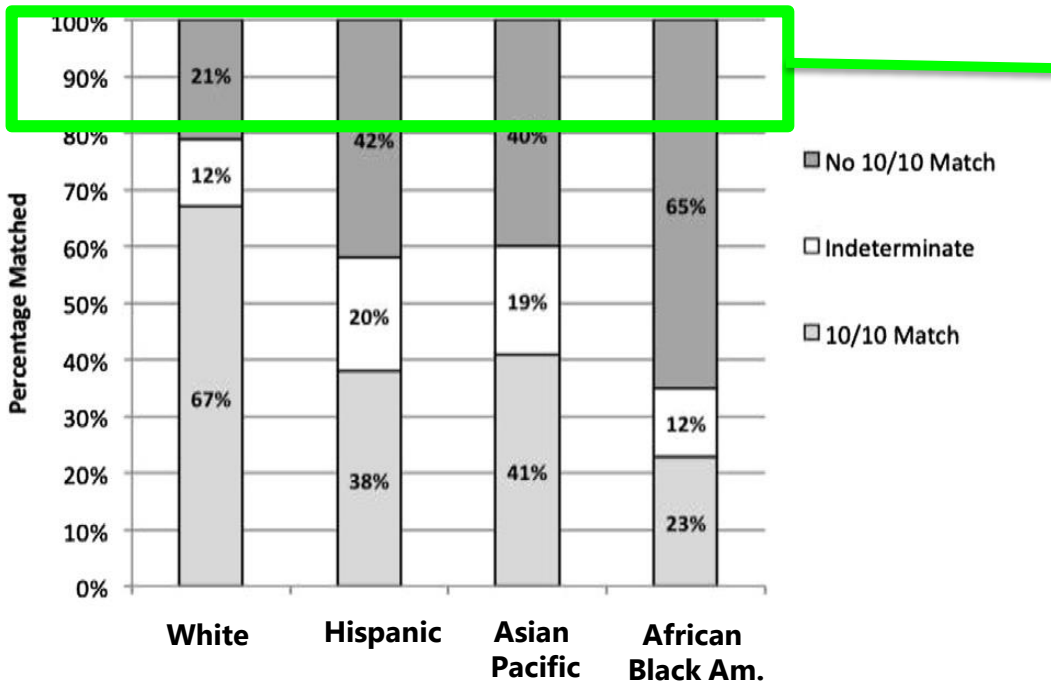
- Population diversity
- Lower number of donors

→ Need for more donors in the registries

Dehn et al., Biol Blood Marrow Transplant 21 (2015)

Probability of finding an HLA-matched donor is dependent on the allele and haplotype frequencies present in de registries.

Chance to have a 10/10 match vary in race



Dehn et al., Biol Blood Marrow Transplant 21 (2015)

More donors won't close the last gap:

- rare alleles
- new alleles
- aberrant linkages
- ....

Define permissable mismatches

Some MM are better tolerated than others…

Important factors are:

- Position of the mismatch in the antigen

Fleischhauer et al., 2012 Lancet Oncol
Pidala et al., 2013 Blood.

- Expression levels: high expression levels may facilitate the presentation of foreign antigens leading to improved immune surveillance.

Petersdorf et al., 2015 N. Eng J Med
Petersdorf et al., 2016 Blood

To conclude: if we want to

Find the best fit donor for each patient

Find appropriate donors for patients with rare, new or non LD alleles

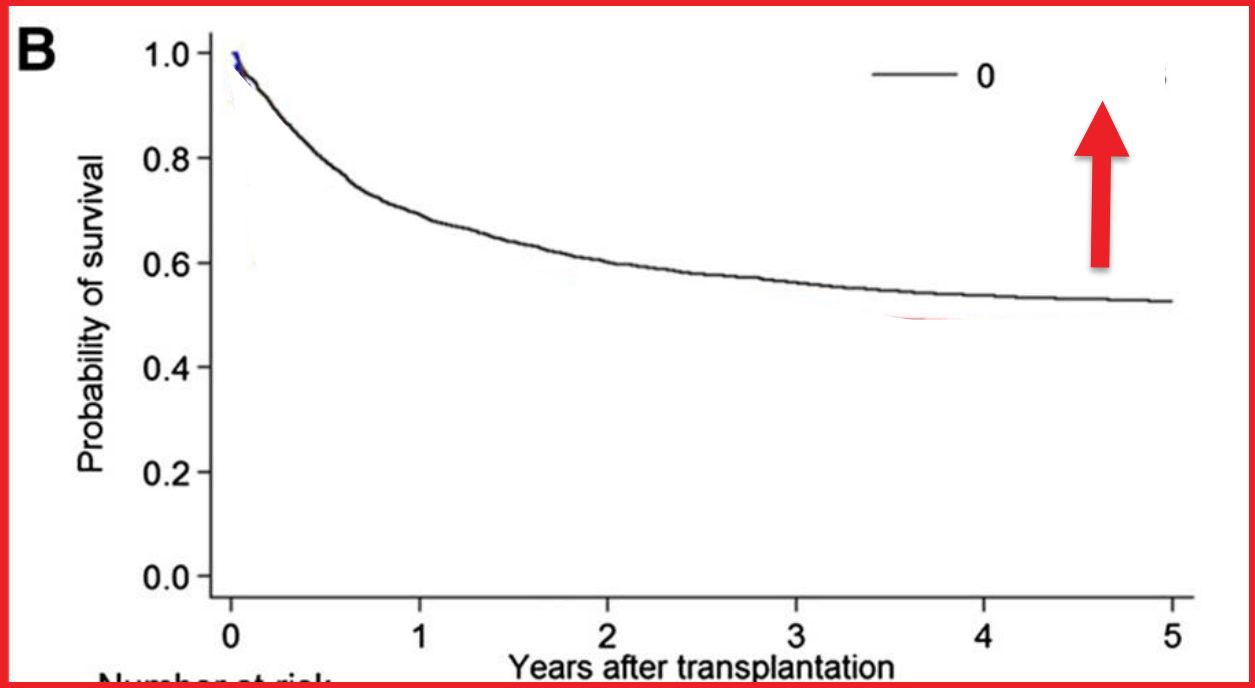
Use the benefit of the Graft Versus Leukemia effect

To adequately understand TX outcome and POST TX monitoring

We need to:

Use the opportunities NGS has to offer

Be careful with data interpretation – allelic resolution



Probability of survival even with 10/10 matched donors can be improved

NGS will further help us to do so

Thanks to

Belgian
Red Cross
Flanders