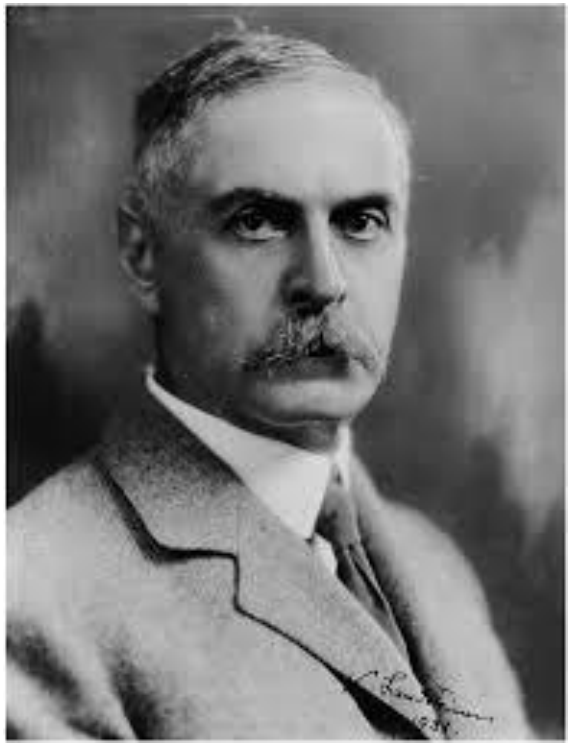
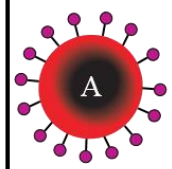
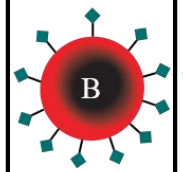
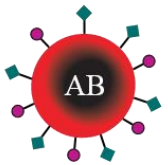
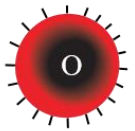
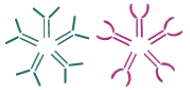
Genotyping of blood donors

Ellen van der Schoot, MD, PhD
Department Experimental Immunohematology
Sanquin Research, Amsterdam
The Netherlands

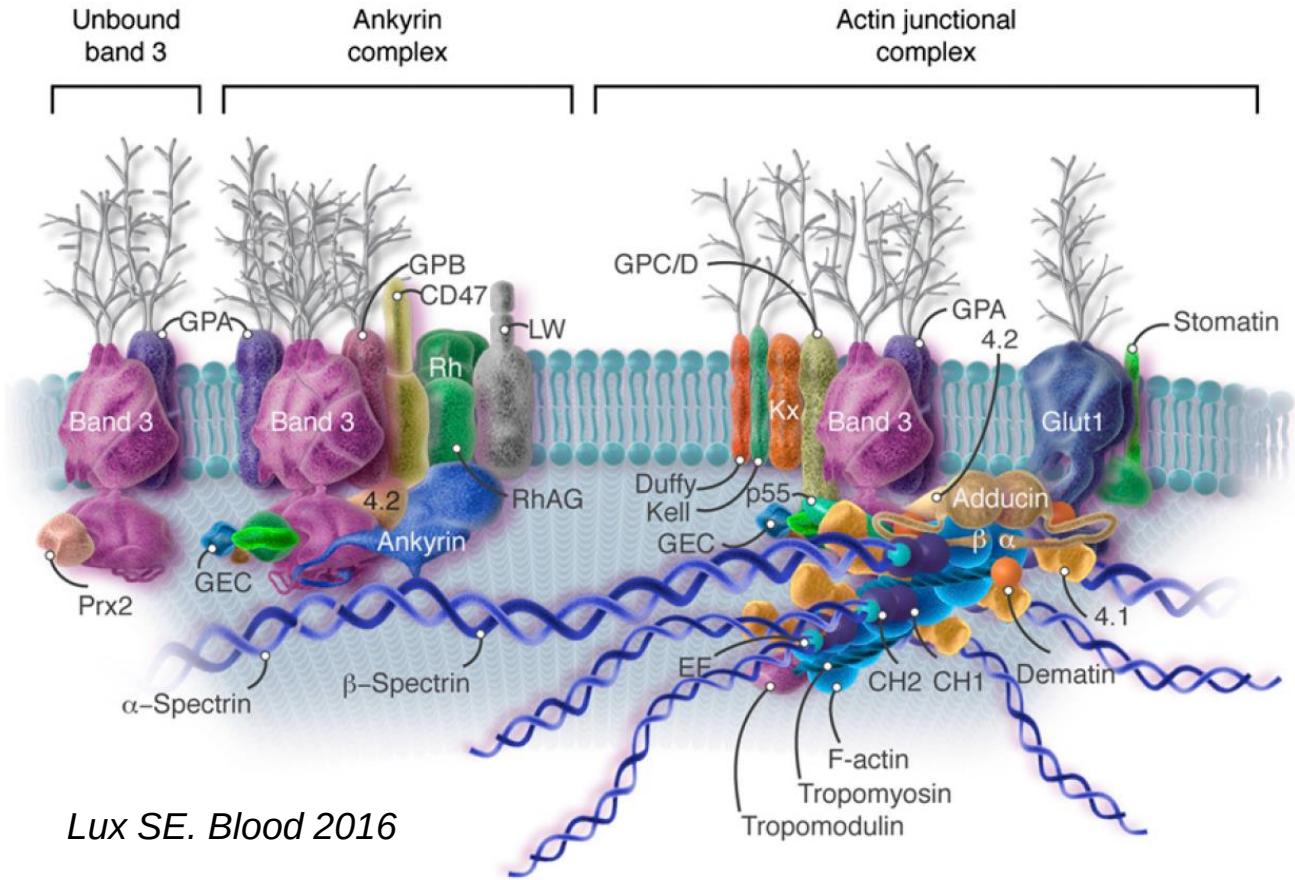
Karl Landsteiner



1901

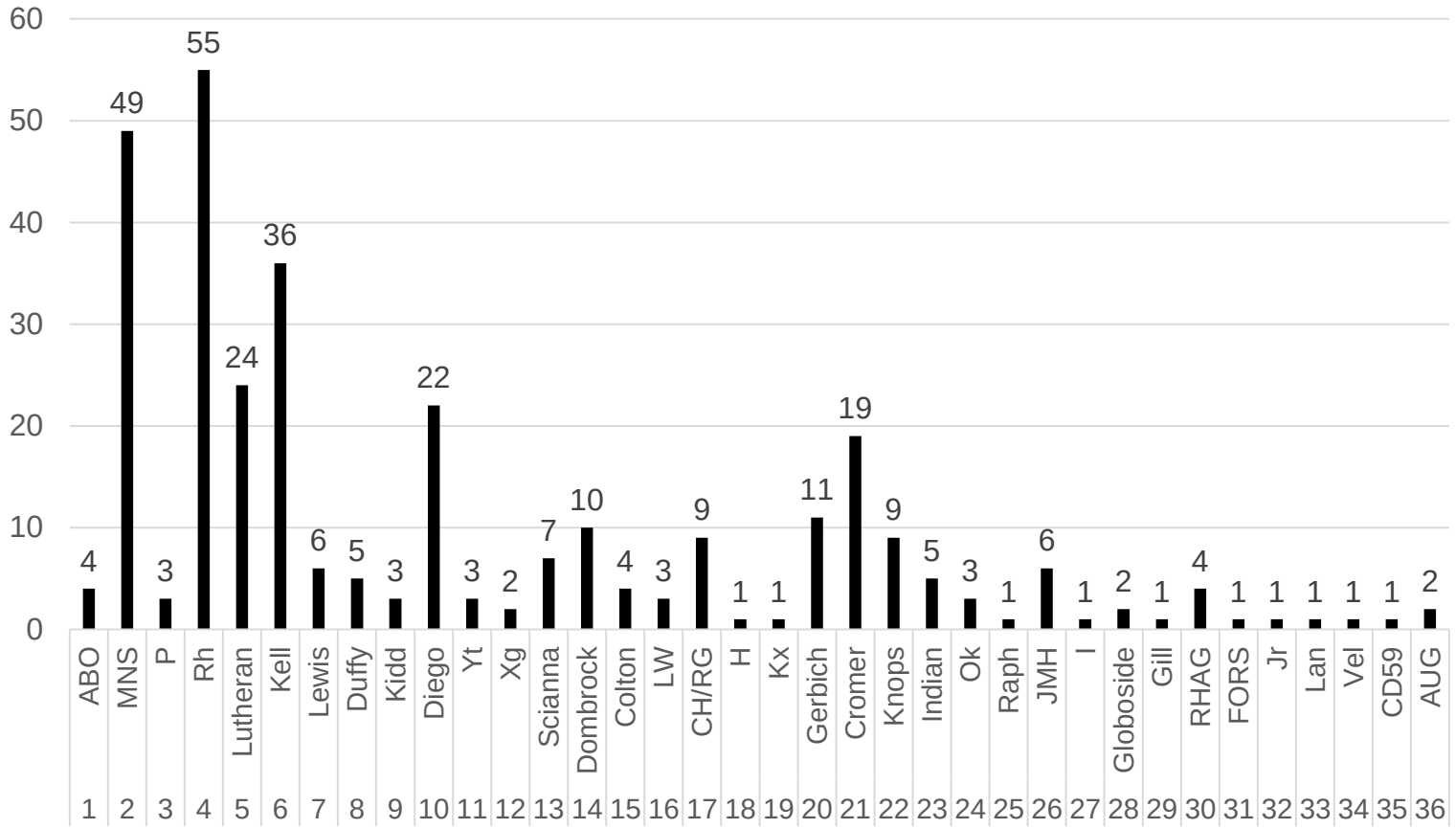
	Group A	Group B	Group AB	Group O
Red blood cell type				
Antibodies in Plasma	 Anti-B	 Anti-A	None	 Anti-A and Anti-B
Antigens in Red Blood Cell	A antigen	 B antigen	 A and B antigens	None

Red cell (glyco)proteins/lipids carry 36 Blood Group systems

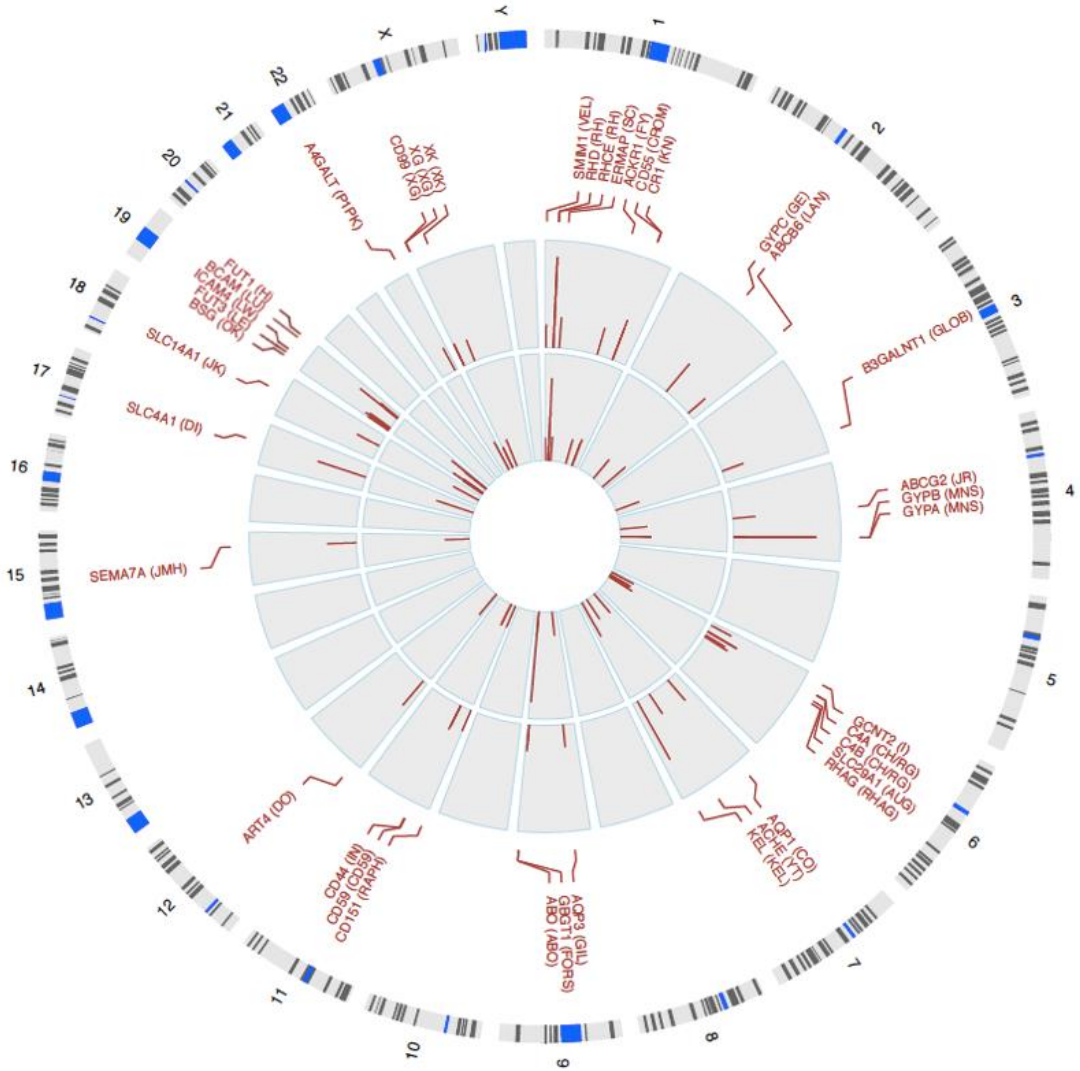


Lux SE. Blood 2016

ISBT nomenclature of blood group antigens

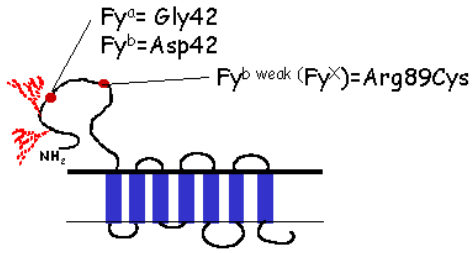


Molecular basis of all Blood Group Systems is known

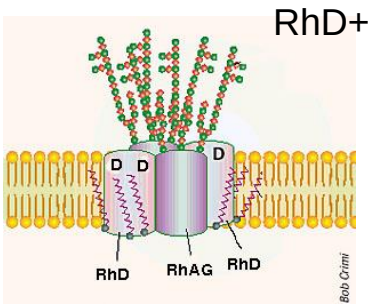


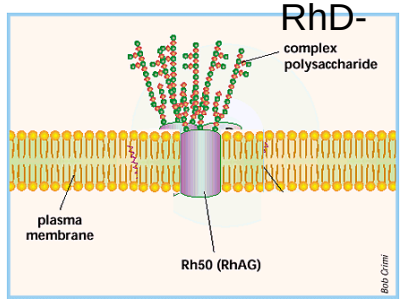
Blood group polymorphisms

- Single (or multiple) amino acid change e.g. Fy(a) / Fy(b)

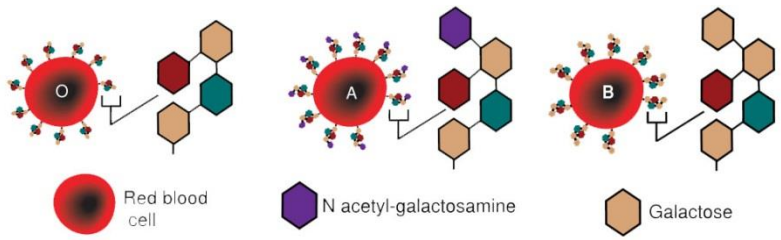


- Presence or absence of protein e.g. RhD+ / RhD-,



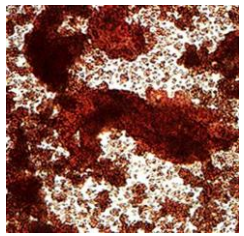


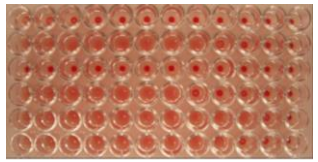
- Carbohydrate difference
- e.g. ABO(A/B)



Serology versus Genomics

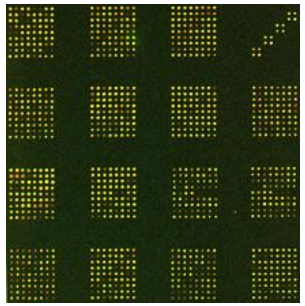
20th century

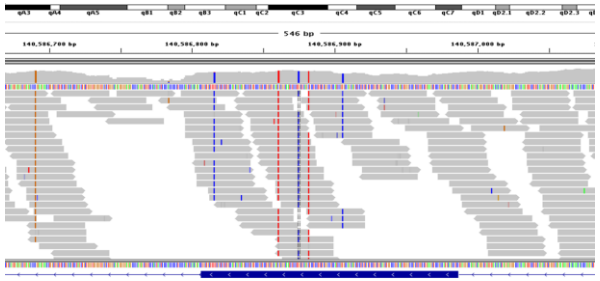




1 single antigen/test
only RBC test

21th century





Upto 10-100 thousands antigens /test
Universal test

Routine methods for blood donor typing anno 2019

Antigens	Presently applied method	Consequence
RBC antigens	Serology	Limited number of antigens
HPA antigens	Genotyping	Limited number of donors*
HNA antigens	Hardly done	
HLA antigens	Genotyping	Limited number of donors*

*Sanquin has 5000 (= 1.4%) HLA + HPA typed donors

Many commercial genotyping assays are available

Supplier	Name	Number of blood group systems	Number of antigens (excl variants)	Number of SNPs	Method	Through-put
Progenika (Grifols)	IDCORE XT	10	37	29	Luminex xMAP	16/ 4hrs
Immucor	PreciseType HEA	11	36	24	Bead micro-array	96/ 5 hrs
MRC-Holland	MLPA	18	54	82	MLPA	32/ 24 hrs
Agena Bioscience	Hemo ID™ DQS	3-12	61	7-33	MALDI-TOF MS	3000/8hrs
Life Technologies	Taqman Open Array	variable	42	32	RQ-PCR	96 / 8 hrs
Beckman	GenomeLab SNP stream	6	19-22	11	Single base extension	384/ day?
Applied Biosystems	SNaPshot	1-10	<26	5-39	Minisequencing	Variable (medium)
AXO	HIFI blood 96	9	22	11	microarray	96/ 4.5 hrs
Innotrain	RBC-Ready Gene	1-7	Several panels	<24	SSP-PCR	96/ 3-4 hrs

Why is donor genotyping still not implemented

- Pricing:
  - Serology is very cheap

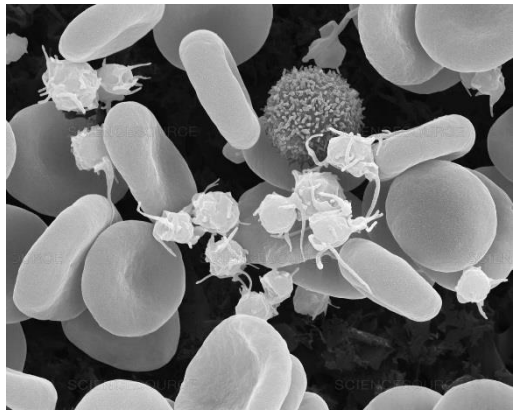
- Conservatism:
  - Serology is very well implemented, and difficult to outcompete
  - Huge impact on equipment and ICT

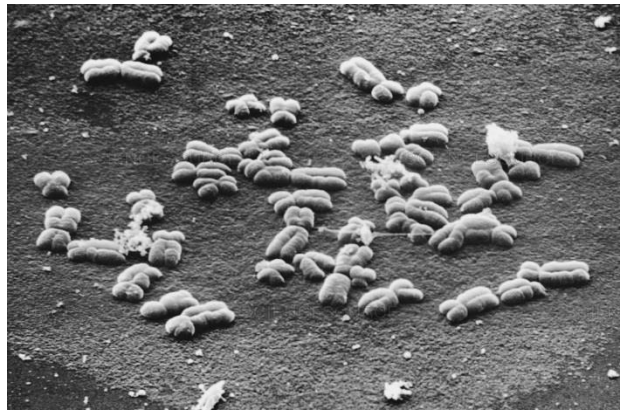
➢ **Genotyping should become better than serology**
  - ➢More comprehensive
  - ➢More reliable
  - ➢Additional advantages
  - ➢Pricing

Aim of the project

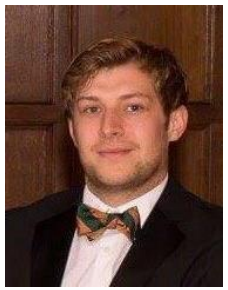
To develop an affordable and comprehensive DNA based donor typing platform for:

- Red blood cell antigens (RBC)
- Human platelet antigens (HPA)
- Human neutrophil antigens (HNA)
- Human leucocyte antigens, class I and II (HLA)
- Beyond blood group antigens





Blood transfusion Genomics Consortium



UNIVERSITY OF CAMBRIDGE

Adam Butterworth
Nick Gleadall
Willem Ouwehand
Karyn Megy

NHS Blood and Transplant

Nick Watkins

Sanquin

Barbera Veldhuisen
Ellen van der Schoot

ThermoFisher SCIENTIFIC
affymetrix

Jeremy Gollub
Laurent Bellon

New York Blood Center

BWH

Connie Westhoff
William Lane

UNIVERSITY OF OXFORD

David Roberts
Gil McVean

LUND UNIVERSITY

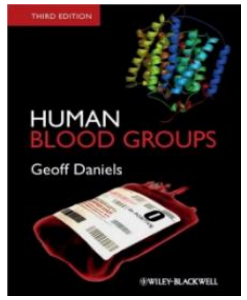
Mattias Möller
Jill Storry
Martin Olsson

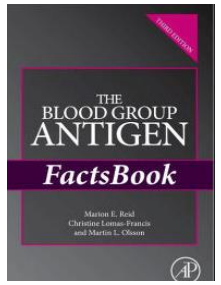
LRG

Aoife McMahon

Custom axiom array ThermoFisher

Selection of variants from databases (Africa Variant Project, 100,000 Genomes Project, gnomAD and TopMED) , books, ISBT, transfusions labs





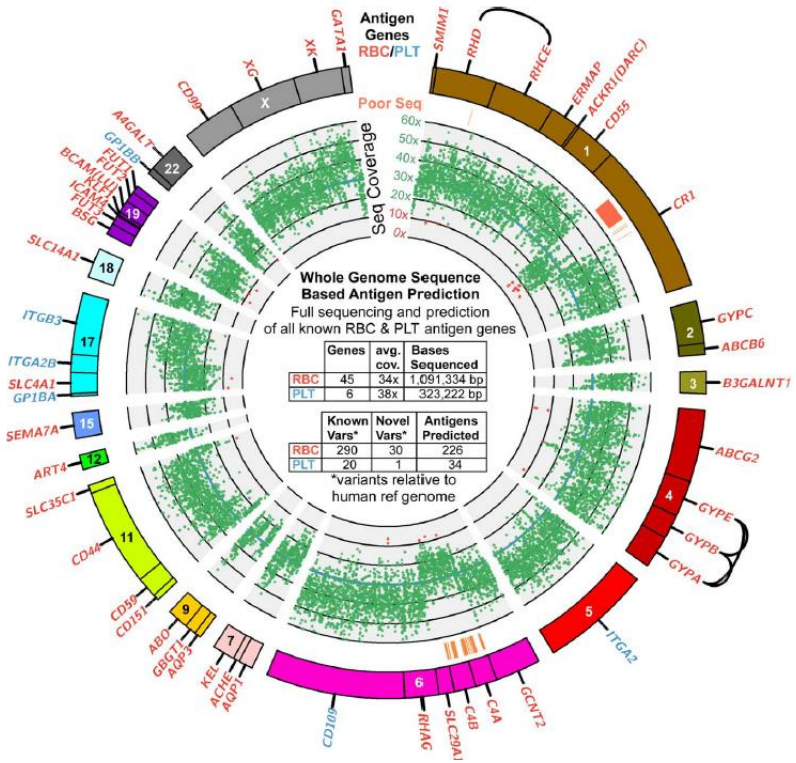
ISBT dbRBC

ERYTHROGENE

BloodAntigens.com

Blood Group Analysis Pipeline from WES data

NGS data



Consolidated Databases

Analysis algorithm

Blood Group Report

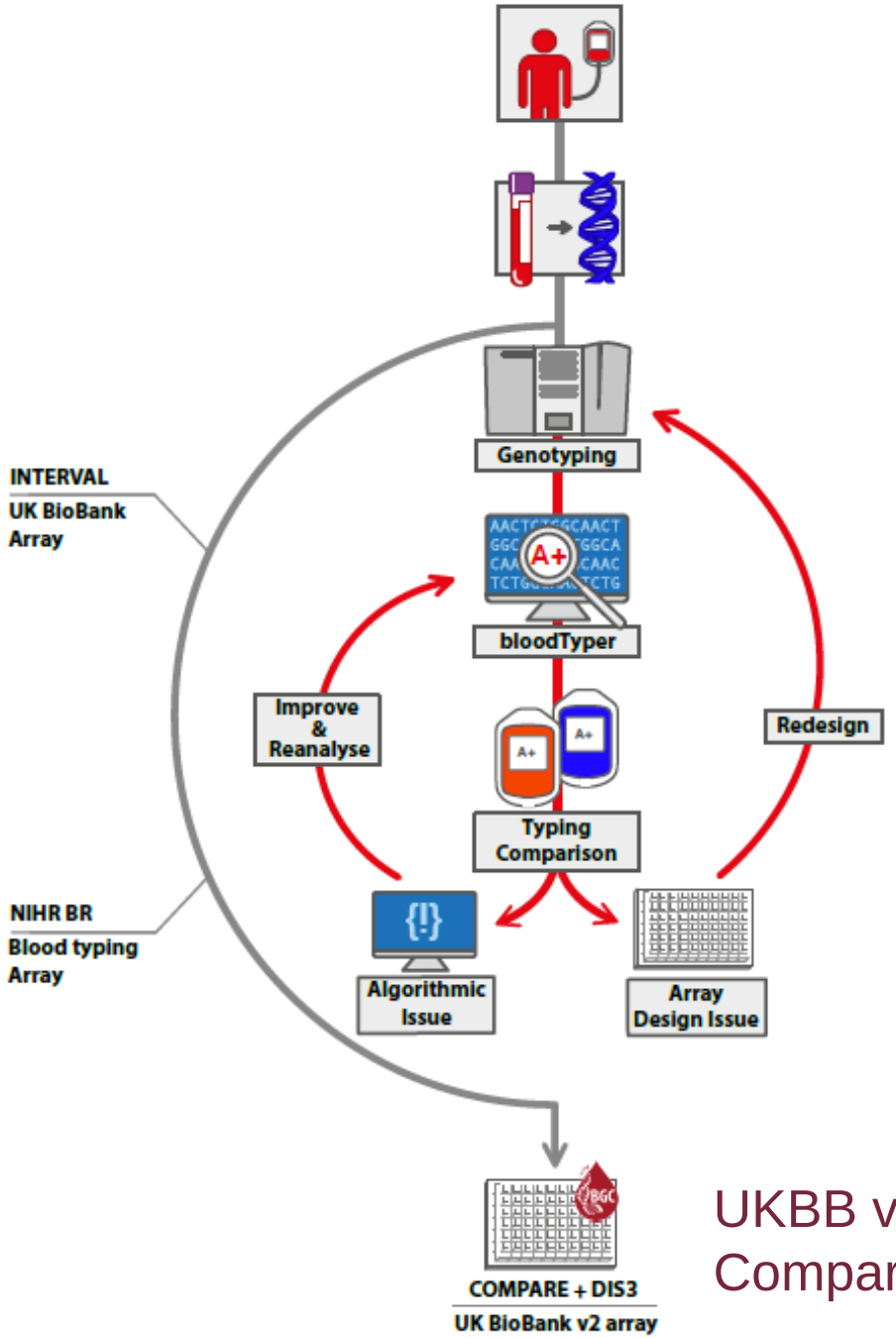
bloodTyper

Lane *et al.*, 2016, *Transfusion*
Lane *et al.*, 2018, *Lancet Haematology*

Workflow

Selection of content of array:
~5000 probesets

- 1572 probesets for RBC, HPA and HNA
- 768 probesets copy number variation: RHD, MNS
- HLA class I and II probesets
- Donor related probesets (iron regulation, storage)



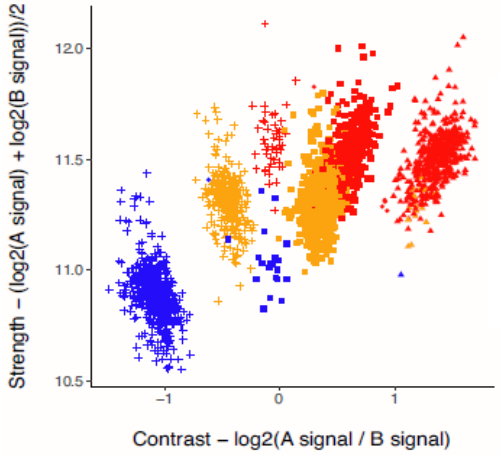
UKBB v2 array
Compare/DIS-III

First results with UK BioBank Axiom Array (0.8M snvs)

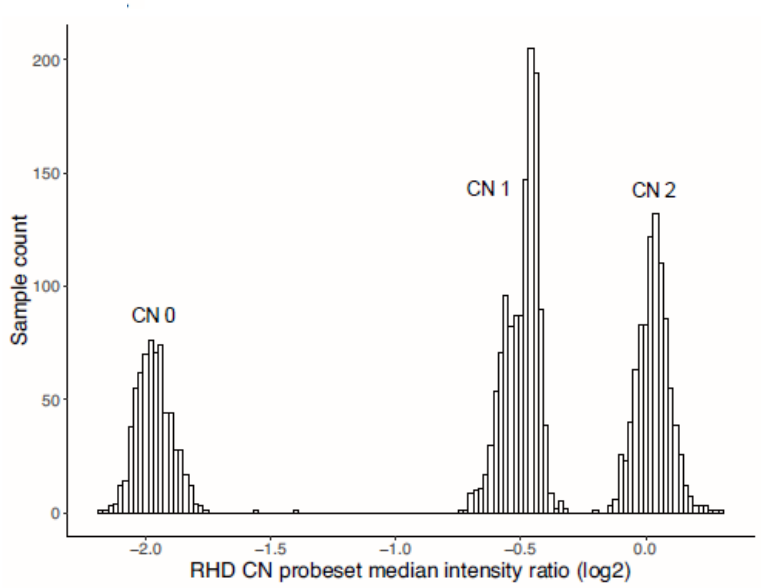
- INTERVAL study : 50.000 donors
- Blood groups retrieved from NHSBT database
- Good results for some antigens
  - S/s – 99.8%
  - C(w) – 100%
  - Lu(b) – 100%
- Some more challenging
  - ABO – 95.4%
  - D - 32.8%
  - M – 27.5%
  - N – 69.6

Blood Group	Individuals Typed
ABO	47,694
Rh D	47,691
Rh C	47,679
RH c	47,686
Rh E	47,686
Rh e	47,681
Rh C(W)	27,682
Kell (K)	47,687
k (k)	5,493
P1	5,493
Jka	26,867
Jkb	26,576
M	26,173
N	3,280
S	24,431
s	17,217
Fya	17,887
Fyb	16,493
Lua	6,243
Lub	10,005
Kpb	9,837
Kpa	6,730
Lea	5,967
Leb	4,387

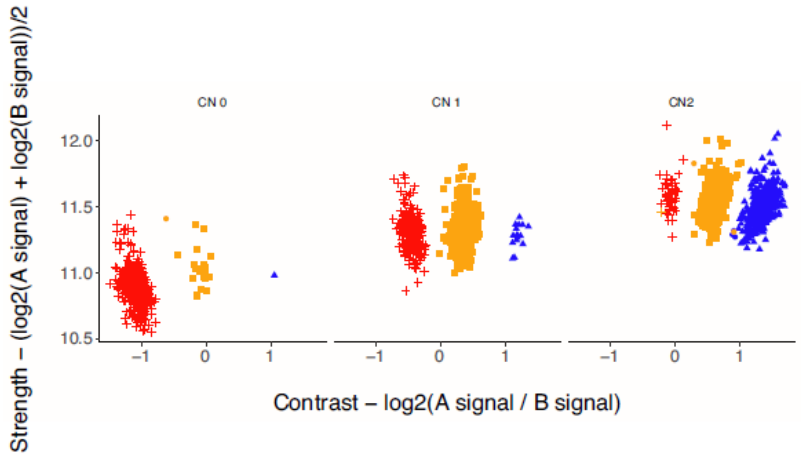
Adding extra probes for homologous genes (RH, MNS)



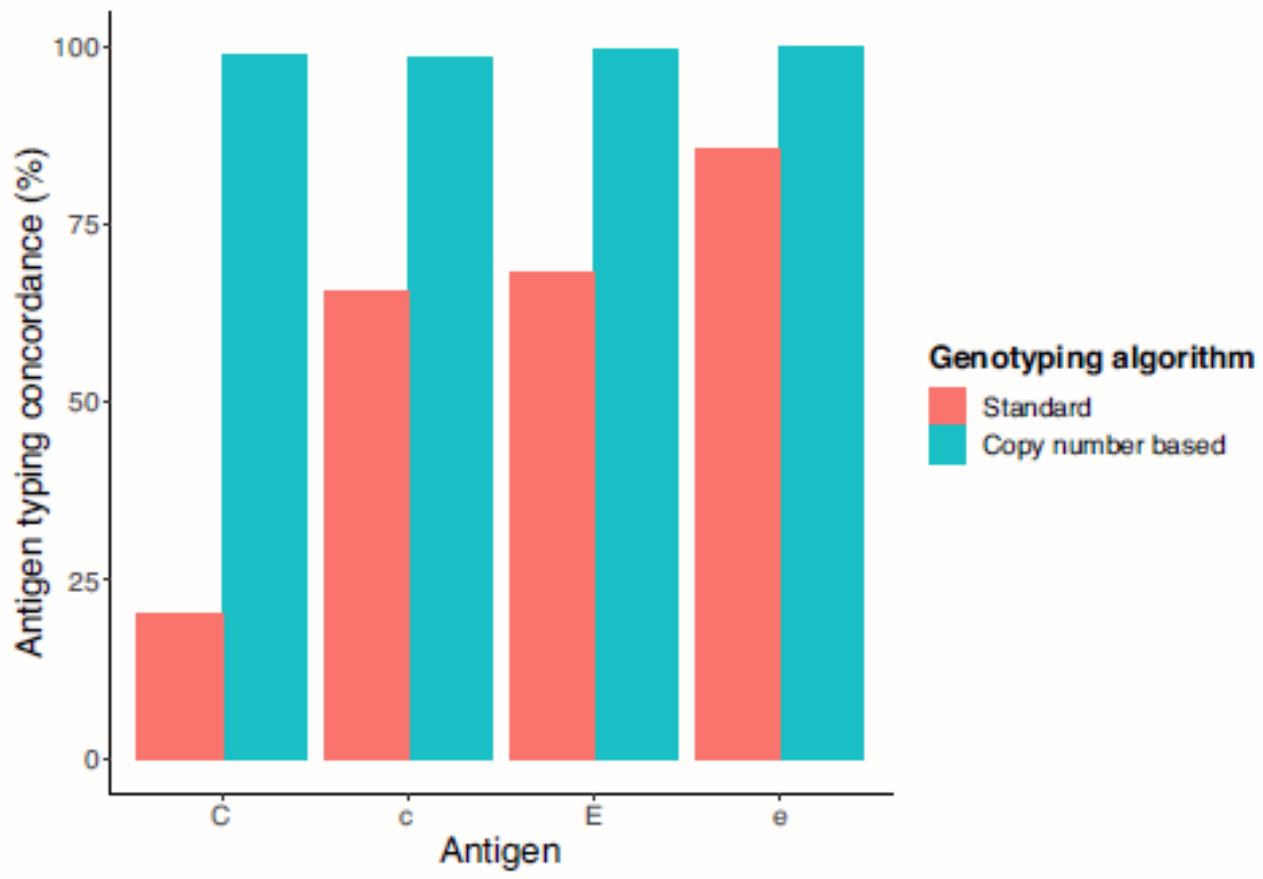
C/c typing



RHD copy number

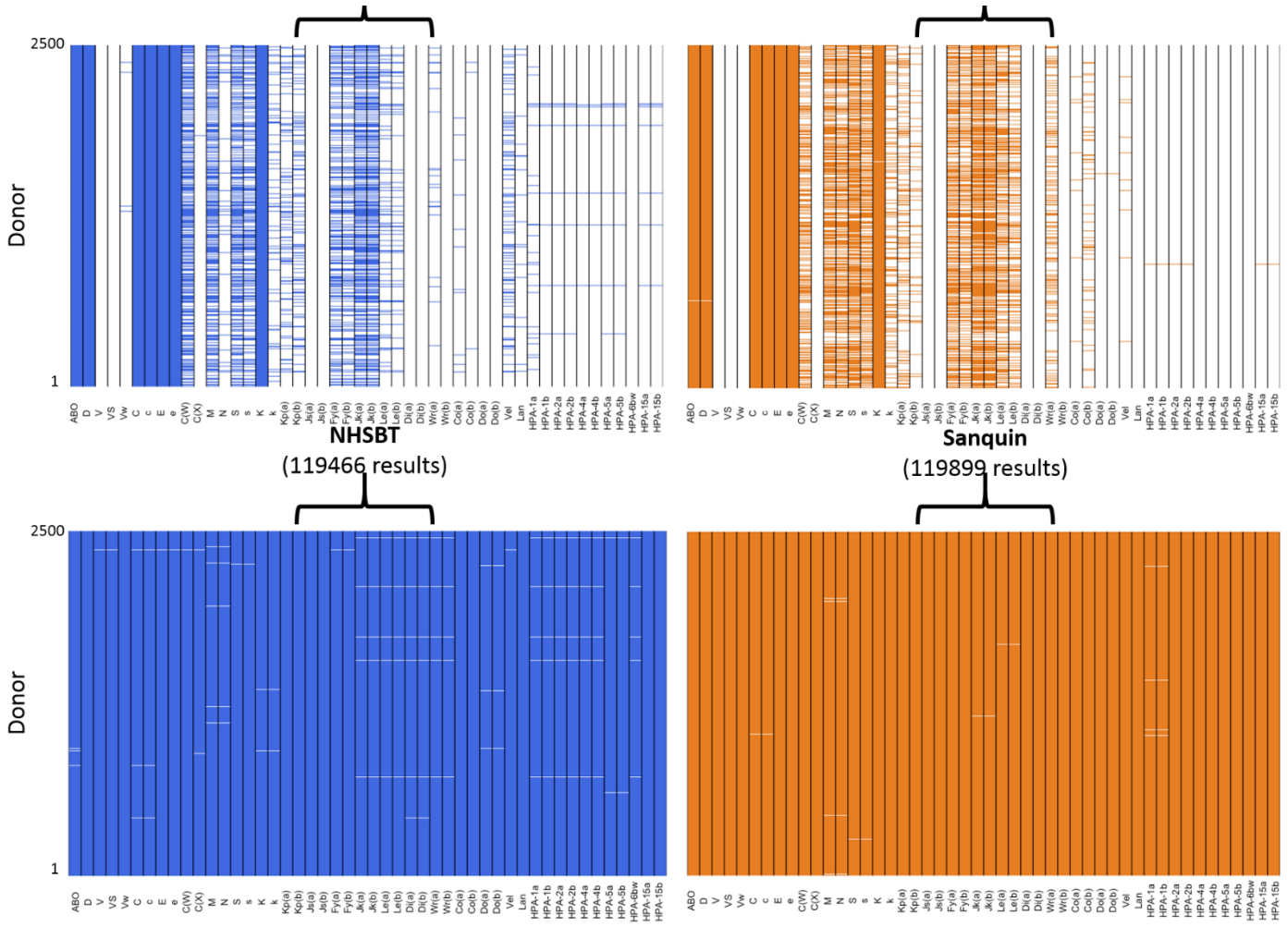


Greatly improved concordance by extra CNV probes



Comparison serology and genotyping in 7473 donors

Compare UK (n=4791) Donor InSight III NL(n=2682)



Tiffany Timmer

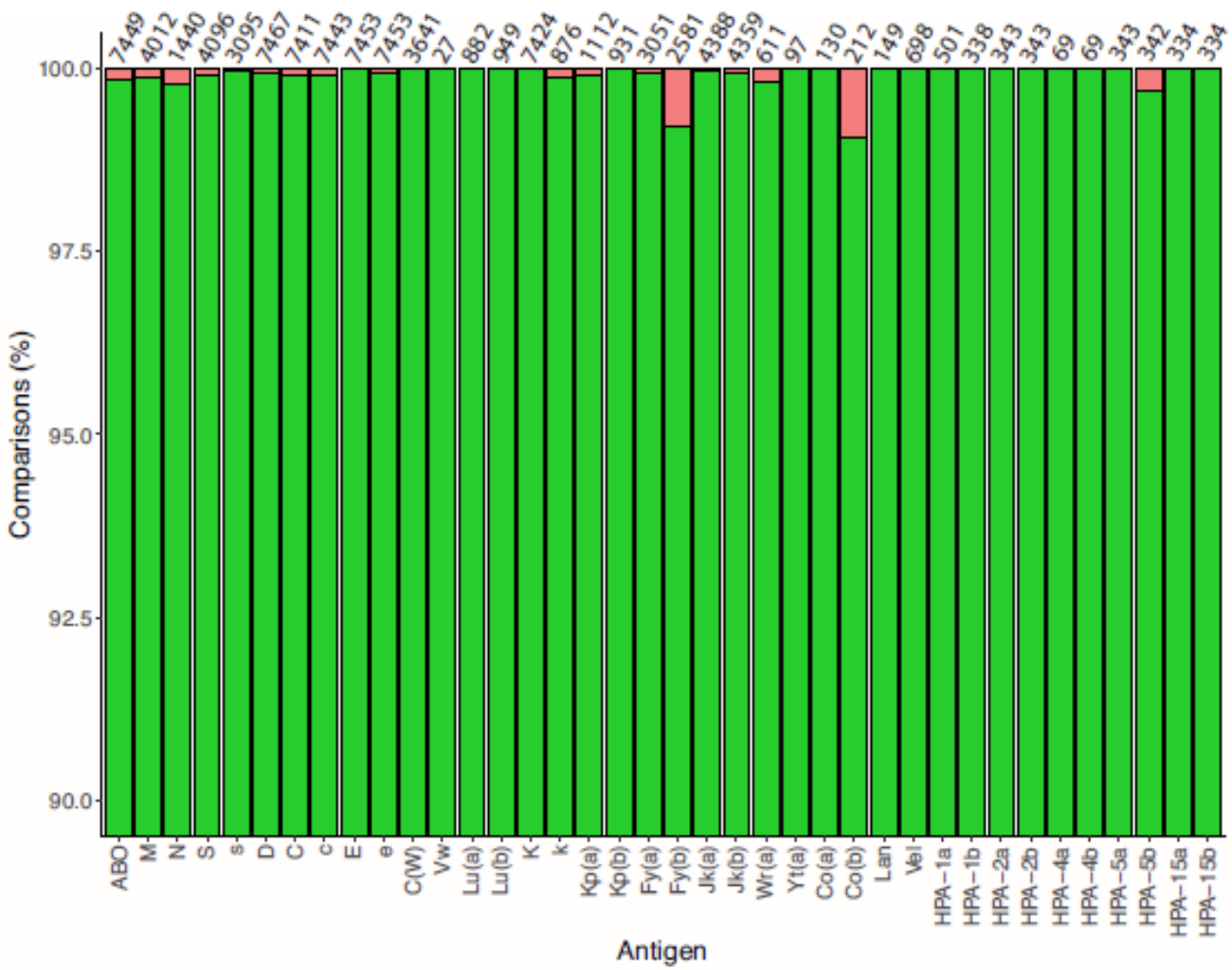
Serology:

13,23/ don

Genotyping:

47,9/donor
+
224 extra / donor

99.92% concordancy between genotype and serology

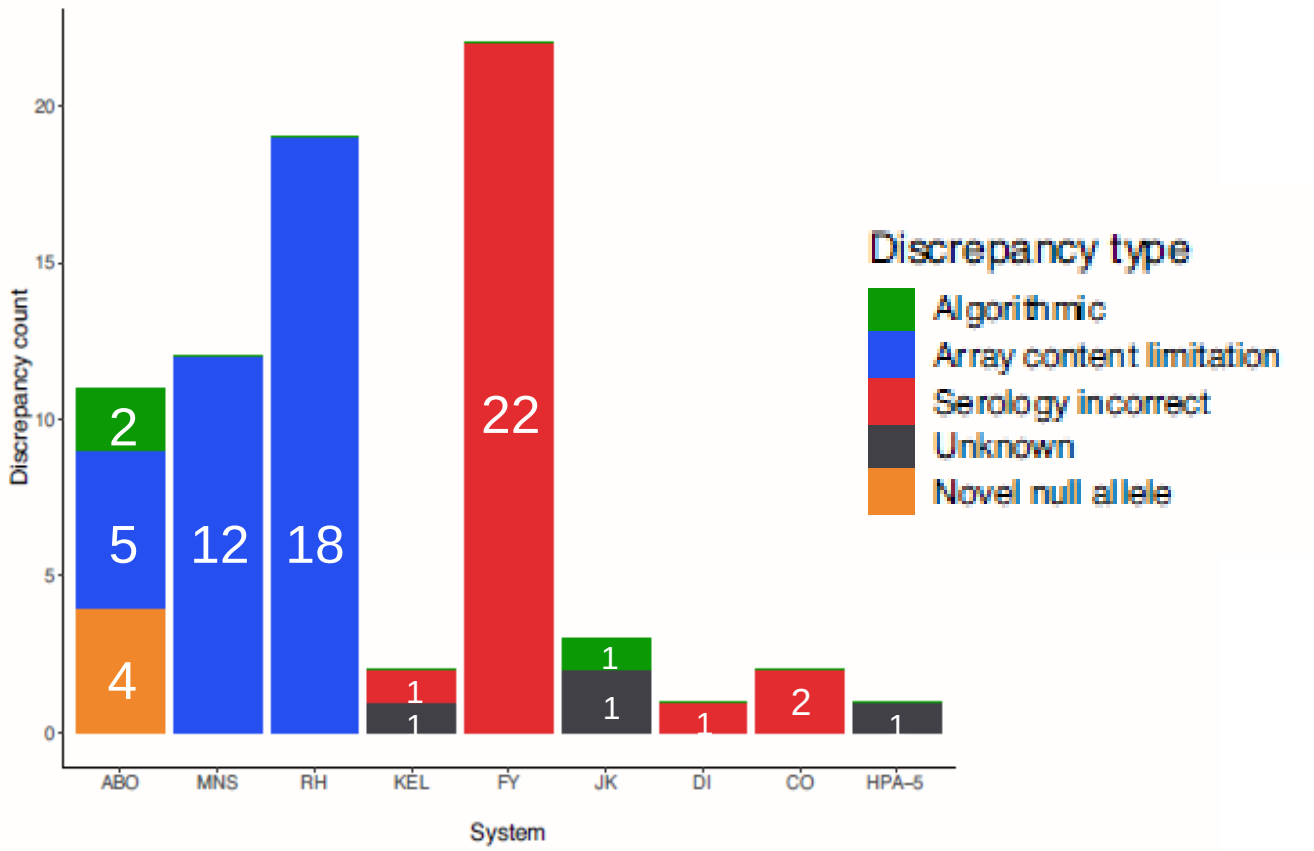


Number of comparisons

73 discrepancies in 92,387 comparisons

38 antigens

Breakdown of 73 discrepancies



Analysis of discrepancies: Sequencing and Additional serology

38 rare donors identified by Bloodtyper array

Rare donors: negative for HFA antigens; phenotype < 1 in 1000

- U−,
- Rh26−,
- CEAG−,
- Lu13−, Lu(b−), Lu8−
- k−, Kp(b−), Js(b−),
- Yt(a−),
- Sc1−,
- Hy−, Jo(a−),
- Co(a−),
- Kn(a−), McC(a−), Sl3−,
- Vel−

Example donor request

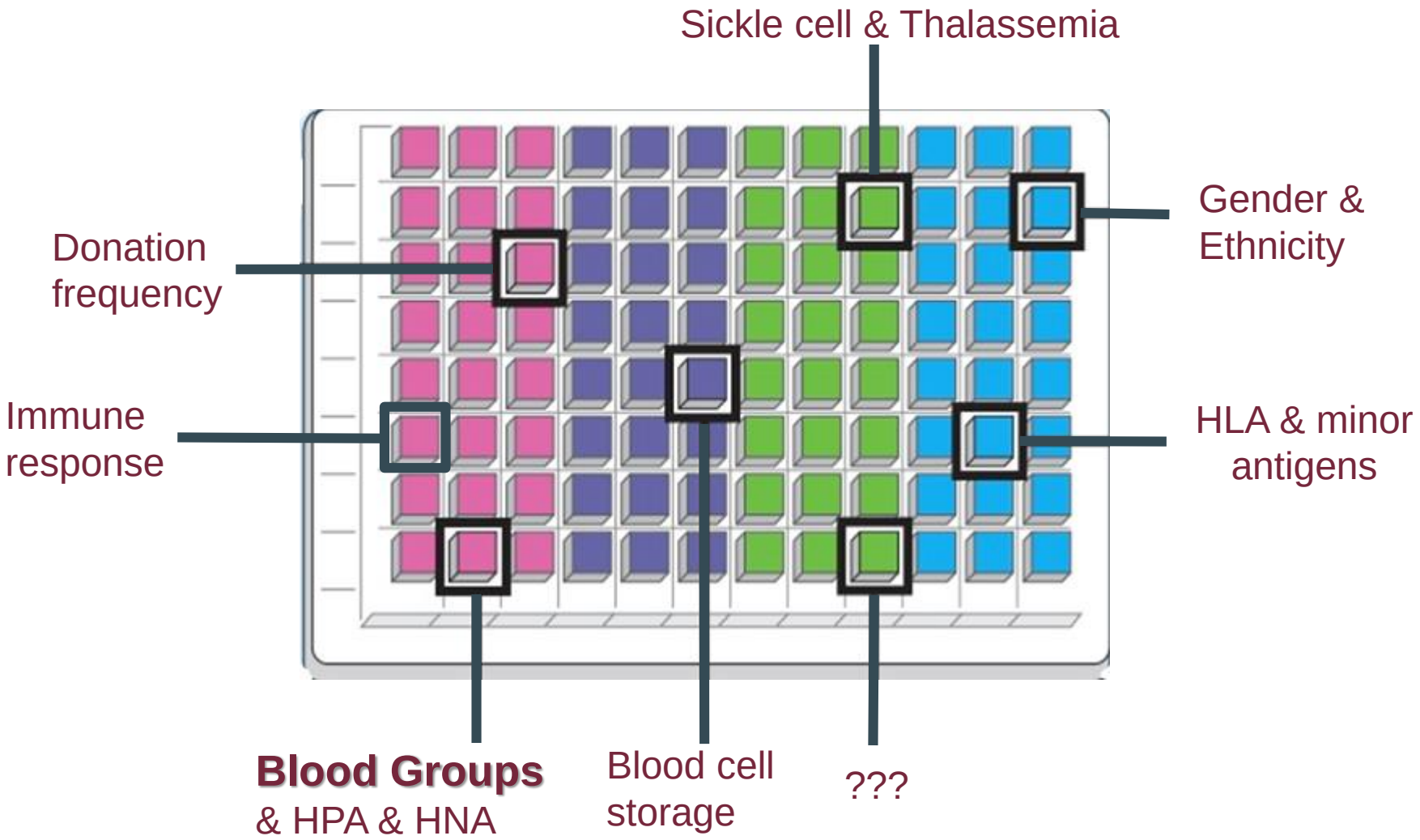
Blood for patient (A and D positive) for regular transfusions:
Coa negative (0.5%) and E negative (71%) donor needed

Blood bank: no donors readily available
Alternative treatment: bone marrow transplantation

Genotyped DISIII cohort:
7 donors of 2682 (0.3%): CoBB and ee
5 donors of compatible with A blood group

Conclusion: without genotyped cohort at least 1000 A and O donors have to be screened serologically for Co(a) antigen to obtain 3 donors to fulfill this request

SNV array not only for blood group typing



HLA typing based on SNV-arrays

- Imputing classical alleles from linked SNP genotype data

Locus	Number of validated	Call rate (2-digit)	Accuracy (2-digit)
HLA-A	816	0.98	0.98
HLA-B	1009	0.98	0.98
HLA-C	635	0.98	0.97
HLA-DRB1	858	0.99	0.98
HLA-DQA1	51	1	0.98
HLA-DQB1	867	1	0.99

Gil McVean, Bioinformatics, 2011, 27, 968-972

=> Algorithm still has to be improved, adding more SNVs from the HLA locus

- Targeted selection of platelet donors:
- Targeted selection of stem cell donors:
  - Possibillity to identify potential rare haplotypes, that are still missing in donor registries

(NB ethical approval)

Conclusions (1)

- Integrated platform for RBC grouping, HLA, HPA and HNA typing will be available by 2019, **<30 euro/assay**

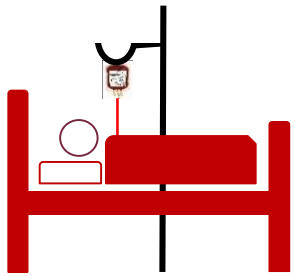
- This platform might allow to get a fully typed blood donor cohort

1) **Identification of rare donors**

2) Preventive matching: 3-5% of all transfusion episodes result in alloimmunization

   - **Availability of fully matched donors?**

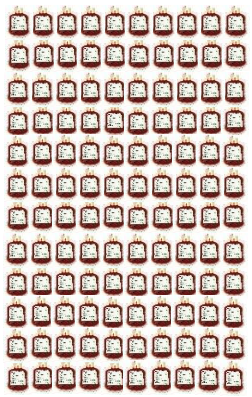
Sanquin



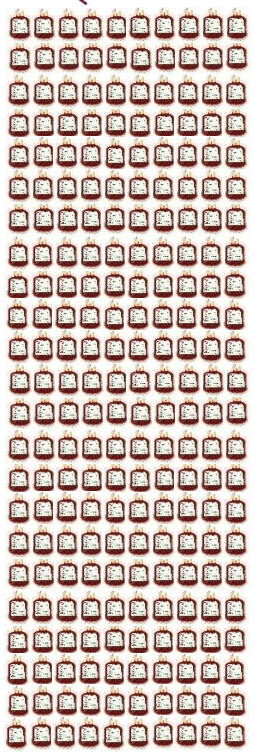
Can a request for a fully matched unit be directly fulfilled from the inventory?

Joost van Sambeeck

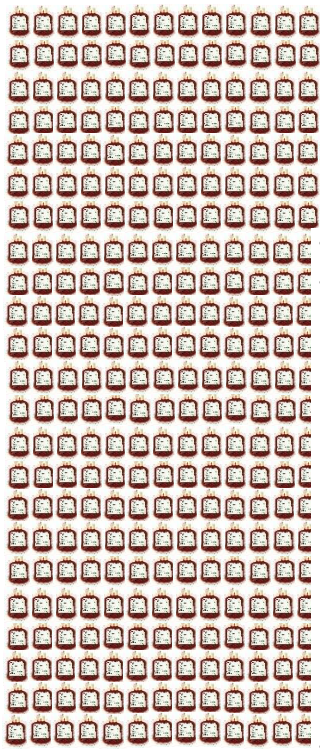
n=60 (small hospital)



n=120 (large hospital)



n=250 (university hosp)



n=1000 (distribution center)

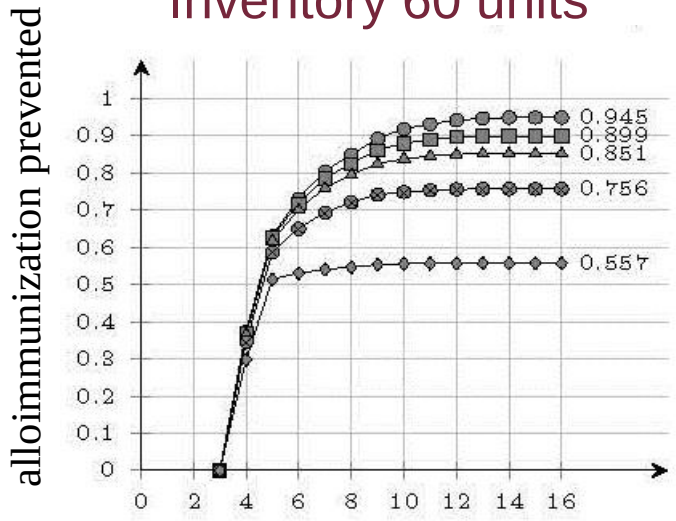
Algorithm 1: Binary search algorithm to find p

Input : n, k, α
Output: p

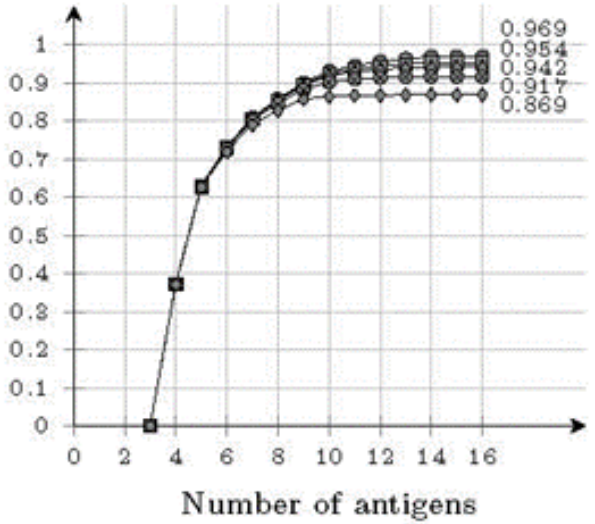
Let $X \sim \text{Bin}(n, p)$ and $F(k) = \mathbb{P}[X \leq k]$;
Define $p_{\text{low}} = 0$, $p_{\text{high}} = 1$, $p = \frac{p_{\text{low}} + p_{\text{high}}}{2}$;
while $|F(k-1) - (1-\alpha)| \geq 10^{-6}$ **do**
  if $F(k-1) < 1 - \alpha$ **then**
    $p_{\text{high}} = p$;
  else
    $p_{\text{low}} = p$;
  end
  $p = \frac{p_{\text{low}} + p_{\text{high}}}{2}$;
end

Availability of fully genotyped donor cohort makes preventive matching feasible

Small hospital
Inventory 60 units



Distribution center
Inventory 1000 units



1 unit
2 units
3 units
5 units
10 units

Static model
Only high risk patients

Dynamic model
All patients

Conclusions / Summary

- If all donors and transfusion recipients are fully typed, **extensive preventive matching** for all transfusion recipients **is feasible**

- Alloimmunization prevented:

		number of units requested (k)		
		1	**2**	**3**
Inventory size (n)	**60**	78%	68%	60%
	120	85%	79%	74%
	250	91%	87%	84%
	1000	97%	95%	94%

- Optimal order: (transfusion recipients **typed for a limited number of antigens**)

*	1	2	3	4	5	6	7	8	9	10	11	12	13
ABD	E	K	Jk^a	c	C	Fy^a	e	M	S	Jk^b	Fy^b	s	k

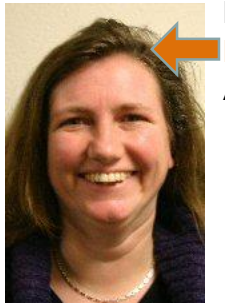
Final concusions

- Availability of SNV arrays optimized for comprehensive blood grouping and extended with more donor information will enable
  - to obtain fully typed donor cohorts, including HLA and HPA
  - to identify rare donors
  - to perform ‘precision donation’

- This will make preventive matching possible at a larger scale
  - Logistics of blood supply might need to change:
    - Larger distribution centers from which blood for elective (planned) transfusions can be supplied
    - Algorithms to supply blood based on phenotypes

  - Patients should be typed more comprehensively
    - Serology for most immunogenic antigens
    - Extraction of data from WGS or WES data
    - (Recognition of high responders)

BloodMatch project team



Dept. Exp.Immunohematology
Barbera Veldhuisen
Ahmad Javadi
Henk Schonewille
Ellen van der Schoot

Dept. Donor Studies
Joost van Sambeeck
Mart Janssen

Blood Bank
Rianne Koopman
Marian van Kraaij
Jessie Luken

Clinical Transfusion Research
Dorotea Evers
Jaap jan Zwaginga
Anske van der Bom
Masja de Haas

Blood Transfusion Genomics Consortium

UNIVERSITY OF CAMBRIDGE

Adam Butterworth
Nick Gleadall
Willem H Ouwehand
Karyn Megy
Luigi Grassi
Ernest Turro
Christopher Penkett
Kathy Stirrups

UNIVERSITY OF OXFORD

David Roberts
Mike Murphy
Gil McVean

NHS Blood and Transplant

Nick Watkins
Alan Grey
Shane Grimsley

Connie Westhoff
William Lane

UCL

Steven Marsh

Sanquin

Barbera Veldhuisen
Ellen van der Schoot

LUND UNIVERSITY

Mattias Moller
Jill Storry
Martin L Olsson

affymetrix

Laurent Bellon
Jeremy Gollub
Claire Bloore

Cambridge University Hospitals NHS Foundation Trust

Kim Brügger

LRG

Aoife McMahon

Development of blood group genotyping array

- Selection of SNPs in blood group genes : **n = 1,372**
  - Based on ISBT allele tables
  - + SNPs found in data from over 230,000 individuals (Africa Variant Project, 100,000 Genomes Project, gnomAD and TopMED)

- SNPs in 60 genes relevant to RBC, HP and HN antigen expression

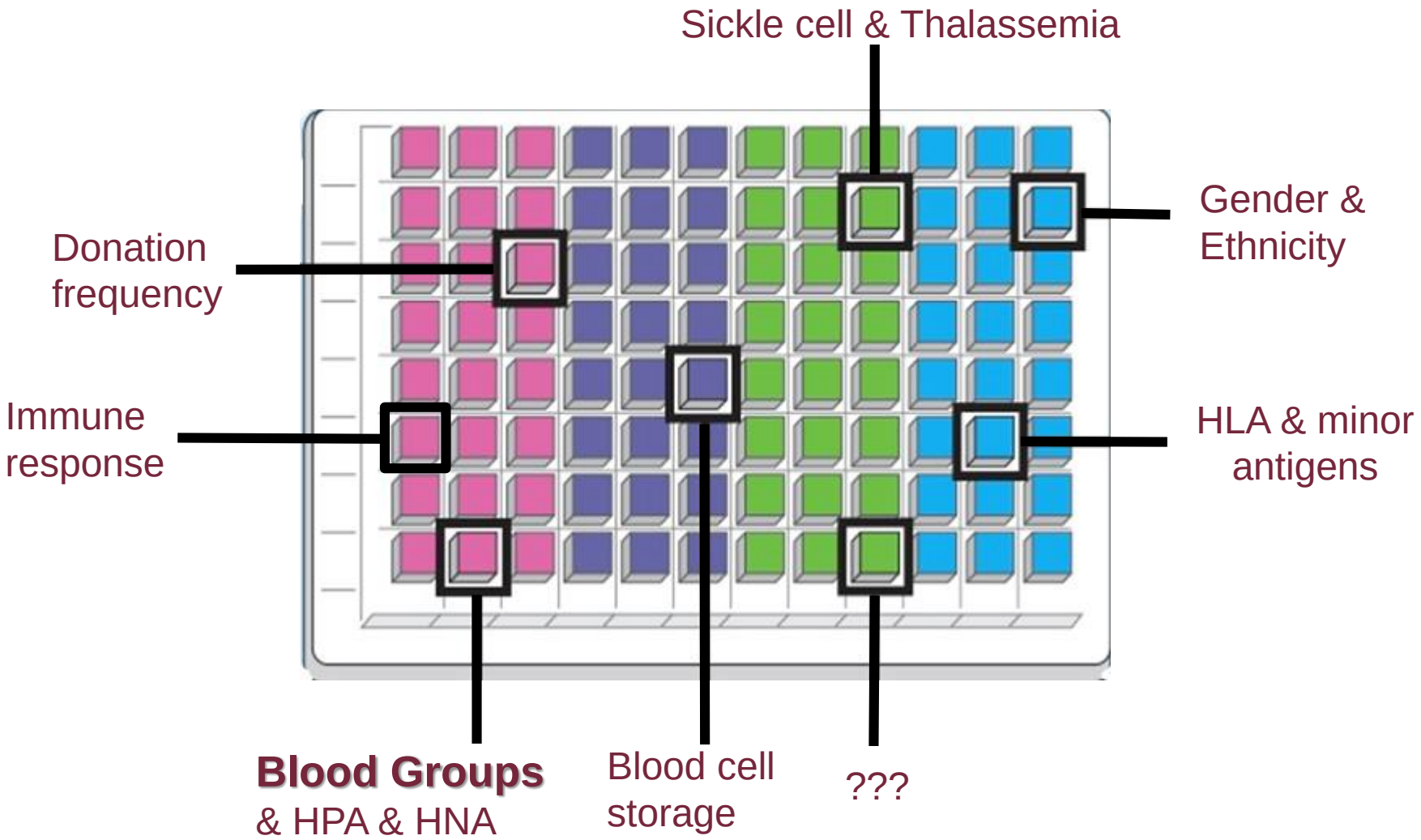
Beyond blood group antigens

- SNPs in HLA system

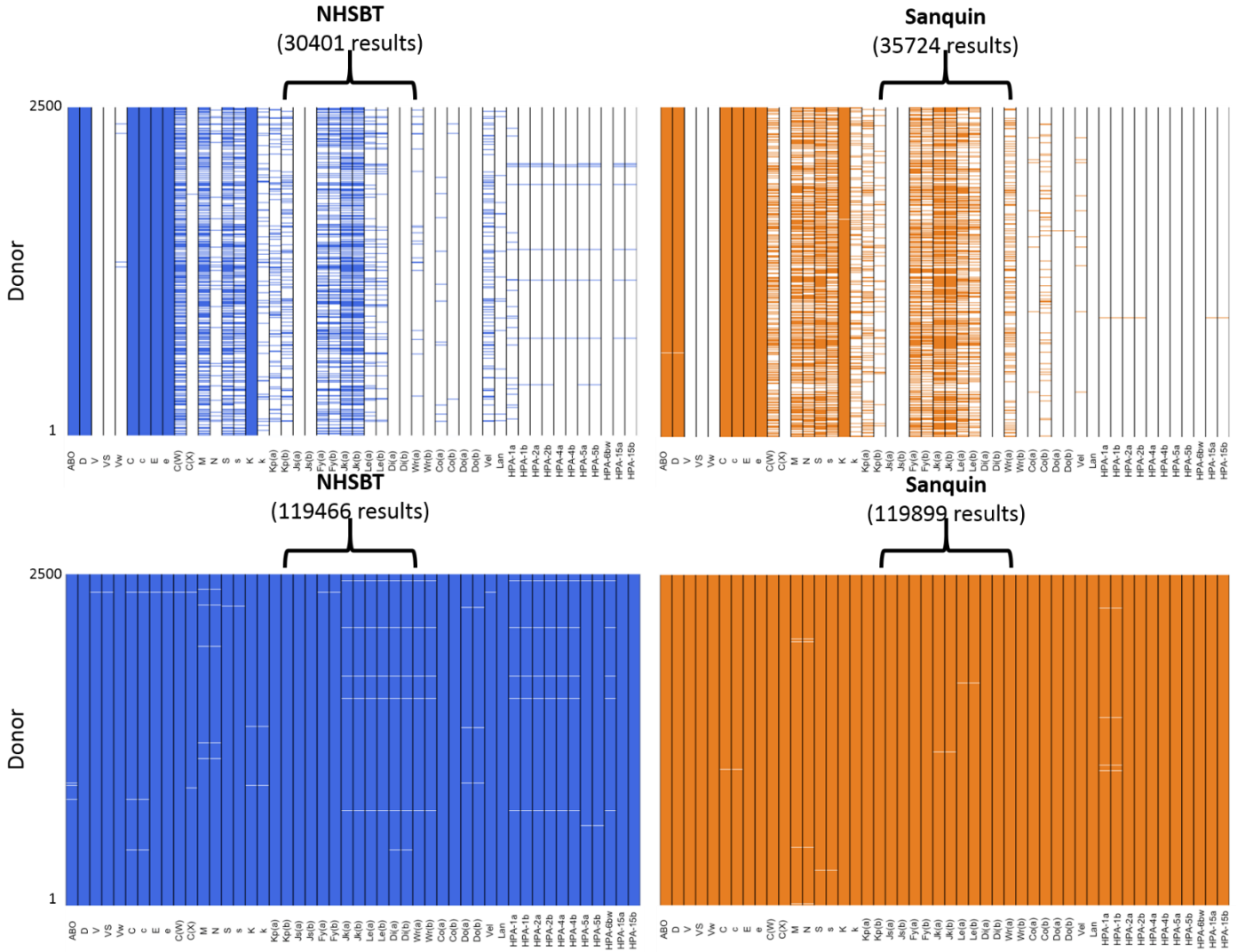
- SNPs in Human Platelet Antigen (HPA) systems, and in Human Neutrophil Antigen (HNA) systems

- SNPs in 304 genes found to be relevant for donor health

SNV array not only for blood group typing



Comparison serology and genotyping in 5000 donors



Serology:
66125 typing results
13,23/ donor

Genotyping:
239365 typing results
47,87/donor
+
224 extra / donors

20 m

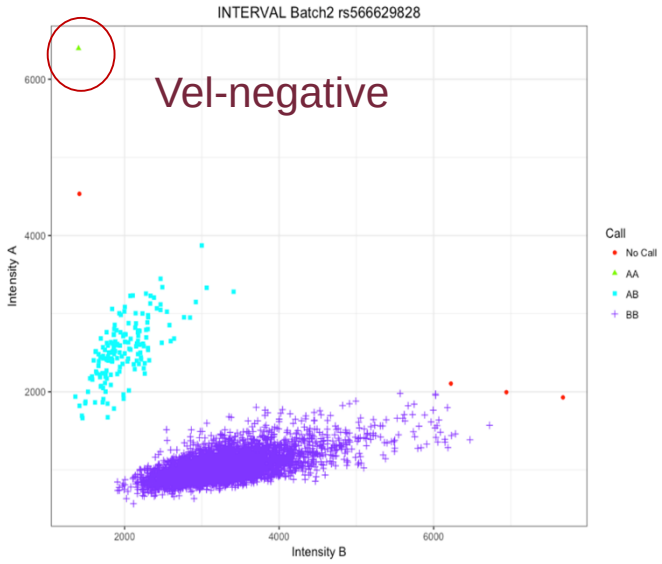
Blood Transfusion Genomics

99.9% concordancy between genotype and serology

Antigen	comparisons	discrepant	concordancy	Antigen	comparisons	discrepant	concordancy
ABO	4772	6	99.85	N	288	1	99.65
D	4787	3	99.94	S	2476	3	99.88
c	4780	6	99.87	s	1919	1	99.95
C	4780	4	99.92	Lu(b)	884	0	100.00
e	4784	1	99.98	Kp(b)	854	0	100.00
E	4784	0	100.00	Kp(a)	755	1	99.87
C(W)	2854	0	100.00	Le(a)	709	7	99.01
Fy(a)	1795	2	99.89	Vel	678	0	100.00
Fy(b)	1610	14	99.13	Lu(a)	637	0	100.00
K	4761	0	100.00	Le(b)	501	17	96.61
k	498	1	99.80	HPA-1a	250	0	100.00
Jk(a)	2743	1	99.96	Wr(a)	197	0	100.00
Jk(b)	2734	0	100.00	Lan	149	0	100.00
M	2669	2	99.93	Co(a)	115	0	100.00
				Total	**59676**	**76**	**99.87**

Identification of rare donors

- in 50,000 interval samples:
  - VEL- negative : 3
  - k- (KK) : 74
  - Js(b)-: 6
  - U- 1
  - Di(a)- 1
  - LAN- 4



HLA typing based on SNV-arrays

- Imputing classical alleles from linked SNP genotype data

Locus	Number of validated	Call rate (2-digit)	Accuracy (2-digit)
HLA-A	816	0.98	0.98
HLA-B	1009	0.98	0.98
HLA-C	635	0.98	0.97
HLA-DRB1	858	0.99	0.98
HLA-DQA1	51	1	0.98
HLA-DQB1	867	1	0.99

Gil McVean, Bioinformatics, 2011, 27, 968-972

=> Algorithm still has to be improved, adding more SNVs from the HLA locus

➢Targeted selection of platelet donors:

➢Targeted selection of stem cell donors:

- Possibillity to identify potential rare haplotypes, that are still missing in donor registries (NB ethical approval)

Conclusions (1)

- Integrated platform for RBC grouping, HLA, HPA and HNA typing will be available by 2019, **<30 euro/assay**

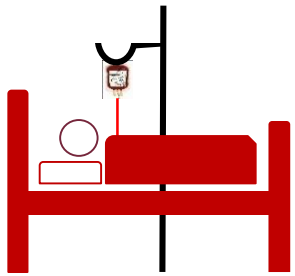
- This platform might allow to get a fully typed blood donor cohort

1) **Identification of rare donors**

2) Preventive matching: 3-5% of all transfusion episodes result in alloimmunization

   - **Availability of fully matched donors?**

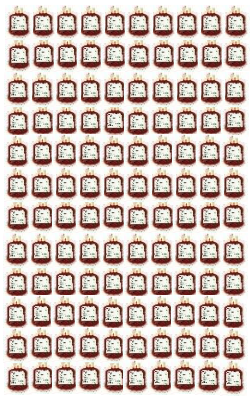
Sanquin



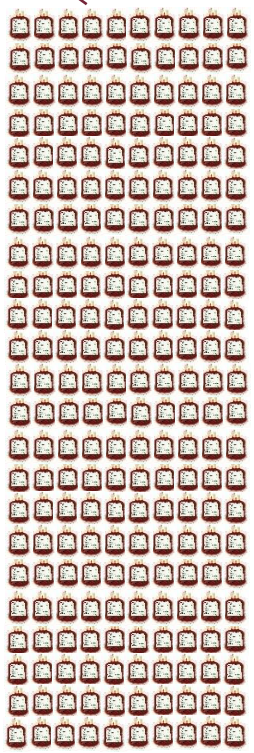
Can a request for a fully matched unit be directly fulfilled from the inventory?

Joost van Sambeeck

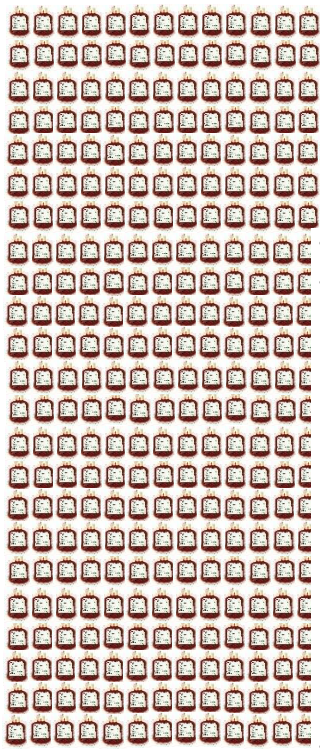
n=60 (small hospital)



n=120 (large hospital)



n=250 (university hosp)



n=1000 (distribution center)

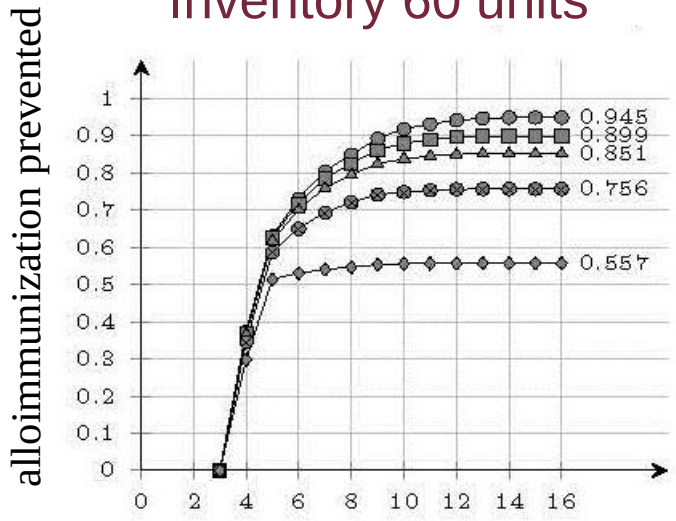
Algorithm 1: Binary search algorithm to find p

Input : n, k, α
Output: p

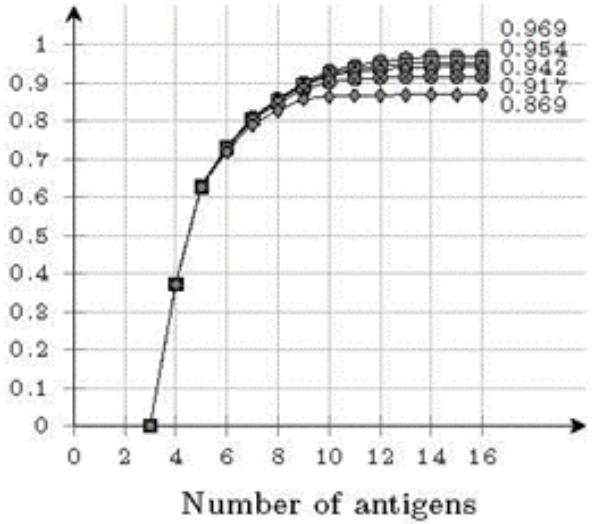
Let $X \sim \text{Bin}(n, p)$ and $F(k) = \mathbb{P}[X \leq k]$;
Define $p_{\text{low}} = 0$, $p_{\text{high}} = 1$, $p = \frac{p_{\text{low}} + p_{\text{high}}}{2}$;
while $|F(k-1) - (1-\alpha)| \geq 10^{-6}$ **do**
  if $F(k-1) < 1 - \alpha$ **then**
    $p_{\text{high}} = p$;
  else
    $p_{\text{low}} = p$;
  end
  $p = \frac{p_{\text{low}} + p_{\text{high}}}{2}$;
end

Availability of fully genotyped donor cohort makes preventive matching feasible

Small hospital
Inventory 60 units



Distribution center
Inventory 1000 units



1 unit
2 units
3 units
5 units
10 units

Static model
Only high risk patients

Dynamic model
All patients

Conclusions / Summary

- If all donors and transfusion recipients are fully typed, **extensive preventive matching** for all transfusion recipients **is feasible**

- Alloimmunization prevented:

		number of units requested (k)		
		1	**2**	**3**
Inventory size (n)	**60**	78%	68%	60%
	120	85%	79%	74%
	250	91%	87%	84%
	1000	97%	95%	94%

- Optimal order: (transfusion recipients **typed for a limited number of antigens**)

*	1	2	3	4	5	6	7	8	9	10	11	12	13
ABD	E	K	Jk^a	c	C	Fy^a	e	M	S	Jk^b	Fy^b	s	k

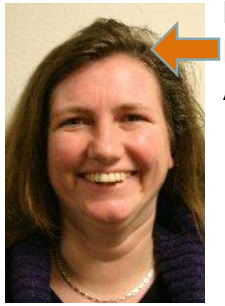
Final concusions

- Availability of SNV arrays optimized for comprehensive blood grouping and extended with more donor information will enable
  - to obtain fully typed donor cohorts
  - to identify rare donors
  - to perform ‘precision donation’

- This will make preventive matching possible at a larger scale
  - Logistics of blood supply might need to change:
    - Larger distribution centers from which blood for elective (planned) transfusions can be supplied
    - Algorithms to supply blood based on phenotypes

  - Patients should be typed more comprehensively
    - Serology for most immunogenic antigens
    - Extraction of data from WGS or WES data
    - (Recognition of high responders)

BloodMatch project team



Dept. Exp.Immunohematology
Barbera Veldhuisen
Ahmad Javadi
Henk Schonewille
Ellen van der Schoot

Dept. Donor Studies
Joost van Sambeeck
Mart Janssen

Blood Bank
Rianne Koopman
Marian van Kraaij
Jessie Luken

Clinical Transfusion Research
Dorotea Evers
Jaap jan Zwaginga
Anske van der Bom
Masja de Haas

Blood Transfusion Genomics Consortium

UNIVERSITY OF CAMBRIDGE

Adam Butterworth
Nick Gleadall
Willem H Ouwehand
Karyn Megy
Luigi Grassi
Ernest Turro
Christopher Penkett
Kathy Stirrups

UNIVERSITY OF OXFORD

David Roberts
Mike Murphy
Gil McVean

NHS Blood and Transplant

Nick Watkins
Alan Grey
Shane Grimsley

Connie Westhoff
William Lane

UCL

Steven Marsh

Sanquin

Barbera Veldhuisen
Ellen van der Schoot

LUND UNIVERSITY

Mattias Moller
Jill Storry
Martin L Olsson

affymetrix

Laurent Bellon
Jeremy Gollub
Claire Bloore

Cambridge University Hospitals NHS Foundation Trust

Kim Brügger

LRG

Aoife McMahon