A multiplex whole-gene HLA typing strategy evaluation on various NGS platforms: both long reads (PacBio, MinION), and short reads (Illumina, Ion Torrent, MGI).



Sake van Wageningen, Pascal van der Weele, Bart Valkenburg, Helma de Bruin, Erik Rozemuller, Maarten Penning

GenDx, Utrecht, The Netherlands

Aim

To investigate which NGS platforms available on the market are suitable for high resolution HLA typing, we performed a platform comparison study (Figure 1). Different NGS Libraries were generated from the same amplicons for the following platforms: Illumina Miseq, ThermoFisher Ion S5, PacBio Sequel IIe, MGI DNBseq-G50 and ONT MinION Mk1B. All studies were performed by GenDx, independently of the NGS platform manufacturers.

Figure 2: Setup of the Study

- HLA diverse 58 sample panel
- Multiplexed amplification; NGSgo-MX11-3

HLA-A	UTR	2	3	4	5 6	7 8 UTR	Amplified exp
HLA-B	UR	2	3	4	5 6	7 UTR	→ NGSgo-N
HLA C	UR 1	2	- 3	- 4	5 6	7 8 018	
DR81	UTL 1		- 2	- 3	-4	5 6 <u>yn</u>	
0Q81	UR 1	•	2	3	-4	5 6 <u>yr</u> e	1
DPB1	UIR 1	← →	z	3	4	S UTR	K./
OPA1			2	3		4 UTR	X
DQA1	UR 1		- 2 -	3		4 UTR	
DR83	UR 1	→	- 2 -	3	-4	8 6 98	
DR84	UTR 1		2	3		S 6 UTR	
OR85	UIII 1	← →	- 2 -	- 3	4	5 6 <u>yr</u>	



Methods

HLA amplicons were generated from a 58 sample HLA diversity panel (GeT-RM, HLA58) from the Coriell institute (Figure 2). Using an 11-gene multiplex amplification, whole gene amplicons were generated for HLA-A, -B, -C, -DRB1, -DQB1, -DPA1, -DPB1, - DQA, -DRB3, -DRB5. For HLA-DRB4 a single amplicon containing exons 2 and 3 was generated (NGSgo®-MX11-3; GenDx). Amplicon pools were used as input for library preparation for the different NGS sequencing platforms. All data analysis was performed using NGSengine software (GenDx).

Results and conclusions

The Illumina Miseq (Fig 3A) and PacBio Sequel IIe (Fig 3B) generated HLA typing results that were fully concordant with the pre-type information. Moreover, the Sequel IIe data analysis in all cases yielded fully phased amplicons. ThermoFisher Ion S5 (Fig 3C) and MGI DNBseq-G50 (Fig 3D) sequencers also generated highly concordant data (98% and 97%). Analysis of some loci in some samples required extra attention and re-analysis with adjusted analysis settings. Ion S5 data quality decreased in loci with homopolymers and in certain regions of DRB loci with high homology to other DRB genes. DNBseq-G50 data showed high sequencing noise at a few specific positions. This artefact will be further investigated in the future. ONT MinION Mk1B (Fig 3E) showed high concordance for class I genes, but data analysis of class II genes was hampered in several cases by high background noise in the data. Higher noise is innate to ONT sequencing, but this effect is exacerbated by homopolymers and STR regions in intronic HLA regions. Newer ONT sequencing chemistry compared to the one used here (R10.3) will be tested in the future for further optimization.

Figure 1: Platforms used in this study Disclaimer; All studies were performed by GenDx, independently of the NGS platform manufacturers.



Figure 3: NGSengine data overview

Locus	Mep	Region	UD	210	Noise	65N	65A	3494	1
NA172	12-510-7	08 [9/9]	1						
BAA	98%	Cores	1572	19 21	445	475	45N 485	1	(CIII) A102-01-0101, A168-01-02-02
HAD	975	Core +	1746	29 40	425	175	45N 48N	-	(ZEC) \$140db0200,8144d200d0 Ref. Bell
HAC	975	Core-	1062	21 29	3.7% 3.7%	475	45N 49%	1	(CI)E C103444101, C105410142
0581	945	Core-	401	17 18	525 535	39% 39%	475 475	1	CEE DIB1/1342-0102, DIB1/14-5401401 048
0882	97%	Care-	227 227	*	2.15	38% 37%	46N 45N	1	CCC 0483'62620143,0483'03456601
20283	97%	Care- Denni	1501 545	17 30	1.0%	425	40N 45N	8	CIE DOBYESIONER, DOBPOSIANER
DGA1	95%	Care+ Dren+	451 340	1	2.6%	455	425 455	-	DOAT-OLIZZIBOA, DOAT-OLIZABOI
0981	97%	Care+ Den+	722 722	76 25	276	45	425	-	CIE DERIVATION DERIVATION
DEA1	97%	Care+	1091	0	105			1	DENTOLOGICO, DENTOLOGICO

B) Sequel IIe 100% concordance (3-field pre-type) Lacue Map Region USD H4P Note 2DN ISA NM 7

1868	995	Core+ Dram+	1197	19 21	11.5%	25%	425	1	XIII AND INTRO, ASSUMPTING
1444	995	Core+ Dram+	9635 9635	29	1125	25%	45	1	[[]]] 0'46/1420(0'4442060 ^{but} , but
HAS	995	Cove+ Don+	622	21	17.2%	22% 22%	475 475	1	EIR COMMOND, COSMOND
a 2881	92%	Cure+ Drat+	418 374	17	1255	22%	65	1	IT OF PRODUCT DEPTH SHOLD THE
a <u>1993</u> a	995	Cure+ Crat+	290	-	645	205	405	1	TT ORBITUCOUNDI ORBITUCIONI
a post	98%	Cure+ Gran+	142	17	435	225	45	1	(III) ogeresolatiot, ogeresolation
a prost	99%	Cure+	342 311	;	1.8%	475	-	1	TEL DOM CHEROLOGI, DOM THOUSAND
a 2001	98%	Core +	63	20	625	59%	-	1	INTEL OPPERATION OF OPPERATION OF
		Cores	560	0	085				

C) Ion S5 98% concordance (3-field pre-type)

AD lice	05,NA	7222 [9/	111						
a <u>BAA</u>	975	Care- Exon-	422 251	18	92N 93N	15N 35N	42% 43%	8	[[]] Arezonatos, Areantozoz
-	975	Cave Extre	182 362	28	7.8% 7.2%	34%	475	8	(III) 8*430182261 8*66822600 Bed, Bud
анас	97%	Care Lon-	143	21	10.4%	31% 21%	ersi ersi	-	RIN CONSISTER CONSISTER
A DEE1	SEN.	Corr.	14	17	158N 124N	27% 27%	45N 48N	8	[E][] 08871342-9101 088714-540141 0883
il pras	90%	Core-	73		11.5N 11.8N	31N 13N	40% 43%	8	[[]]] 0483182420103, 0483103010101
iii 0081	97%	Care-	581 511	17	125N 128N	25N 25N	405 475	3	IZIZI DOBI'95/83/8101, DOBI'96/64/8101
(1 DOM)	05%	Core-	547 314	1	122N 122N	10N 25N	425	2	TTE DOATOLO20184, DOATOLO48101
1890 E	295	Cov-	299	16	105	29%	405	2	(E)E) DPERMITENDED, DPERMITEND
10 DEN1	90%	Care-	1080	1	2.0%			1	DRAFELOLOISE2, DRAFELOLOISE2

D) DNBSEQ-G50 97% concordance (3-field pre-type)

	and the	range of	000				1.00		
NA17222									
A HEAR	205	Corr-	4505	19 21	7.05 8.4%	37% 26%	43%	÷	III AND GEORGE, ANNOUNCE OF
HLAS	27%	Cores	5313	29 40	675 675	37% 11%	47%	1	KIKI \$1400300001, \$144620101 Ref., Bel
HAC	205	Corr+ Augu+	2615	21 29	635	22% 22%	43%	\$	III C133444101, C135410142
A 0885	95	Cert+ Aq1+	1332 264	17	106	345	48%	:	
A (2883	985	Cerro Apple	455 455	216	175	375	45% 41%	8	[1][] DR831024024143, DR831034014541
0081	98%	Carro Light	291.5 2309	17 30	5.75 5.75	37% 26%	48%	8	[[]]] DQ8765030141, DQ8706444501
10041	965	Care Lon-	997 302	1	6.7% 7.0%	47% 29%	42%	8	TE DOATONOLOGIO DOATONOLOGIO
a cent	985	Carro Don-	1302	18 20	6.9%	29% 32%	42%	-	INTE DEPRIVATION DEPRIVATION
a <u>1994 s</u>	985	Cave Done	2004 716	0 0	2025			1	DENTOLOGICO, DENTOLOGICO

E) Minlon; 8 samples processed Class I loci fully concordant Class II loci (DRB1, DQB1) challenging

100.00	- and	Page 1	-		11000		144		
NA0723	2,baro	de06 (11	(713.)						
анаа	845	Care-	9053	19	28.2%	12%	45%		TTT AND/ONCOLO AND/ONCO
		2001+	4181	21	28,2%	12%	45%	•	
		Conte	10001	14	De alto	100	47%		- Ind B
a HV-3	- 85%	Europa.	10005	- 43	30.25	69.	47%	- 2	[2][2] \$*43.0162.01 \$*44.62.0101
		10011	100.00		10.11				
		Care-	1149	21	25.2%	175	47%		10000 and a set of a set of a set of a
HAS	16.2	Dign+	1142	29	21.9%	125	47%		CIN CANOFEEDI CAROEEEDS
	100	Core-	3757	35	18.7%	-10%	22%	10	(will press and press (we)
AT AL		Di01+	3737	45	18.75	-10%	25%	15	CHEAD CHEAT COLOCUL, CHEAT PEPHI
					-	-			
0023	725	Cores	200	- 2	20.0%	12	445	- 5	[x]C] D883*8202/8103.D883*03/81010
		0054	20/2	18	20.0%	-/%	4479		
0 CE24 STN	_	Owne	25		24.05				
	375	from a			34.00		1700	- 21	TETHI DARM. GEOJOPEE DARM. GEORGO
			_						
	1000	Cover-	53	14	55.9%	5%	45%		TATUT DESCRIPTION OF ON ON OWNERS AND AN
16.00	20%	Enin*	53	21	35.9%	276	44%		STEL DISS WORKED AS RECENT
		-			-				
0.081	81%	Carte	4198		410.110				RTC1 DOB1/25/03/01/01 DOB1/06/04/01/0
		2001+	1883	32	661%	-14	49%		
		Care-	2069		16.8%	225	47%		
iii <u>0041</u> 75%	75%	Europa.	1007	- 9	04.16	105	48%	- 2	DQAYEE620664, DQAYEE66608
		-							
il <u>CERI</u> 75%	-	Care-	5674	16	12,7%	24%	47%		COLOR AND AND AN AD AD AD ADD AD AD AD
	100	£101+	\$633	22	21.0%	195	47%		Date of the second of the best of the
0091	TEN	COVE-	3087	- 2	13.0%				DOMPERATORIA DOMPSO AND
IN NUMBER OF STREET		Cunna 4	1004		12.26				

Legend figure 3:

hite (no marking): high quality Ilow: intermediate quality hk: low quality

est Read Depth (LRD) terozygous positions (HP) elta Signal to Noise (ΔSN) timated Second Allele% (ESA) ismatches (MM)

GENDX