

NGS Typing Analysis Software Fails to Identify HLA-A*24:11N Allele due to Presence of Homopolymer Sequence

Anthony Spinelli¹, Min-Jeong Lee¹, Robert Vorhaben¹, Chantale Lacelle²

Clinical Laboratory Services¹ and Department of Pathology², University of Texas Southwestern Medical Center, Dallas, Texas

INTRODUCTION

HLA-A*24:11N differs from HLA-A*24:02 by a single nucleotide insertion at the beginning of exon 4 causing a frameshift mutation resulting in the introduction of a premature stop codon. This results in the lack of surface expression of HLA-A protein encoded by this allele. The A*24:11N has 7 consecutive cytosine while the A*24:02 allele has one less. The HLA-A*24:11N is a well-defined allele in European/European decent and intermediate allele in Asian/Pacific Islander populations based on the CWD v3.0 classification. While performing validation of the One Lambda AllType[™] FASTplex NGS assay we intentionally used a sample that was heterozygous for HLA-A*24:02, 24:11N to assess the performance of the NGS assay in identifying alleles with homopolymer sequence.

FIGURE 4: TypeStream Visual Analysis



METHODS

DNA EXTRACTION AND HLA TYPING: DNA was extracted using the EZ1 (Qiagen) automated DNA extractor and quantified using a Nanodrop (ThermoFisher). Samples were tested using three typing methods: 1. SSO (ThermoFisher-One Lambda). 2. SBT using lab developed reagents and analyzed using the SBT engine software (GenDX, Netherlands), and 3. NGS using One Lambda AllType[™] FASTplex NGS assay analyzed with TypeStream Visual.

RESULTS FIGURE 1: Sequence Alignment Comparing A*24:02:01:01 and A*24:11N Exon 4 **Result:** Analysis of a reference samples that is A*24:02, 24:11N using TypeStream Visual fails to identify potential A*24:11N allele. HLA-A*24:11N has an additional cytosine compared to A*24:02.

FIGURE 5: TypeStream Visual with Manual addition of the A*24:11N Reference Sequence



AA Codon	185	190	195	200	205
A*24:02:01:01	AC CCC CCC .AAG ACA CA	AT ATG ACC CAC CAC CCC A	ATC TCT GAC CAT GAG GCC	C ACT CTG AGA TGC TG	G GCC CTG GGC
A*24:11N	C				
AA Codon	210	215	220	225	230
A*24:02:01:01	TTC TAC CCT GCG GAG AT	C ACA CTG ACC TGG CAG CO	GG GAT GGG GAG GAC CAG	ACC CAG GAC ACG GAG	G CTT GTG GAG
A*24:11N					
AA Codon	235	240	245	250	255
A*24:02:01:01	ACC AGG CCT GCA GGG GA	E GGA ACC TTC CAG AAG TO	GG GCA GCT GTG GTG GTA	CCT TCT GGA GAG GAG	G CAG AGA TAC
A*24:11N					
AA Codon	260	265	270		
A*24:02:01:01	ACC TGC CAT GTG CAG CAT	F GAG GGT CTG CCC AAG CO	CC CTC ACC CTG AGA TGG	G	
A*24:11N				_	

Result: HLA-A*24:02:01:01 differs from A*24:11N by one nucleotide insertion in exon 4. A*24:02:01:01 has 6 consecutive cytosine while A*24:11N has 7 consecutive cytosine resulting in the introduction of a premature stop codon.

FIGURE 2: Protein Sequence Comparing A*24:02:01:01 and A*24:11N

AA Pos.	10	20	30	40	50	60	70	80	90	100
A*24:02:01:01	GSHSMRYFST	SVSRPGRGEP	RFIAVGYVDD	TQFVRFDSDA	ASQRMEPRAP	WIEQEGPEYW	DEETGKVKAH	SQTDRENLRI	ALRYYNQSEA	GSHTLQMMFG
A*24:11N										
AA Pos.	110	120	130	140	150	160	170	180	190	200
A*24:02:01:01	CDVGSDGRFL	RGYHQYAYDG	KDYIALKEDL	RSWTAADMAA	QITKRKWEAA	HVAEQQRAYL	EGTCVDGLRR	YLENGKETLQ	RTDPPKTHMT	HHPISDHEAT
A*24:11N									QDTYD	PP-HLX
AA Pos.	210	220	230	240) 250	260) 270	280	290	300
A*24:02:01:01	LRCWALGFYP	AEITLTWQRD	GED.QTQDTEI	L VETRPAGDG	FQKWAAVVVI	SGEEQRYTCH	I VQHEGLPKPI	L TLRWEPSSQF	, TVPIVGIIAG	, LVLLGAVITG
A*24:11N										
AA Pos.	310	320	330	340						
A*24:02:01:01	AVVAAVMWRR	NSSDRKGGSY	SQAASSDSAQ	GSDVSLTACK	V					
A*24:11N										



As a majority of laboratories are now using next generation sequencing it becomes particularly important to identify scenarios where software may underperform and warrant manual review of sequences. One should be extremely vigilant to manually review all homopolymer sequences to ensure proper alleles are identified since sequences which differ in a homopolymer region may be missed. While for our patient this would have had little clinical impact since the typing missed a null allele, this may be of great clinical significance in instances where two distinct expressed allele are present. Our data indicates that selection of samples for test and software validation is extremely important and that manual review of NGS data by qualified HLA staff is essential.

Result: HLA-A*24:11N has a premature stop codon resulting in truncated protein sequence.