GENDX

personalising diagnostics

Improved SBTengine® batch analysis module.

Introduction

HLA typing by Sanger Sequencing is a high-resolution typing method which depend on the correct analysis of the sequence traces. Using a combination of Core sequencing and Group Specific Sequencing reactions (GSSPs) results in allele level typing in the vast majority of samples. However, the reliability of the results depend on the sequence quality and the accuracy of the manual inspection and editing.

SBTengine® is a software package which optimizes the SBT analysis workflow. CWD lists and Genotype expected frequencies are applied to focus on the crucial steps of the analysis. GSSP advisor and GPS predictor are tools used to minimize the work to come to the resolution required. However, an accurate manual inspection and editing still are required.

Method

To reduce the manual inspection and editing work, we improved the basic analysis algorithms in SBTengine®. The analysis includes an extensive evaluation of the sequence quality at every position in the context of the other traces of a sample. This means that the results are based on the combined quality of all traces. The improved algorithm result in a huge reduction in manual editing steps after opening of the sample (Table 1).

Results

To further facilitate efficient HLA typing of large numbers of samples, we developed a module for batch analysis, which analyzes Sanger SBT data to come to an HLA typing result. The module also optimizes the result in the context of the expected frequencies of the genotyping results (Figure 1).

Upon opening a folder containing sequence data, SBTengine® automatically analysis the data and returns a typing result. In the upper right corner of the screen, information of typing success is summarized.

The quality of the traces directly influences the possibilities of batch analysis, which is a balance between acceptance and rejection of some data based on quality criteria. One can expect that not all data can be analyzed automatically. The results of the batch analysis depend on the basic quality of the traces. With good quality data, in the majority of cases the correct genotype is identified, which makes the module applicable for registry typing. A small percentage of samples may not come to the correct genotype, although it is expected that this concerns only minor subtypes. For donor-recipient typing, the module can perform the first analysis, which always has to be followed by an intensive evaluation by the lab supervisor (Table 2).

Table 1: Manual edits in 19 samples.

Manual typing edits in SBTengine version 2161, containing the improved analysis algorithm, was compared to the number of manual edits with the original analysis algorithm (SBTengine version 2014). The table clearly shows that the number of edits is reduced dramatically.

SBTengine® version	2014	2161
Sample	# of manual edits	
1	23	0
2	4	2
3	16	2
4	9	0
5	14	1
6	15	0
8	17	0
9	0	0
10	9	0
11	11	0
12	4	3
13	2	0
14	3	0
15	17	1
16	8	2
17	0	0
18	6	2
19	0	0
20	24	0

Figure 1. SBTengine® batch analysis in action.

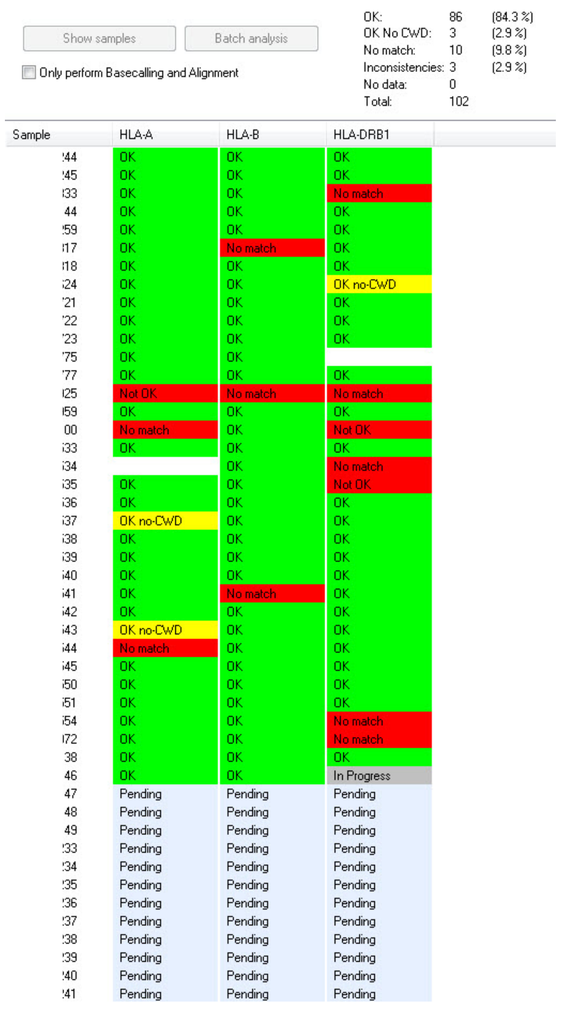

Table 2. SBTengine® batch analysis results.

In this example, 277 sample / loci were analyzed using the batch analysis module. In 10 minutes, 86% samples were typed automatically. As a control, all samples were also typed manually. Typing concordance between manual analysis and the batch analysis was 99%.

In 16 cases (6%), there were still sequencing inconsistencies that could not be resolved automatically 15 of these concerned DRB1. In 23 instances (8%), no inconsistencies between sequences and IMGT were left, but no matching genotype was found. Also in this case, most of these concerned DRB1 (14 from 23).

	HLA-A		HLA-B		HLA-DRB1		Total	
	Count	%	Count	%	Count	%	Count	%
Time/sample	3,3		1,5		1,5		2,3	
Typing	88	95,7%	87	93,5%	63	68,5%	238	85,9%
No match	3	3,3%	6	6,5%	14	15,2%	23	8,3%
Inconsistencies	1	1,1%	0	0,0%	15	16,3%	16	5,8%
	92		93		92		277	

Conclusions

The improved analysis algorithm reduces the number of necessary manual edits.

The batch analysis leads to a further reduction of editing, and offers a reliable method for increased efficiency of Sanger SBT data analysis.

Key Features SBTengine®

- Clear and comprehensive overview
- Homozygous and heterozygous sequence traces are analysed simultaneously
- GSSP Prediction System (GPS) predicts possible ambiguities
- Quality filtering

Erik H. Rozemuller, Job Geerligs, Wietse Mulder, Maarten T. Penning

GenDx, Yalelaan 48, 3584 CM Utrecht, the Netherlands