Validation Study Using HLA Data of the British 1958 Cohort Study

Aug 01, 2012

Table 1: The sensitivity (SEN), specificity (SPE), positive predictive value (PPV) and negative predictive value (NPV) for each four-digit HLA allele without call threshold. GSK data of European ancestry were used to build the prediction models, and HLA data from the British 1958 Cohort Study were used as validation samples. The SNP markers in the models are common to the Illumina Human1M, Omni, OmniExpress, 660K and 550K platforms.

	Allele	Num. Train	Freq. Train	Num. Valid.	Freq. Valid.	CR[1] (%)	ACC[1] (%)	SEN (%)	SPE (%)	PPV (%)	NPV (%)	Miscall[2] (%)
HLA–A (# of training samples: 1857, # of validation samples: 884)												
Overall accuracy: 98.1%												
1	01:01	541	0.1457	487	0.2755	100.0	99.6	99.6	99.6	98.9	99.8	
2	01:02	1	0.0003	0	–	–	–	–	–	–	–	–
3	01:25	1	0.0003	0	–	–	–	–	–	–	–	–
4	02:01	1036	0.2789	0	–	–	–	–	–	–	–	–
5	02:02	6	0.0016	2	0.0011	100.0	100.0	100.0	100.0	100.0	100.0	–
6	02:05	38	0.0102	30	0.0170	100.0	99.9	100.0	99.9	98.4	100.0	–
7	02:06	11	0.0030	2	0.0011	100.0	100.0	100.0	100.0	100.0	100.0	–
8	02:17	1	0.0003	0	–	–	–	–	–	–	–	–
9	02:20	1	0.0003	0	–	–	–	–	–	–	–	–
10	02:22	1	0.0003	0	–	–	–	–	–	–	–	–
11	02:35	1	0.0003	0	–	–	–	–	–	–	–	–
12	03:01	486	0.1309	354	0.2002	100.0	99.8	99.4	99.9	99.7	99.9	01:01 (100)
13	03:02	5	0.0013	1	0.0006	100.0	99.9	0.0	100.0	–	99.9	03:01 (100)
14	03:17	1	0.0003	0	–	–	–	–	–	–	–	–
15	11:01	247	0.0665	154	0.0871	100.0	100.0	100.0	100.0	100.0	100.0	–
16	11:02	1	0.0003	0	–	–	–	–	–	–	–	–
17	11:25	1	0.0003	0	–	–	–	–	–	–	–	–
18	23:01	81	0.0218	38	0.0215	100.0	99.9	97.4	100.0	100.0	99.9	24:02 (50)
19	24:02	311	0.0837	178	0.1007	100.0	99.7	99.4	99.7	98.1	99.9	02:05 (50)
20	24:03	10	0.0027	3	0.0017	100.0	99.8	0.0	100.0	–	99.8	24:02 (100)
21	24:32	1	0.0003	0	–	–	–	–	–	–	–	–
22	25:01	105	0.0283	45	0.0255	100.0	99.6	93.3	99.8	91.3	99.8	26:01 (100)

Continued on next page ...

Table 1 – continued from previous page

	Allele	Num. Train	Freq. Train	Num. Valid.	Freq. Valid.	CR[1] (%)	ACC[1] (%)	SEN (%)	SPE (%)	PPV (%)	NPV (%)	Miscall[2] (%)
23	26:01	144	0.0388	51	0.0288	100.0	99.2	92.2	99.4	82.5	99.8	25:01 (100)
24	26:08	1	0.0003	5	0.0028	100.0	99.7	0.0	100.0	–	99.7	26:01 (100)
25	29:01	20	0.0054	0	–	–	–	–	–	–	–	–
26	29:02	120	0.0323	94	0.0532	100.0	99.9	97.9	100.0	100.0	99.9	32:01 (50)
27	29:10	1	0.0003	0	–	–	–	–	–	–	–	–
28	30:01	57	0.0153	32	0.0181	100.0	100.0	100.0	100.0	100.0	100.0	–
29	30:02	36	0.0097	23	0.0130	100.0	99.8	87.0	100.0	100.0	99.8	01:01 (67)
30	30:04	3	0.0008	1	0.0006	100.0	100.0	100.0	100.0	100.0	100.0	–
31	31:01	91	0.0245	71	0.0402	100.0	100.0	100.0	100.0	100.0	100.0	–
32	31:08	1	0.0003	0	–	–	–	–	–	–	–	–
33	32:01	135	0.0363	93	0.0526	100.0	99.8	100.0	99.8	96.9	100.0	–
34	33:01	25	0.0067	0	–	–	–	–	–	–	–	–
35	33:03	18	0.0048	1	0.0006	100.0	100.0	100.0	100.0	100.0	100.0	–
36	33:05	1	0.0003	0	–	–	–	–	–	–	–	–
37	34:02	3	0.0008	2	0.0011	100.0	100.0	100.0	100.0	100.0	100.0	–
38	36:01	1	0.0003	0	–	–	–	–	–	–	–	–
39	66:01	23	0.0062	8	0.0045	100.0	99.9	87.5	100.0	100.0	99.9	26:01 (100)
40	66:02	1	0.0003	0	–	–	–	–	–	–	–	–
41	68:01	112	0.0302	83	0.0469	100.0	99.8	96.4	99.9	99.4	99.8	01:01 (33)
42	68:02	24	0.0065	8	0.0045	100.0	100.0	100.0	100.0	100.0	100.0	–
43	68:24	1	0.0003	0	–	–	–	–	–	–	–	–
44	68:35	1	0.0003	0	–	–	–	–	–	–	–	–
45	69:01	4	0.0011	0	–	–	–	–	–	–	–	–
46	74:01	1	0.0003	0	–	–	–	–	–	–	–	–
47	74:03	2	0.0005	0	–	–	–	–	–	–	–	–
48	74:06	–	–	1	0.0006	100.0	99.9	0.0	100.0	–	99.9	32:01 (100)
49	74:10	–	–	1	0.0006	100.0	99.9	0.0	100.0	–	99.9	32:01 (100)
50	80:01	1	0.0003	0	–	–	–	–	–	–	–	–
HLA–B (# of training samples: 2572, # of validation samples: 1532)												
Overall accuracy: 96.9%												
1	07:02	610	0.1186	452	0.1475	100.0	99.5	98.0	99.7	98.3	99.7	08:01 (22)
2	07:04	2	0.0004	0	–	–	–	–	–	–	–	–
3	07:05	15	0.0029	0	–	–	–	–	–	–	–	–
4	07:10	1	0.0002	0	–	–	–	–	–	–	–	–
5	07:17	1	0.0002	0	–	–	–	–	–	–	–	–
6	08:01	501	0.0974	448	0.1462	100.0	99.6	98.9	99.8	98.7	99.8	07:02 (30)
7	13:02	172	0.0334	56	0.0183	100.0	99.9	96.4	99.9	96.4	99.9	07:02 (50)
8	14:01	34	0.0066	45	0.0147	100.0	100.0	100.0	100.0	100.0	100.0	–
9	14:02	121	0.0235	84	0.0274	100.0	99.9	98.8	100.0	98.8	100.0	39:01 (100)
10	15:01	307	0.0597	217	0.0708	100.0	99.6	98.2	99.7	96.2	99.9	35:01 (25)
11	15:02	2	0.0004	0	–	–	–	–	–	–	–	–
12	15:03	4	0.0008	5	0.0016	100.0	100.0	100.0	100.0	100.0	100.0	–
13	15:04	1	0.0002	0	–	–	–	–	–	–	–	–
14	15:07	1	0.0002	1	0.0003	100.0	100.0	0.0	100.0	–	100.0	15:01 (100)
15	15:08	2	0.0004	0	–	–	–	–	–	–	–	–
16	15:09	1	0.0002	0	–	–	–	–	–	–	–	–

Continued on next page ...

Table 1 – continued from previous page

	Allele	Num. Train	Freq. Train	Num. Valid.	Freq. Valid.	CR[1] (%)	ACC[1] (%)	SEN (%)	SPE (%)	PPV (%)	NPV (%)	Miscall[2] (%)
17	15:10	4	0.0008	0	–	–	–	–	–	–	–	–
18	15:15	–	–	1	0.0003	100.0	100.0	0.0	100.0	–	100.0	15:01 (100)
19	15:16	4	0.0008	3	0.0010	100.0	100.0	100.0	100.0	100.0	100.0	–
20	15:17	20	0.0039	4	0.0013	100.0	100.0	100.0	100.0	100.0	100.0	–
21	15:18	9	0.0017	12	0.0039	100.0	99.9	83.3	100.0	100.0	99.9	15:01 (100)
22	15:24	–	–	1	0.0003	100.0	100.0	0.0	100.0	–	100.0	15:01 (100)
23	18:01	292	0.0568	102	0.0333	100.0	99.8	98.0	99.9	96.6	99.9	51:01 (50)
24	18:02	1	0.0002	0	–	–	–	–	–	–	–	–
25	18:03	12	0.0023	0	–	–	–	–	–	–	–	–
26	18:11	1	0.0002	0	–	–	–	–	–	–	–	–
27	18:18	1	0.0002	0	–	–	–	–	–	–	–	–
28	27:01	–	–	1	0.0003	100.0	100.0	0.0	100.0	–	100.0	27:05 (100)
29	27:02	47	0.0091	1	0.0003	100.0	100.0	100.0	100.0	100.0	100.0	–
30	27:03	1	0.0002	0	–	–	–	–	–	–	–	–
31	27:05	193	0.0375	138	0.0450	100.0	99.8	99.3	99.8	96.8	100.0	35:01 (100)
32	27:07	1	0.0002	0	–	–	–	–	–	–	–	–
33	27:09	1	0.0002	0	–	–	–	–	–	–	–	–
34	35:01	274	0.0533	110	0.0359	100.0	99.4	93.6	99.7	90.7	99.8	35:08 (43)
35	35:02	48	0.0093	8	0.0026	100.0	99.9	100.0	99.9	72.7	100.0	–
36	35:03	115	0.0224	16	0.0052	100.0	99.8	81.2	99.9	86.7	99.9	35:01 (100)
37	35:08	37	0.0072	11	0.0036	100.0	99.8	72.7	99.9	72.7	99.9	35:02 (67)
38	35:09	1	0.0002	0	–	–	–	–	–	–	–	–
39	35:17	2	0.0004	0	–	–	–	–	–	–	–	–
40	35:41	1	0.0002	0	–	–	–	–	–	–	–	–
41	35:55	1	0.0002	0	–	–	–	–	–	–	–	–
42	37:01	54	0.0105	49	0.0160	100.0	99.9	95.9	99.9	96.9	99.9	08:01 (25)
43	38:01	144	0.0280	27	0.0088	100.0	99.9	96.3	99.9	89.7	100.0	18:01 (100)
44	39:01	73	0.0142	32	0.0104	100.0	99.9	90.6	100.0	96.7	99.9	38:01 (67)
45	39:06	23	0.0045	25	0.0082	100.0	100.0	96.0	100.0	100.0	100.0	07:02 (50)
46	39:10	1	0.0002	0	–	–	–	–	–	–	–	–
47	39:24	3	0.0006	0	–	–	–	–	–	–	–	–
48	39:29	1	0.0002	0	–	–	–	–	–	–	–	–
49	40:01	203	0.0395	211	0.0689	100.0	99.7	96.7	99.9	98.8	99.8	08:01 (29)
50	40:02	89	0.0173	22	0.0072	100.0	100.0	100.0	100.0	95.7	100.0	–
51	40:06	6	0.0012	0	–	–	–	–	–	–	–	–
52	40:10	1	0.0002	0	–	–	–	–	–	–	–	–
53	40:27	1	0.0002	0	–	–	–	–	–	–	–	–
54	40:32	1	0.0002	0	–	–	–	–	–	–	–	–
55	41:01	23	0.0045	9	0.0029	100.0	100.0	88.9	100.0	100.0	100.0	40:01 (100)
56	41:02	32	0.0062	9	0.0029	100.0	100.0	100.0	100.0	100.0	100.0	–
57	42:01	1	0.0002	0	–	–	–	–	–	–	–	–
58	42:02	1	0.0002	0	–	–	–	–	–	–	–	–
59	44:02	438	0.0851	332	0.1084	100.0	99.5	97.0	99.8	98.3	99.6	44:05 (30)
60	44:03	252	0.0490	158	0.0516	100.0	99.6	97.5	99.8	95.7	99.9	07:02 (38)
61	44:04	2	0.0004	3	0.0010	100.0	99.9	0.0	100.0	–	99.9	44:03 (100)
62	44:05	29	0.0056	1	0.0003	100.0	99.9	100.0	99.9	25.0	100.0	–

Continued on next page ...

Table 1 – continued from previous page

	Allele	Num. Train	Freq. Train	Num. Valid.	Freq. Valid.	CR[1] (%)	ACC[1] (%)	SEN (%)	SPE (%)	PPV (%)	NPV (%)	Miscall[2] (%)
63	44:06	1	0.0002	0	–	–	–	–	–	–	–	–
64	44:08	1	0.0002	0	–	–	–	–	–	–	–	–
65	44:27	1	0.0002	0	–	–	–	–	–	–	–	–
66	45:01	25	0.0049	26	0.0085	100.0	100.0	96.2	100.0	100.0	100.0	15:01 (50)
67	46:01	1	0.0002	0	–	–	–	–	–	–	–	–
68	47:01	14	0.0027	15	0.0049	100.0	100.0	93.3	100.0	100.0	100.0	07:02 (100)
69	47:02	1	0.0002	0	–	–	–	–	–	–	–	–
70	48:01	5	0.0010	0	–	–	–	–	–	–	–	–
71	49:01	88	0.0171	35	0.0114	100.0	99.9	91.4	100.0	97.0	99.9	44:02 (33)
72	50:01	72	0.0140	32	0.0104	100.0	99.8	90.6	99.9	93.5	99.9	44:03 (33)
73	50:02	1	0.0002	0	–	–	–	–	–	–	–	–
74	51:01	276	0.0537	119	0.0388	100.0	99.7	98.3	99.8	93.6	99.9	07:02 (25)
75	51:05	3	0.0006	0	–	–	–	–	–	–	–	–
76	51:07	5	0.0010	0	–	–	–	–	–	–	–	–
77	51:08	4	0.0008	1	0.0003	100.0	100.0	0.0	100.0	–	100.0	51:01 (100)
78	51:09	1	0.0002	0	–	–	–	–	–	–	–	–
79	52:01	65	0.0126	20	0.0065	100.0	99.9	90.0	100.0	94.7	99.9	51:01 (100)
80	53:01	13	0.0025	6	0.0020	100.0	99.9	83.3	100.0	83.3	100.0	18:01 (100)
81	54:01	3	0.0006	0	–	–	–	–	–	–	–	–
82	55:01	86	0.0167	61	0.0199	100.0	100.0	100.0	100.0	98.4	100.0	–
83	55:02	3	0.0006	0	–	–	–	–	–	–	–	–
84	55:05	1	0.0002	0	–	–	–	–	–	–	–	–
85	56:01	46	0.0089	11	0.0036	100.0	99.9	90.9	100.0	95.2	100.0	55:01 (100)
86	57:01	153	0.0297	131	0.0428	100.0	99.8	98.5	99.9	97.4	99.9	07:02 (25)
87	57:02	–	–	1	0.0003	100.0	100.0	0.0	100.0	–	100.0	57:01 (100)
88	57:03	1	0.0002	0	–	–	–	–	–	–	–	–
89	58:01	47	0.0091	12	0.0039	100.0	100.0	91.7	100.0	100.0	100.0	35:01 (100)
90	70:20	2	0.0004	0	–	–	–	–	–	–	–	–
91	73:01	5	0.0010	0	–	–	–	–	–	–	–	–
92	81:01	1	0.0002	0	–	–	–	–	–	–	–	–
HLA–C (# of training samples: 1866, # of validation samples: 840)												
Overall accuracy: 96.5%												
1	01:02	157	0.0421	89	0.0530	100.0	99.8	96.6	99.9	98.9	99.8	02:02 (33)
2	02:02	208	0.0557	74	0.0440	100.0	99.5	95.9	99.7	94.0	99.8	03:04 (100)
3	02:10	2	0.0005	0	–	–	–	–	–	–	–	–
4	03:02	8	0.0021	1	0.0006	100.0	100.0	100.0	100.0	100.0	100.0	–
5	03:03	182	0.0488	108	0.0643	100.0	98.8	94.4	99.1	88.3	99.6	07:01 (50)
6	03:04	219	0.0587	196	0.1167	100.0	98.7	90.8	99.7	97.5	98.8	03:03 (67)
7	03:05	2	0.0005	0	–	–	–	–	–	–	–	–
8	03:06	1	0.0003	0	–	–	–	–	–	–	–	–
9	03:10	1	0.0003	0	–	–	–	–	–	–	–	–
10	04:01	427	0.1144	167	0.0994	100.0	99.8	99.4	99.8	98.2	99.9	16:01 (100)
11	04:03	1	0.0003	0	–	–	–	–	–	–	–	–
12	04:09	1	0.0003	0	–	–	–	–	–	–	–	–
13	05:01	300	0.0804	210	0.1250	100.0	99.4	98.1	99.6	97.2	99.7	04:01 (38)
14	06:02	350	0.0938	207	0.1232	100.0	99.6	97.6	99.9	99.0	99.7	07:02 (60)

Continued on next page ...

Table 1 – continued from previous page

	Allele	Num. Train	Freq. Train	Num. Valid.	Freq. Valid.	CR[1] (%)	ACC[1] (%)	SEN (%)	SPE (%)	PPV (%)	NPV (%)	Miscall[2] (%)
15	07:01	551	0.1476	9	0.0054	100.0	99.6	100.0	99.6	60.0	100.0	–
16	07:02	485	0.1300	261	0.1554	100.0	99.3	98.9	99.4	95.5	99.8	02:02 (33)
17	07:04	75	0.0201	36	0.0214	100.0	99.8	94.4	99.9	94.4	99.9	02:02 (50)
18	08:01	3	0.0008	1	0.0006	100.0	100.0	100.0	100.0	100.0	100.0	–
19	08:02	97	0.0260	88	0.0524	100.0	99.6	93.2	100.0	100.0	99.6	05:01 (62)
20	08:03	3	0.0008	0	–	–	–	–	–	–	–	–
21	12:02	50	0.0134	14	0.0083	100.0	100.0	100.0	100.0	100.0	100.0	–
22	12:03	262	0.0702	63	0.0375	100.0	99.8	96.8	99.9	99.2	99.9	07:01 (25)
23	12:05	1	0.0003	0	–	–	–	–	–	–	–	–
24	12:13	1	0.0003	0	–	–	–	–	–	–	–	–
25	14:02	42	0.0113	19	0.0113	100.0	99.9	94.7	99.9	100.0	99.9	01:02 (50)
26	14:03	2	0.0005	0	–	–	–	–	–	–	–	–
27	14:04	1	0.0003	0	–	–	–	–	–	–	–	–
28	15:02	88	0.0236	40	0.0238	100.0	99.8	95.0	99.9	96.2	99.9	02:02 (50)
29	15:04	4	0.0011	0	–	–	–	–	–	–	–	–
30	15:05	16	0.0043	4	0.0024	100.0	99.9	75.0	100.0	100.0	99.9	15:02 (100)
31	15:11	1	0.0003	0	–	–	–	–	–	–	–	–
32	16:01	125	0.0335	74	0.0440	100.0	99.8	98.6	99.9	98.0	99.9	07:01 (50)
33	16:02	11	0.0029	6	0.0036	100.0	100.0	100.0	100.0	100.0	100.0	–
34	16:04	9	0.0024	0	–	–	–	–	–	–	–	–
35	17:01	40	0.0107	13	0.0077	100.0	100.0	100.0	100.0	100.0	100.0	–
36	17:03	5	0.0013	0	–	–	–	–	–	–	–	–
37	18:01	1	0.0003	0	–	–	–	–	–	–	–	–
HLA–DRB1 (# of training samples: 2436, # of validation samples: 1129)												
Overall accuracy: 92.2%												
1	01:01	406	0.0833	98	0.0434	100.0	98.6	100.0	98.5	73.8	100.0	–
2	01:02	68	0.0140	22	0.0097	100.0	100.0	95.5	100.0	100.0	100.0	01:01 (100)
3	01:03	36	0.0074	57	0.0252	100.0	98.8	50.9	100.0	100.0	98.7	01:01 (100)
4	01:04	–	–	1	0.0004	100.0	100.0	0.0	100.0	–	100.0	01:01 (100)
5	03:01	554	0.1137	429	0.1900	100.0	99.6	98.1	99.9	99.8	99.6	11:01 (50)
6	03:02	1	0.0002	0	–	–	–	–	–	–	–	–
7	04:01	377	0.0774	282	0.1249	100.0	97.6	92.9	98.3	87.5	99.0	04:07 (35)
8	04:02	57	0.0117	11	0.0049	100.0	99.7	72.7	99.9	80.0	99.9	04:04 (100)
9	04:03	56	0.0115	46	0.0204	100.0	98.0	8.7	99.9	60.6	98.1	04:01 (41)
10	04:04	138	0.0283	117	0.0518	100.0	98.5	85.5	99.2	86.2	99.2	04:01 (39)
11	04:05	36	0.0074	16	0.0071	100.0	99.8	87.5	99.9	87.5	99.9	04:01 (100)
12	04:06	2	0.0004	0	–	–	–	–	–	–	–	–
13	04:07	53	0.0109	23	0.0102	100.0	98.8	82.6	99.0	46.4	99.8	04:01 (50)
14	04:08	17	0.0035	0	–	–	–	–	–	–	–	–
15	04:09	1	0.0002	0	–	–	–	–	–	–	–	–
16	04:10	2	0.0004	0	–	–	–	–	–	–	–	–
17	04:59	1	0.0002	0	–	–	–	–	–	–	–	–
18	07:01	643	0.1320	421	0.1864	100.0	99.5	98.3	99.7	99.0	99.6	04:01 (57)
19	08:01	121	0.0248	42	0.0186	100.0	99.9	97.6	99.9	95.3	100.0	08:03 (100)
20	08:02	8	0.0016	0	–	–	–	–	–	–	–	–
21	08:03	7	0.0014	3	0.0013	100.0	99.9	66.7	100.0	66.7	100.0	

Continued on next page ...

Table 1 – continued from previous page

	Allele	Num. Train	Freq. Train	Num. Valid.	Freq. Valid.	CR[1] (%)	ACC[1] (%)	SEN (%)	SPE (%)	PPV (%)	NPV (%)	Miscall[2] (%)
22	08:04	9	0.0018	5	0.0022	100.0	99.9	60.0	100.0	100.0	99.9	08:01 (50)
23	08:10	2	0.0004	0	–	–	–	–	–	–	–	–
24	09:01	48	0.0099	38	0.0168	100.0	100.0	97.4	100.0	100.0	100.0	04:01 (100)
25	10:01	33	0.0068	13	0.0058	100.0	100.0	92.3	100.0	100.0	100.0	07:01 (50)
26	11:01	342	0.0702	72	0.0319	100.0	98.8	94.4	98.9	73.1	99.8	11:04 (50)
27	11:02	12	0.0025	11	0.0049	100.0	99.9	72.7	100.0	100.0	99.9	11:01 (33)
28	11:03	41	0.0084	10	0.0044	100.0	99.6	10.0	100.0	50.0	99.6	11:01 (100)
29	11:04	206	0.0423	24	0.0106	100.0	99.5	62.5	99.9	83.3	99.6	11:01 (78)
30	11:06	1	0.0002	0	–	–	–	–	–	–	–	–
31	11:08	1	0.0002	0	–	–	–	–	–	–	–	–
32	11:12	2	0.0004	0	–	–	–	–	–	–	–	–
33	11:15	1	0.0002	0	–	–	–	–	–	–	–	–
34	11:28	1	0.0002	0	–	–	–	–	–	–	–	–
35	11:39	1	0.0002	0	–	–	–	–	–	–	–	–
36	11:43	1	0.0002	0	–	–	–	–	–	–	–	–
37	12:01	74	0.0152	35	0.0155	100.0	99.8	94.3	99.9	91.7	99.9	04:01 (25)
38	12:02	3	0.0006	2	0.0009	100.0	99.9	0.0	100.0	–	99.9	12:01 (100)
39	13:01	312	0.0640	100	0.0443	100.0	99.8	98.0	99.9	98.0	99.9	11:01 (50)
40	13:02	176	0.0361	71	0.0314	100.0	99.9	98.6	99.9	97.2	100.0	07:01 (50)
41	13:03	61	0.0125	25	0.0111	100.0	99.9	92.0	100.0	100.0	99.9	08:01 (50)
42	13:05	7	0.0014	1	0.0004	100.0	100.0	0.0	100.0	–	100.0	11:01 (100)
43	13:10	1	0.0002	1	0.0004	100.0	100.0	0.0	100.0	–	100.0	13:01 (100)
44	13:15	1	0.0002	0	–	–	–	–	–	–	–	–
45	13:17	1	0.0002	0	–	–	–	–	–	–	–	–
46	13:18	1	0.0002	0	–	–	–	–	–	–	–	–
47	14:01	135	0.0277	69	0.0306	100.0	99.8	100.0	99.8	97.2	100.0	–
48	14:02	2	0.0004	0	–	–	–	–	–	–	–	–
49	14:04	7	0.0014	1	0.0004	100.0	100.0	0.0	100.0	–	100.0	14:01 (100)
50	14:06	2	0.0004	0	–	–	–	–	–	–	–	–
51	15:01	594	0.1219	184	0.0815	100.0	99.8	100.0	99.8	97.9	100.0	–
52	15:02	56	0.0115	10	0.0044	100.0	99.9	100.0	99.9	87.0	100.0	–
53	15:03	2	0.0004	0	–	–	–	–	–	–	–	–
54	16:01	136	0.0279	16	0.0071	100.0	99.9	93.8	100.0	96.8	100.0	16:02 (100)
55	16:02	14	0.0029	2	0.0009	100.0	100.0	100.0	100.0	66.7	100.0	–
56	16:05	2	0.0004	0	–	–	–	–	–	–	–	–
HLA–DQB1 (# of training samples: 1924, # of validation samples: 1004)												
Overall accuracy: 97.8%												
1	02:01	424	0.1102	27	0.0134	100.0	99.6	88.9	99.7	82.3	99.8	03:02 (67)
2	02:02	400	0.1040	24	0.0120	100.0	99.6	95.8	99.6	73.4	99.9	03:01 (100)
3	02:03	–	–	1	0.0005	100.0	100.0	0.0	100.0	–	100.0	02:02 (100)
4	02:04	1	0.0003	0	–	–	–	–	–	–	–	–
5	03:01	777	0.2019	508	0.2530	100.0	99.3	97.8	99.7	99.3	99.3	03:03 (45)
6	03:02	374	0.0972	236	0.1175	100.0	99.6	97.9	99.8	98.3	99.7	02:01 (53)
7	03:03	167	0.0434	148	0.0737	100.0	99.6	97.3	99.7	96.3	99.8	02:02 (50)
8	03:04	10	0.0026	1	0.0005	100.0	100.0	100.0	100.0	100.0	100.0	–
9	03:05	4	0.0010	1	0.0005	100.0	100.0	0.0	100.0	–	100.0	02:02 (50)

Continued on next page ...

Table 1 – continued from previous page

	Allele	Num. Train	Freq. Train	Num. Valid.	Freq. Valid.	CR[1] (%)	ACC[1] (%)	SEN (%)	SPE (%)	PPV (%)	NPV (%)	Miscall[2] (%)
10	03:19	4	0.0010	0	–	–	–	–	–	–	–	–
11	04:02	121	0.0314	61	0.0304	100.0	100.0	100.0	100.0	100.0	100.0	–
12	05:01	433	0.1125	314	0.1564	100.0	99.9	99.7	99.9	99.7	99.9	03:03 (50)
13	05:02	132	0.0343	20	0.0100	100.0	99.9	95.0	99.9	95.0	99.9	05:04 (100)
14	05:03	111	0.0288	65	0.0324	100.0	99.9	98.5	99.9	97.0	99.9	03:01 (50)
15	05:04	2	0.0005	3	0.0015	100.0	100.0	100.0	100.0	75.0	100.0	–
16	06:01	46	0.0120	15	0.0075	100.0	100.0	100.0	99.9	93.8	100.0	–
17	06:02	462	0.1201	340	0.1693	100.0	99.5	98.5	99.6	98.4	99.7	06:03 (40)
18	06:03	244	0.0634	141	0.0702	100.0	99.4	94.3	99.8	97.1	99.6	06:02 (62)
19	06:04	107	0.0278	79	0.0393	100.0	99.8	97.5	99.9	97.5	99.9	06:03 (100)
20	06:09	27	0.0070	24	0.0120	100.0	100.0	100.0	100.0	100.0	100.0	–
21	06:14	1	0.0003	0	–	–	–	–	–	–	–	–
22	06:16	1	0.0003	0	–	–	–	–	–	–	–	–

[1]: CR – call rate; ACC – allele accuracy.
[2]: the most likely miscalled allele and the proportion of the most likely miscalled allele in all miscalled alleles.

Table 2: The sensitivity (SEN), specificity (SPE), positive predictive value (PPV) and negative predictive value (NPV) for each four-digit HLA allele with call threshold 0.5. GSK data of European ancestry were used to build the prediction models, and HLA data from the British 1958 Cohort Study were used as validation samples. The SNP markers in the models are common to the Illumina Human1M, Omni, OmniExpress, 660K and 550K platforms.

	Allele	Num. Train	Freq. Train	Num. Valid.	Freq. Valid.	CR[1] (%)	ACC[1] (%)	SEN (%)	SPE (%)	PPV (%)	NPV (%)	Miscall[2] (%)
HLA–A (# of training samples: 1857, # of validation samples: 884)												
Overall accuracy: 98.2%, Call rate: 99.4%												
1	01:01	541	0.1457	487	0.2755	99.8	99.6	99.6	99.6	98.9	99.8	
2	01:02	1	0.0003	0	–	–	–	–	–	–	–	–
3	01:25	1	0.0003	0	–	–	–	–	–	–	–	–
4	02:01	1036	0.2789	0	–	–	–	–	–	–	–	–
5	02:02	6	0.0016	2	0.0011	100.0	100.0	100.0	100.0	100.0	100.0	–
6	02:05	38	0.0102	30	0.0170	100.0	99.9	100.0	99.9	98.4	100.0	–
7	02:06	11	0.0030	2	0.0011	100.0	100.0	100.0	100.0	100.0	100.0	–
8	02:17	1	0.0003	0	–	–	–	–	–	–	–	–
9	02:20	1	0.0003	0	–	–	–	–	–	–	–	–
10	02:22	1	0.0003	0	–	–	–	–	–	–	–	–
11	02:35	1	0.0003	0	–	–	–	–	–	–	–	–
12	03:01	486	0.1309	354	0.2002	99.7	99.8	99.4	99.9	99.7	99.9	01:01 (100)
13	03:02	5	0.0013	1	0.0006	100.0	99.9	0.0	100.0	–	99.9	03:01 (100)
14	03:17	1	0.0003	0	–	–	–	–	–	–	–	–
15	11:01	247	0.0665	154	0.0871	100.0	100.0	100.0	100.0	100.0	100.0	–
16	11:02	1	0.0003	0	–	–	–	–	–	–	–	–
17	11:25	1	0.0003	0	–	–	–	–	–	–	–	–
18	23:01	81	0.0218	38	0.0215	100.0	99.9	97.4	100.0	100.0	99.9	24:02 (50)
19	24:02	311	0.0837	178	0.1007	99.4	99.7	99.4	99.7	98.1	99.9	02:05 (50)
20	24:03	10	0.0027	3	0.0017	100.0	99.8	0.0	100.0	–	99.8	24:02 (100)
21	24:32	1	0.0003	0	–	–	–	–	–	–	–	–
22	25:01	105	0.0283	45	0.0255	95.6	99.6	93.0	99.8	90.9	99.8	26:01 (100)
23	26:01	144	0.0388	51	0.0288	98.0	99.2	92.0	99.4	82.1	99.8	25:01 (100)
24	26:08	1	0.0003	5	0.0028	100.0	99.7	0.0	100.0	–	99.7	26:01 (100)
25	29:01	20	0.0054	0	–	–	–	–	–	–	–	–
26	29:02	120	0.0323	94	0.0532	100.0	99.9	97.9	100.0	100.0	99.9	32:01 (50)
27	29:10	1	0.0003	0	–	–	–	–	–	–	–	–
28	30:01	57	0.0153	32	0.0181	100.0	100.0	100.0	100.0	100.0	100.0	–
29	30:02	36	0.0097	23	0.0130	95.7	99.8	86.4	100.0	100.0	99.8	01:01 (67)
30	30:04	3	0.0008	1	0.0006	100.0	100.0	100.0	100.0	100.0	100.0	–
31	31:01	91	0.0245	71	0.0402	100.0	100.0	100.0	100.0	100.0	100.0	–
32	31:08	1	0.0003	0	–	–	–	–	–	–	–	–
33	32:01	135	0.0363	93	0.0526	98.9	99.8	100.0	99.8	96.8	100.0	–
34	33:01	25	0.0067	0	–	–	–	–	–	–	–	–
35	33:03	18	0.0048	1	0.0006	100.0	100.0	100.0	100.0	100.0	100.0	–
36	33:05	1	0.0003	0	–	–	–	–	–	–	–	–
37	34:02	3	0.0008	2	0.0011	100.0	100.0	100.0	100.0	100.0	100.0	–
38	36:01	1	0.0003	0	–	–	–	–	–	–	–	–

Continued on next page ...

Table 2 – continued from previous page

	Allele	Num. Train	Freq. Train	Num. Valid.	Freq. Valid.	CR[1] (%)	ACC[1] (%)	SEN (%)	SPE (%)	PPV (%)	NPV (%)	Miscall[2] (%)
39	66:01	23	0.0062	8	0.0045	87.5	99.9	85.7	100.0	100.0	99.9	26:01 (100)
40	66:02	1	0.0003	0	–	–	–	–	–	–	–	–
41	68:01	112	0.0302	83	0.0469	98.8	99.8	97.6	99.9	99.4	99.9	01:01 (50)
42	68:02	24	0.0065	8	0.0045	100.0	100.0	100.0	100.0	100.0	100.0	–
43	68:24	1	0.0003	0	–	–	–	–	–	–	–	–
44	68:35	1	0.0003	0	–	–	–	–	–	–	–	–
45	69:01	4	0.0011	0	–	–	–	–	–	–	–	–
46	74:01	1	0.0003	0	–	–	–	–	–	–	–	–
47	74:03	2	0.0005	0	–	–	–	–	–	–	–	–
48	74:06	–	–	1	0.0006	100.0	99.9	0.0	100.0	–	99.9	32:01 (100)
49	74:10	–	–	1	0.0006	100.0	99.9	0.0	100.0	–	99.9	32:01 (100)
50	80:01	1	0.0003	0	–	–	–	–	–	–	–	–
HLA–B (# of training samples: 2572, # of validation samples: 1532)												
Overall accuracy: 97.4%, Call rate: 97.3%												
1	07:02	610	0.1186	452	0.1475	98.7	99.5	98.2	99.8	98.5	99.7	08:01 (25)
2	07:04	2	0.0004	0	–	–	–	–	–	–	–	–
3	07:05	15	0.0029	0	–	–	–	–	–	–	–	–
4	07:10	1	0.0002	0	–	–	–	–	–	–	–	–
5	07:17	1	0.0002	0	–	–	–	–	–	–	–	–
6	08:01	501	0.0974	448	0.1462	98.7	99.6	98.9	99.8	98.6	99.8	07:02 (30)
7	13:02	172	0.0334	56	0.0183	96.4	99.9	100.0	99.9	96.4	100.0	–
8	14:01	34	0.0066	45	0.0147	97.8	100.0	100.0	100.0	100.0	100.0	–
9	14:02	121	0.0235	84	0.0274	97.6	99.9	98.8	100.0	98.8	100.0	39:01 (100)
10	15:01	307	0.0597	217	0.0708	98.6	99.6	98.1	99.7	96.1	99.9	35:01 (25)
11	15:02	2	0.0004	0	–	–	–	–	–	–	–	–
12	15:03	4	0.0008	5	0.0016	60.0	100.0	100.0	100.0	100.0	100.0	–
13	15:04	1	0.0002	0	–	–	–	–	–	–	–	–
14	15:07	1	0.0002	1	0.0003	100.0	100.0	0.0	100.0	–	100.0	15:01 (100)
15	15:08	2	0.0004	0	–	–	–	–	–	–	–	–
16	15:09	1	0.0002	0	–	–	–	–	–	–	–	–
17	15:10	4	0.0008	0	–	–	–	–	–	–	–	–
18	15:15	–	–	1	0.0003	100.0	100.0	0.0	100.0	–	100.0	15:01 (100)
19	15:16	4	0.0008	3	0.0010	100.0	100.0	100.0	100.0	100.0	100.0	–
20	15:17	20	0.0039	4	0.0013	75.0	100.0	100.0	100.0	100.0	100.0	–
21	15:18	9	0.0017	12	0.0039	91.7	99.9	81.8	100.0	100.0	99.9	15:01 (100)
22	15:24	–	–	1	0.0003	100.0	100.0	0.0	100.0	–	100.0	15:01 (100)
23	18:01	292	0.0568	102	0.0333	96.1	99.9	99.0	99.9	97.5	100.0	15:01 (50)
24	18:02	1	0.0002	0	–	–	–	–	–	–	–	–
25	18:03	12	0.0023	0	–	–	–	–	–	–	–	–
26	18:11	1	0.0002	0	–	–	–	–	–	–	–	–
27	18:18	1	0.0002	0	–	–	–	–	–	–	–	–
28	27:01	–	–	1	0.0003	100.0	100.0	0.0	100.0	–	100.0	27:05 (100)
29	27:02	47	0.0091	1	0.0003	100.0	100.0	100.0	100.0	100.0	100.0	–
30	27:03	1	0.0002	0	–	–	–	–	–	–	–	–
31	27:05	193	0.0375	138	0.0450	97.8	99.8	99.3	99.8	96.8	100.0	35:01 (100)
32	27:07	1	0.0002	0	–	–	–	–	–	–	–	–

Continued on next page ...

Table 2 – continued from previous page

	Allele	Num. Train	Freq. Train	Num. Valid.	Freq. Valid.	CR[1] (%)	ACC[1] (%)	SEN (%)	SPE (%)	PPV (%)	NPV (%)	Miscall[2] (%)
33	27:09	1	0.0002	0	–	–	–	–	–	–	–	–
34	35:01	274	0.0533	110	0.0359	92.7	99.6	96.1	99.8	92.9	99.9	35:03 (50)
35	35:02	48	0.0093	8	0.0026	87.5	100.0	100.0	100.0	100.0	100.0	–
36	35:03	115	0.0224	16	0.0052	87.5	99.9	85.7	99.9	85.7	99.9	35:01 (100)
37	35:08	37	0.0072	11	0.0036	54.5	99.9	100.0	99.9	75.0	100.0	–
38	35:09	1	0.0002	0	–	–	–	–	–	–	–	–
39	35:17	2	0.0004	0	–	–	–	–	–	–	–	–
40	35:41	1	0.0002	0	–	–	–	–	–	–	–	–
41	35:55	1	0.0002	0	–	–	–	–	–	–	–	–
42	37:01	54	0.0105	49	0.0160	98.0	99.9	95.8	99.9	96.8	99.9	08:01 (25)
43	38:01	144	0.0280	27	0.0088	96.3	99.9	96.2	100.0	96.2	100.0	18:01 (100)
44	39:01	73	0.0142	32	0.0104	81.2	99.9	92.3	100.0	96.0	99.9	18:01 (50)
45	39:06	23	0.0045	25	0.0082	96.0	100.0	95.8	100.0	100.0	100.0	07:02 (50)
46	39:10	1	0.0002	0	–	–	–	–	–	–	–	–
47	39:24	3	0.0006	0	–	–	–	–	–	–	–	–
48	39:29	1	0.0002	0	–	–	–	–	–	–	–	–
49	40:01	203	0.0395	211	0.0689	97.6	99.7	96.6	99.9	98.8	99.8	08:01 (29)
50	40:02	89	0.0173	22	0.0072	86.4	100.0	100.0	100.0	95.0	100.0	–
51	40:06	6	0.0012	0	–	–	–	–	–	–	–	–
52	40:10	1	0.0002	0	–	–	–	–	–	–	–	–
53	40:27	1	0.0002	0	–	–	–	–	–	–	–	–
54	40:32	1	0.0002	0	–	–	–	–	–	–	–	–
55	41:01	23	0.0045	9	0.0029	77.8	100.0	85.7	100.0	100.0	100.0	40:01 (100)
56	41:02	32	0.0062	9	0.0029	100.0	100.0	100.0	100.0	100.0	100.0	–
57	42:01	1	0.0002	0	–	–	–	–	–	–	–	–
58	42:02	1	0.0002	0	–	–	–	–	–	–	–	–
59	44:02	438	0.0851	332	0.1084	98.8	99.5	97.0	99.8	98.3	99.6	44:05 (30)
60	44:03	252	0.0490	158	0.0516	98.7	99.7	97.4	99.8	96.8	99.9	07:02 (38)
61	44:04	2	0.0004	3	0.0010	100.0	99.9	0.0	100.0	–	99.9	44:03 (100)
62	44:05	29	0.0056	1	0.0003	100.0	99.9	100.0	99.9	25.0	100.0	–
63	44:06	1	0.0002	0	–	–	–	–	–	–	–	–
64	44:08	1	0.0002	0	–	–	–	–	–	–	–	–
65	44:27	1	0.0002	0	–	–	–	–	–	–	–	–
66	45:01	25	0.0049	26	0.0085	96.2	100.0	96.0	100.0	100.0	100.0	15:01 (50)
67	46:01	1	0.0002	0	–	–	–	–	–	–	–	–
68	47:01	14	0.0027	15	0.0049	100.0	100.0	93.3	100.0	100.0	100.0	07:02 (100)
69	47:02	1	0.0002	0	–	–	–	–	–	–	–	–
70	48:01	5	0.0010	0	–	–	–	–	–	–	–	–
71	49:01	88	0.0171	35	0.0114	94.3	99.9	97.0	100.0	97.0	100.0	44:02 (100)
72	50:01	72	0.0140	32	0.0104	96.9	99.9	93.5	100.0	96.7	99.9	49:01 (50)
73	50:02	1	0.0002	0	–	–	–	–	–	–	–	–
74	51:01	276	0.0537	119	0.0388	98.3	99.8	98.3	99.8	95.0	99.9	07:02 (25)
75	51:05	3	0.0006	0	–	–	–	–	–	–	–	–
76	51:07	5	0.0010	0	–	–	–	–	–	–	–	–
77	51:08	4	0.0008	1	0.0003	100.0	100.0	0.0	100.0	–	100.0	51:01 (100)
78	51:09	1	0.0002	0	–	–	–	–	–	–	–	–

Continued on next page ...

Table 2 – continued from previous page

	Allele	Num. Train	Freq. Train	Num. Valid.	Freq. Valid.	CR[1] (%)	ACC[1] (%)	SEN (%)	SPE (%)	PPV (%)	NPV (%)	Miscall[2] (%)
79	52:01	65	0.0126	20	0.0065	90.0	99.9	94.4	100.0	94.4	100.0	51:01 (100)
80	53:01	13	0.0025	6	0.0020	83.3	100.0	100.0	100.0	100.0	100.0	–
81	54:01	3	0.0006	0	–	–	–	–	–	–	–	–
82	55:01	86	0.0167	61	0.0199	100.0	100.0	100.0	100.0	98.4	100.0	–
83	55:02	3	0.0006	0	–	–	–	–	–	–	–	–
84	55:05	1	0.0002	0	–	–	–	–	–	–	–	–
85	56:01	46	0.0089	11	0.0036	100.0	99.9	90.9	100.0	95.2	100.0	55:01 (100)
86	57:01	153	0.0297	131	0.0428	97.7	99.8	98.4	99.9	97.3	99.9	07:02 (25)
87	57:02	–	–	1	0.0003	100.0	100.0	0.0	100.0	–	100.0	57:01 (100)
88	57:03	1	0.0002	0	–	–	–	–	–	–	–	–
89	58:01	47	0.0091	12	0.0039	100.0	100.0	91.7	100.0	100.0	100.0	35:01 (100)
90	70:20	2	0.0004	0	–	–	–	–	–	–	–	–
91	73:01	5	0.0010	0	–	–	–	–	–	–	–	–
92	81:01	1	0.0002	0	–	–	–	–	–	–	–	–
HLA–C (# of training samples: 1866, # of validation samples: 840)												
Overall accuracy: 96.6%, Call rate: 99.5%												
1	01:02	157	0.0421	89	0.0530	100.0	99.8	96.6	99.9	98.9	99.8	02:02 (33)
2	02:02	208	0.0557	74	0.0440	98.6	99.5	95.9	99.7	94.0	99.8	03:04 (100)
3	02:10	2	0.0005	0	–	–	–	–	–	–	–	–
4	03:02	8	0.0021	1	0.0006	100.0	100.0	100.0	100.0	100.0	100.0	–
5	03:03	182	0.0488	108	0.0643	99.1	98.9	95.3	99.1	88.3	99.7	07:01 (60)
6	03:04	219	0.0587	196	0.1167	99.5	98.7	90.8	99.7	97.5	98.8	03:03 (67)
7	03:05	2	0.0005	0	–	–	–	–	–	–	–	–
8	03:06	1	0.0003	0	–	–	–	–	–	–	–	–
9	03:10	1	0.0003	0	–	–	–	–	–	–	–	–
10	04:01	427	0.1144	167	0.0994	100.0	99.8	99.4	99.8	98.2	99.9	16:01 (100)
11	04:03	1	0.0003	0	–	–	–	–	–	–	–	–
12	04:09	1	0.0003	0	–	–	–	–	–	–	–	–
13	05:01	300	0.0804	210	0.1250	99.5	99.4	98.1	99.6	97.2	99.7	04:01 (38)
14	06:02	350	0.0938	207	0.1232	100.0	99.6	97.6	99.9	99.5	99.7	07:02 (60)
15	07:01	551	0.1476	9	0.0054	100.0	99.6	100.0	99.6	60.0	100.0	–
16	07:02	485	0.1300	261	0.1554	100.0	99.3	98.9	99.4	95.5	99.8	02:02 (33)
17	07:04	75	0.0201	36	0.0214	100.0	99.8	94.4	99.9	94.4	99.9	02:02 (50)
18	08:01	3	0.0008	1	0.0006	0.0	–	–	–	–	–	–
19	08:02	97	0.0260	88	0.0524	97.7	99.6	93.0	100.0	100.0	99.6	05:01 (62)
20	08:03	3	0.0008	0	–	–	–	–	–	–	–	–
21	12:02	50	0.0134	14	0.0083	100.0	100.0	100.0	100.0	100.0	100.0	–
22	12:03	262	0.0702	63	0.0375	98.4	99.8	96.8	99.9	99.2	99.9	07:01 (25)
23	12:05	1	0.0003	0	–	–	–	–	–	–	–	–
24	12:13	1	0.0003	0	–	–	–	–	–	–	–	–
25	14:02	42	0.0113	19	0.0113	100.0	99.9	94.7	99.9	100.0	99.9	01:02 (50)
26	14:03	2	0.0005	0	–	–	–	–	–	–	–	–
27	14:04	1	0.0003	0	–	–	–	–	–	–	–	–
28	15:02	88	0.0236	40	0.0238	100.0	99.8	95.0	99.9	96.2	99.9	02:02 (50)
29	15:04	4	0.0011	0	–	–	–	–	–	–	–	–
30	15:05	16	0.0043	4	0.0024	100.0	99.9	75.0	100.0	100.0	99.9	15:02 (100)

Continued on next page ...

Table 2 – continued from previous page

	Allele	Num. Train	Freq. Train	Num. Valid.	Freq. Valid.	CR[1] (%)	ACC[1] (%)	SEN (%)	SPE (%)	PPV (%)	NPV (%)	Miscall[2] (%)
31	15:11	1	0.0003	0	–	–	–	–	–	–	–	–
32	16:01	125	0.0335	74	0.0440	100.0	99.8	98.6	99.9	98.0	99.9	07:01 (50)
33	16:02	11	0.0029	6	0.0036	100.0	100.0	100.0	100.0	100.0	100.0	–
34	16:04	9	0.0024	0	–	–	–	–	–	–	–	–
35	17:01	40	0.0107	13	0.0077	100.0	100.0	100.0	100.0	100.0	100.0	–
36	17:03	5	0.0013	0	–	–	–	–	–	–	–	–
37	18:01	1	0.0003	0	–	–	–	–	–	–	–	–
HLA–DRB1 (# of training samples: 2436, # of validation samples: 1129)												
Overall accuracy: 94.0%, Call rate: 94.6%												
1	01:01	406	0.0833	98	0.0434	99.0	98.7	100.0	98.7	77.3	100.0	–
2	01:02	68	0.0140	22	0.0097	100.0	100.0	95.5	100.0	100.0	100.0	01:01 (100)
3	01:03	36	0.0074	57	0.0252	87.7	98.9	54.0	100.0	100.0	99.0	01:01 (100)
4	01:04	–	–	1	0.0004	100.0	100.0	0.0	100.0	–	100.0	01:01 (100)
5	03:01	554	0.1137	429	0.1900	96.3	99.7	98.5	99.9	99.8	99.7	04:01 (33)
6	03:02	1	0.0002	0	–	–	–	–	–	–	–	–
7	04:01	377	0.0774	282	0.1249	93.6	98.1	95.8	98.5	88.7	99.4	04:07 (36)
8	04:02	57	0.0117	11	0.0049	72.7	99.8	87.5	99.9	77.8	100.0	04:04 (100)
9	04:03	56	0.0115	46	0.0204	78.3	98.4	8.3	100.0	75.0	98.5	04:01 (51)
10	04:04	138	0.0283	117	0.0518	88.0	99.2	90.3	99.7	93.7	99.5	
11	04:05	36	0.0074	16	0.0071	81.2	99.9	92.3	99.9	85.7	100.0	04:01 (100)
12	04:06	2	0.0004	0	–	–	–	–	–	–	–	–
13	04:07	53	0.0109	23	0.0102	82.6	99.2	78.9	99.3	51.7	99.8	04:01 (50)
14	04:08	17	0.0035	0	–	–	–	–	–	–	–	–
15	04:09	1	0.0002	0	–	–	–	–	–	–	–	–
16	04:10	2	0.0004	0	–	–	–	–	–	–	–	–
17	04:59	1	0.0002	0	–	–	–	–	–	–	–	–
18	07:01	643	0.1320	421	0.1864	97.9	99.6	98.5	99.8	99.5	99.7	04:01 (67)
19	08:01	121	0.0248	42	0.0186	97.6	99.9	97.6	100.0	97.6	100.0	08:03 (100)
20	08:02	8	0.0016	0	–	–	–	–	–	–	–	–
21	08:03	7	0.0014	3	0.0013	100.0	99.9	66.7	100.0	66.7	100.0	
22	08:04	9	0.0018	5	0.0022	100.0	99.9	60.0	100.0	100.0	99.9	08:01 (50)
23	08:10	2	0.0004	0	–	–	–	–	–	–	–	–
24	09:01	48	0.0099	38	0.0168	94.7	100.0	97.2	100.0	100.0	100.0	04:01 (100)
25	10:01	33	0.0068	13	0.0058	92.3	100.0	91.7	100.0	100.0	100.0	07:01 (50)
26	11:01	342	0.0702	72	0.0319	88.9	99.3	98.4	99.3	79.7	100.0	11:04 (100)
27	11:02	12	0.0025	11	0.0049	72.7	99.9	75.0	100.0	100.0	99.9	12:01 (50)
28	11:03	41	0.0084	10	0.0044	80.0	99.7	12.5	100.0	100.0	99.7	11:01 (100)
29	11:04	206	0.0423	24	0.0106	91.7	99.6	68.2	99.9	88.2	99.7	11:01 (71)
30	11:06	1	0.0002	0	–	–	–	–	–	–	–	–
31	11:08	1	0.0002	0	–	–	–	–	–	–	–	–
32	11:12	2	0.0004	0	–	–	–	–	–	–	–	–
33	11:15	1	0.0002	0	–	–	–	–	–	–	–	–
34	11:28	1	0.0002	0	–	–	–	–	–	–	–	–
35	11:39	1	0.0002	0	–	–	–	–	–	–	–	–
36	11:43	1	0.0002	0	–	–	–	–	–	–	–	–
37	12:01	74	0.0152	35	0.0155	94.3	99.8	93.9	99.9	91.2	99.9	04:01 (25)

Continued on next page ...

Table 2 – continued from previous page

	Allele	Num. Train	Freq. Train	Num. Valid.	Freq. Valid.	CR[1] (%)	ACC[1] (%)	SEN (%)	SPE (%)	PPV (%)	NPV (%)	Miscall[2] (%)
38	12:02	3	0.0006	2	0.0009	100.0	99.9	0.0	100.0	–	99.9	12:01 (100)
39	13:01	312	0.0640	100	0.0443	98.0	99.9	100.0	99.9	98.0	100.0	–
40	13:02	176	0.0361	71	0.0314	98.6	100.0	98.6	100.0	100.0	100.0	07:01 (50)
41	13:03	61	0.0125	25	0.0111	92.0	100.0	100.0	100.0	100.0	100.0	–
42	13:05	7	0.0014	1	0.0004	0.0	–	–	–	–	–	–
43	13:10	1	0.0002	1	0.0004	100.0	100.0	0.0	100.0	–	100.0	13:01 (100)
44	13:15	1	0.0002	0	–	–	–	–	–	–	–	–
45	13:17	1	0.0002	0	–	–	–	–	–	–	–	–
46	13:18	1	0.0002	0	–	–	–	–	–	–	–	–
47	14:01	135	0.0277	69	0.0306	94.2	99.8	100.0	99.8	96.3	100.0	–
48	14:02	2	0.0004	0	–	–	–	–	–	–	–	–
49	14:04	7	0.0014	1	0.0004	100.0	100.0	0.0	100.0	–	100.0	14:01 (100)
50	14:06	2	0.0004	0	–	–	–	–	–	–	–	–
51	15:01	594	0.1219	184	0.0815	97.3	99.8	100.0	99.8	97.5	100.0	–
52	15:02	56	0.0115	10	0.0044	100.0	100.0	100.0	100.0	95.2	100.0	–
53	15:03	2	0.0004	0	–	–	–	–	–	–	–	–
54	16:01	136	0.0279	16	0.0071	93.8	99.9	93.3	100.0	96.6	100.0	16:02 (100)
55	16:02	14	0.0029	2	0.0009	100.0	100.0	100.0	100.0	66.7	100.0	–
56	16:05	2	0.0004	0	–	–	–	–	–	–	–	–
HLA–DQB1 (# of training samples: 1924, # of validation samples: 1004)												
Overall accuracy: 98.0%, Call rate: 99.0%												
1	02:01	424	0.1102	27	0.0134	100.0	99.6	88.9	99.7	82.3	99.8	03:02 (67)
2	02:02	400	0.1040	24	0.0120	91.7	99.6	100.0	99.6	72.5	100.0	–
3	02:03	–	–	1	0.0005	100.0	99.9	0.0	100.0	–	100.0	02:02 (100)
4	02:04	1	0.0003	0	–	–	–	–	–	–	–	–
5	03:01	777	0.2019	508	0.2530	99.2	99.3	97.8	99.8	99.5	99.3	03:03 (45)
6	03:02	374	0.0972	236	0.1175	100.0	99.5	97.9	99.8	98.3	99.7	02:01 (53)
7	03:03	167	0.0434	148	0.0737	99.3	99.5	97.3	99.7	96.3	99.8	02:02 (50)
8	03:04	10	0.0026	1	0.0005	100.0	100.0	100.0	100.0	100.0	100.0	–
9	03:05	4	0.0010	1	0.0005	100.0	99.9	0.0	100.0	–	100.0	02:02 (50)
10	03:19	4	0.0010	0	–	–	–	–	–	–	–	–
11	04:02	121	0.0314	61	0.0304	100.0	100.0	100.0	100.0	100.0	100.0	–
12	05:01	433	0.1125	314	0.1564	99.4	99.9	99.7	99.9	99.7	99.9	03:03 (50)
13	05:02	132	0.0343	20	0.0100	100.0	99.9	95.0	99.9	95.0	99.9	05:04 (100)
14	05:03	111	0.0288	65	0.0324	98.5	99.8	98.4	99.9	96.9	99.9	03:01 (50)
15	05:04	2	0.0005	3	0.0015	100.0	99.9	100.0	99.9	75.0	100.0	–
16	06:01	46	0.0120	15	0.0075	100.0	100.0	100.0	100.0	100.0	100.0	–
17	06:02	462	0.1201	340	0.1693	98.8	99.5	98.5	99.7	98.7	99.7	06:03 (40)
18	06:03	244	0.0634	141	0.0702	96.5	99.5	96.3	99.8	97.0	99.7	06:02 (80)
19	06:04	107	0.0278	79	0.0393	98.7	99.8	97.4	99.9	98.7	99.9	06:03 (100)
20	06:09	27	0.0070	24	0.0120	100.0	100.0	100.0	100.0	100.0	100.0	–
21	06:14	1	0.0003	0	–	–	–	–	–	–	–	–
22	06:16	1	0.0003	0	–	–	–	–	–	–	–	–

[1]: CR – call rate; ACC – allele accuracy.
[2]: the most likely miscalled allele and the proportion of the most likely miscalled allele in all miscalled alleles.