Supplementary Material

HLA imputation in an admixed population: an assessment of the 1000 genomes data as a reference panel

Kelly Nunes[a], Xiuwen Zheng[b], Margareth Torres[c], Maria Elisa Moraes[c], Bruno Z. Piovezan[c], Gerlandia N. Pontes[c], Lilian Kimura[a], Juliana E.P. Carnavalli[a], Regina C. Mingroni Netto[a], Diogo Meyer[a]

TABLES – Supplementary Material

Table S1. QUI Sample description: communities and sample size.

Quilombo communities	Sample size
Abobral	59
André Lopes	41
Galvão	24
Ivaporanduva	30
Maria Rosa	09
Nhunguara	31
Pedro Cubas	35
Pilões	21
Poças	26
Reginaldo	28
São Pedro	23
Sapatu	38
Total	365*

* Lemes et al. (2014) showed that the levels of substructure of those *quilombo* communities are very small, therefore, the set of communities here reported can be considered as a single population aggregate.

Table S2. 1000g Sample description: population, population origin and sample size.

Population	Population Origin	Sample size
Luhya from Webuye, Kenya (LWK)	African	87
Yoruba from Ibadan,Nigeria (YRI)	African	39
British from England and Scotland, UK (GRB)	European	89
Finnish, Finland (FIN)	European	93
Northern and Western European from Utah, USA (CEU)	European	45
Italian from Tuscany, Italy (TSI)	European	90
Han from south, China (CHS)	East Asian	100
Han Chinese from Beijing, China (CHB) + Japanese from Tokyo, Japan (JPT)	East Asian	165
African Ancestry from Southwest, USA (ASW)	Admixed	53
Colombian from Medellin, Colombia (CLM)	Admixed	60
Mexican Ancestry from Los Angeles California, USA (MXL)	Admixed	55
Puerto Rican, Puerto Rico (PUR)	Admixed	55
TOTAL		931

Table S3. UW_2014 Sample description: population, population origin and sample size.

Dataset	Population	Population Origin	Sample size
HAPMAP	Yoruban	African	30
	CEU	European	30
	Han Chinese	Asian	45
	Japan	Asian	45
HLARES	African	African	173
	European	European	2668
	Asian	Asian	720
	Hispanic	Admixed	439

Table S4. Imputation models characteristics

1000g	1-field	2-fields	1-field	2-fields	1-field	2-fields	1-field	2-fields
Individuals in training panel	931	931	931	931	931	931	931	931
HLA alleles in training panel	20	54	36	103	17	36	13	52
SNPs in training panel	249	249	401	401	456	456	395	395
Quilombo samples with known HLA	146	146	146	146	144	144	146	146
1000g+QUI								
Individuals in training panel	988	988	988	988	986	986	988	988
HLA alleles in training panel	20	58	36	105	17	37	13	52
SNPs in training panel	245	245	397	397	451	451	379	379
Quilombo samples with known HLA	89*	89*	89*	89*	89*	89*	89*	89*
UW								
Individuals in training panel	3090	2901	3977	3886	3102	2916	3905	3713
HLA alleles in training panel	25	83	35	142	14	49	13	79
SNPs in training panel	143	143	176	176	192	192	125	125
Quilombo samples with known HLA	146	146	146	146	144	144	146	146
1000g_UW_shared_SNPs								
Individuals in training panel	931	931	931	931	931	931	931	931
HLA alleles in training panel	20	54	36	103	17	36	13	52
SNPs in training panel	101	101	174	174	190	190	125	125
Quilombo samples with known HLA	146	146	146	146	144	144	146	146

- Without overlapping sample (146 QUI – 57 unrelated QUI = 89 samples)

Table S5. Number of samples with maximum kinship coefficient with any sample in the other subset in the specified range

Kinship coefficient	57 subset	89 subset
<1/32	22	29
1/32 ≤ 1/16	21	41
1/16 ≤ <1/8	7	11
1/8 ≤ 1/4	4	5
≥1/4	3	3
Total	57	89

Table S6. Model accuracy by posterior probability threshold for HLA allele imputation two-fields (4 digits).

Post prob	HLA-A		HLA-B		HLA-C		HLA-DRB1	
	N	Accuracy	N	Accuracy	N	Accuracy	N	Accuracy
1000g (two-fields resolution)								
>10	146	82.9	144	82.3	144	94.8	146	86.6
>20	146	82.9	132	83.7	144	94.8	144	87.9
>30	141	84.0	116	84.9	141	95.0	132	89.8
>40	138	84.4	97	88.7	134	96.3	115	93.0
>50	130	86.5	85	90.0	130	96.2	101	94.1
>60	122	87.3	68	90.4	124	96.4	91	94.5
>70	96	94.8	53	90.1	111	96.8	70	97.8
>80	87	95.4	40	90.0	105	97.5	63	99.2
>90	60	94.4	23	91.3	79	97.5	40	98.8
1000g+QUI								
>10	146	84.6	144	83.3	144	95.1	146	86.6
>20	146	84.6	132	84.5	144	95.1	141	87.9
>30	141	86.2	116	85.3	141	95.4	132	89.8
>40	138	86.6	97	88.6	136	96.3	115	93.5
>50	131	88.5	85	90.0	130	96.1	101	94.5
>60	122	88.9	68	89.7	124	96.4	91	95.0
>70	97	95.9	53	90.1	111	96.8	70	97.8
>80	88	96.0	40	90.0	105	97.6	63	99.2
>90	61	95.1	23	91.3	79	97.5	40	98.7
UW								
>10	146	84.2	145	65.9	144	86.8	145	78.3
>20	144	85.1	118	72.5	142	87.3	130	81.5
>30	136	87.5	99	77.8	135	88.5	114	85.9
>40	132	89.0	81	84.6	118	92.4	98	90.3
>50	126	89.3	63	90.5	116	92.7	83	92.2
>60	121	90.1	50	92.0	109	94.5	71	94.4
>70	111	90.1	40	91.2	94	95.2	59	94.9
>80	94	92.5	25	94.0	73	95.6	51	95.1
>90	72	94.4	18	94.4	61	96.7	37	97.3

Table S7. *Test* of non-independence of imputation accuracy across *HLA* loci pairs. The p-values were obtained using a Fisher Exact test.

Locus pair	***p-value (Fisher test)***
HLA-A and -B	0.64
HLA-A and -C	0.49
HLA-A and -DRB1	0.05
HLA-B and -C	0.0001
HLA-B and -DRB1	0.38
HLA-C and -DRB1	0.44

S8. The sensitivity (SEN), specificity (SPE), positive predictive value (PPV) and negative predictive value (NPV) calculated from validation samples for each four-digit HLA allele to 1000g model.

Allele[1]	Num. Valid.	Freq. Valid.	CR[2] (%)	ACC[3] (%)	SEN (%)	SPE (%)	PPV (%)	NPV (%)	Miscall[4] (%)
HLA-A									
1:01	13	0.0445	100.0	94.2	61.5	95.7	39.0	98.2	36:01 (80)
1:02	1	0.0034	100.0	99.7	0.0	100.0	--	99.7	01:01 (100)
1:03	11	0.0377	100.0	96.2	0.0	100.0	--	96.2	01:01 (100)
2:01	40	0.1370	100.0	99.3	97.5	99.6	97.5	99.6	... (100)
2:02	4	0.0137	100.0	100.0	100.0	100.0	100.0	100.0	--
2:05	2	0.0068	100.0	99.7	100.0	99.7	80.0	100.0	--
2:05	1	0.0034	100.0	99.7	0.0	100.0	--	99.7	01:01 (50)
2:11	1	0.0034	100.0	99.7	0.0	100.0	--	99.7	... (100)
3:01	25	0.0856	100.0	99.0	100.0	98.9	89.3	100.0	--
11:01	9	0.0308	100.0	100.0	100.0	100.0	100.0	100.0	--
23:01	38	0.1301	100.0	98.3	100.0	98.0	88.4	100.0	--
23:05	5	0.0171	100.0	98.3	0.0	100.0	--	98.3	23:01 (100)
24:02:00	18	0.0616	100.0	99.7	100.0	99.6	94.7	100.0	--
24:03:00	1	0.0034	100.0	99.7	0.0	100.0	--	99.7	24:02 (100)
25:01:00	1	0.0034	100.0	100.0	100.0	100.0	100.0	100.0	--
26:01:00	5	0.0171	100.0	96.9	100.0	96.9	35.7	100.0	--
26:08:00	3	0.0103	100.0	99.0	0.0	100.0	--	99.0	26:01 (100)
29:02:00	7	0.0240	100.0	100.0	100.0	100.0	100.0	100.0	--
30:01:00	16	0.0548	100.0	100.0	100.0	100.0	100.0	100.0	--
30:02:00	13	0.0445	100.0	99.7	100.0	99.6	92.9	100.0	--
30:04:00	2	0.0068	100.0	99.3	0.0	100.0	--	99.3	03:01 (100)
31:01:00	7	0.0240	100.0	99.3	100.0	99.3	77.8	100.0	--
32:01:00	2	0.0068	100.0	100.0	100.0	100.0	100.0	100.0	--
33:01:00	5	0.0171	100.0	99.7	80.0	100.0	100.0	99.7	30:02 (50)
33:03:00	5	0.0171	100.0	99.7	100.0	99.7	83.3	100.0	--
34:02:00	5	0.0171	100.0	99.3	60.0	100.0	100.0	99.3	02:01 (50)
36:01:00	6	0.0205	100.0	97.3	100.0	97.2	42.9	100.0	--
66:01:00	7	0.0240	100.0	97.6	0.0	100.0	--	97.6	26:01 (86)
66:03:00	5	0.0171	100.0	98.3	0.0	100.0	--	98.3	36:01 (80)
68:01:00	5	0.0171	100.0	100.0	100.0	100.0	100.0	100.0	--
68:02:00	11	0.0377	100.0	100.0	100.0	100.0	100.0	100.0	--
74:01:00	10	0.0342	100.0	98.6	100.0	98.6	71.4	100.0	--
74:10:00	4	0.0137	100.0	98.6	0.0	100.0	--	98.6	74:01 (100)
80:01:00	4	0.0137	100.0	100.0	100.0	100.0	100.0	100.0	--
HLA-B									
7:02	17	0.0582	100.0	99.3	94.1	99.6	94.1	99.6	44:03 (100)
8:01	19	0.0651	100.0	97.9	73.7	99.6	93.3	98.2	... (60)
8:01	1	0.0034	100.0	99.7	0.0	100.0	--	99.7	08:01 (100)
13:02	4	0.0137	100.0	100.0	100.0	100.0	100.0	100.0	--

14:01	1	0.0034	100.0	100.0	100.0	100.0	100.0	100.0	--
14:02	10	0.0342	100.0	99.3	90.0	99.6	90.0	99.6	15:10 (50)
14:03	1	0.0034	100.0	99.7	0.0	100.0	--	99.7	14:02 (100)
15:01	4	0.0137	100.0	99.0	100.0	99.0	61.5	100.0	--
15:03	17	0.0582	100.0	99.7	94.1	100.0	100.0	99.6	07:02 (100)
15:04	2	0.0068	100.0	99.3	0.0	100.0	--	99.3	15:01 (100)
15:10	6	0.0205	100.0	99.3	83.3	99.7	83.3	99.7	15:01 (50)
15:16	5	0.0171	100.0	100.0	100.0	100.0	100.0	100.0	--
15:17	3	0.0103	100.0	100.0	100.0	100.0	100.0	100.0	--
15:18	1	0.0034	100.0	100.0	100.0	100.0	100.0	100.0	--
15:20	3	0.0103	100.0	99.0	0.0	100.0	--	99.0	51:01 (100)
18:01	7	0.0240	100.0	99.0	71.4	99.6	83.3	99.3	42:01 (50)
35:01:00	17	0.0582	100.0	95.5	94.1	95.6	57.1	99.6	35:02 (100)
35:02:00	6	0.0205	100.0	99.3	83.3	99.7	83.3	99.7	35:01 (100)
35:03:00	8	0.0274	100.0	97.9	37.5	99.6	75.0	98.3	35:01 (100)
35:05:00	2	0.0068	100.0	99.3	0.0	100.0	--	99.3	35:01 (100)
35:20:00	1	0.0034	100.0	99.7	0.0	100.0	--	99.7	35:01 (100)
37:01:00	5	0.0171	100.0	98.3	0.0	100.0	--	98.3	... (80)
38:01:00	6	0.0205	100.0	99.3	100.0	99.3	75.0	100.0	--
39:05:00	3	0.0103	100.0	99.7	100.0	99.7	75.0	100.0	--
39:06:00	1	0.0034	100.0	100.0	100.0	100.0	100.0	100.0	--
39:10:00	2	0.0068	100.0	99.3	0.0	100.0	--	99.3	38:01 (100)
40:01:00	1	0.0034	100.0	100.0	100.0	100.0	100.0	100.0	--
40:04:00	1	0.0034	100.0	99.7	0.0	100.0	--	99.7	... (100)
41:01:00	8	0.0274	100.0	98.6	75.0	99.3	75.0	99.3	51:01 (50)
41:02:00	1	0.0034	100.0	99.7	0.0	100.0	--	99.7	41:01 (100)
42:01:00	19	0.0651	100.0	99.0	94.7	99.3	90.0	99.6	41:01 (100)
44:02:00	12	0.0411	100.0	100.0	100.0	100.0	100.0	100.0	--
44:03:00	16	0.0548	100.0	99.0	100.0	98.9	84.2	100.0	--
45:01:00	28	0.0959	100.0	99.3	96.4	99.6	96.4	99.6	18:01 (50)
48:02:00	1	0.0034	100.0	100.0	100.0	100.0	100.0	100.0	--
49:01:00	8	0.0274	100.0	99.0	87.5	99.3	77.8	99.6	45:01 (100)
50:01:00	2	0.0068	100.0	100.0	100.0	100.0	100.0	100.0	--
50:02:00	1	0.0034	100.0	99.7	0.0	100.0	--	99.7	44:03 (100)
51:01:00	11	0.0377	100.0	97.3	100.0	97.2	56.4	100.0	--
52:01:00	7	0.0240	100.0	98.3	28.6	100.0	100.0	98.3	51:01 (80)
53:01:00	9	0.0308	100.0	99.3	77.8	100.0	100.0	99.3	35:01 (100)
57:01:00	4	0.0137	100.0	100.0	100.0	100.0	100.0	100.0	--
57:03:00	2	0.0068	100.0	99.3	0.0	100.0	--	99.3	49:01 (100)
58:01:00	6	0.0205	100.0	99.7	83.3	100.0	100.0	99.7	35:01 (100)
81:01:00	3	0.0103	100.0	99.7	100.0	99.7	75.0	100.0	--
HLA-C									
1:02	4	0.0139	100.0	100.0	100.0	100.0	100.0	100.0	--
2:02	8	0.0278	100.0	99.3	87.5	99.6	87.5	99.6	02:10 (100)
2:10	22	0.0764	100.0	99.3	95.5	99.6	95.5	99.6	02:02 (100)
3:02	3	0.0104	100.0	100.0	100.0	100.0	100.0	100.0	--

3:03	4	0.0139	100.0	98.6	100.0	98.6	50.0	100.0	--
3:04	30	0.1042	100.0	97.9	86.7	99.2	92.9	98.5	03:03 (100)
4:01	46	0.1597	100.0	99.7	100.0	99.6	97.9	100.0	--
4:13	1	0.0035	100.0	99.7	0.0	100.0	--	99.7	04:01 (100)
5:01	14	0.0486	100.0	100.0	100.0	100.0	100.0	100.0	--
6:02	9	0.0312	100.0	100.0	100.0	100.0	100.0	100.0	--
7:01	23	0.0799	100.0	99.7	100.0	99.6	95.8	100.0	--
7:02	17	0.0590	100.0	100.0	100.0	100.0	100.0	100.0	--
7:04	2	0.0069	100.0	100.0	100.0	100.0	100.0	100.0	--
8:02	13	0.0451	100.0	99.7	92.3	100.0	100.0	99.6	03:04 (100)
8:04	3	0.0104	100.0	100.0	100.0	100.0	100.0	100.0	--
12:02	2	0.0069	100.0	100.0	100.0	100.0	100.0	100.0	--
12:03	10	0.0347	100.0	100.0	100.0	100.0	100.0	100.0	--
14:02	5	0.0174	100.0	100.0	100.0	100.0	100.0	100.0	--
15:02	5	0.0174	100.0	98.3	100.0	98.2	50.0	100.0	--
15:04	1	0.0035	100.0	99.7	0.0	100.0	--	99.7	15:02 (100)
15:39	3	0.0104	100.0	99.0	0.0	100.0	--	99.0	15:02 (100)
16:01	21	0.0729	100.0	100.0	100.0	100.0	100.0	100.0	--
16:02	1	0.0035	100.0	99.7	0.0	100.0	--	99.7	15:02 (100)
17:01	39	0.1354	100.0	99.3	94.9	100.0	100.0	99.2	03:04 (50)
18:01	2	0.0069	100.0	100.0	100.0	100.0	100.0	100.0	--
HLA-DRB1									
1:01	14	0.0479	100.0	100.0	100.0	100.0	100.0	100.0	--
1:02	23	0.0788	100.0	100.0	100.0	100.0	100.0	100.0	--
3:01	11	0.0377	100.0	98.6	90.9	98.9	76.9	99.6	12:01 (50)
3:02	17	0.0582	100.0	100.0	100.0	100.0	100.0	100.0	--
4:01	4	0.0137	100.0	97.9	0.0	99.3	0.0	98.6	04:05 (75)
4:02	3	0.0103	100.0	100.0	100.0	100.0	100.0	100.0	--
4:04	2	0.0068	100.0	99.7	100.0	99.7	66.7	100.0	--
4:05	1	0.0034	100.0	99.0	100.0	99.0	25.0	100.0	--
4:11	3	0.0103	100.0	99.0	0.0	100.0	--	99.0	... (100)
7:01	34	0.1164	100.0	98.6	88.2	100.0	100.0	98.5	04:01 (25)
8:01	3	0.0103	100.0	100.0	100.0	100.0	100.0	100.0	--
8:02	9	0.0308	100.0	99.0	100.0	98.9	75.0	100.0	--
8:04	25	0.0856	100.0	98.6	88.0	99.6	95.7	98.9	08:02 (67)
9:01	1	0.0034	100.0	99.7	0.0	100.0	--	99.7	04:01 (100)
10:01	3	0.0103	100.0	99.7	100.0	99.7	75.0	100.0	--
11:01	27	0.0925	100.0	96.2	96.3	96.2	72.2	99.6	03:01 (100)
11:02	9	0.0308	100.0	99.0	88.9	99.3	80.0	99.6	08:04 (100)
11:04	7	0.0240	100.0	97.6	0.0	100.0	--	97.6	11:01 (100)
12:01	11	0.0377	100.0	98.6	100.0	98.6	73.3	100.0	--
12:02	2	0.0068	100.0	99.3	0.0	100.0	--	99.3	03:01 (100)
13:01	15	0.0514	100.0	99.3	100.0	99.3	88.2	100.0	--
13:02	7	0.0240	100.0	99.7	85.7	100.0	100.0	99.7	13:01 (100)
13:04	4	0.0137	100.0	98.6	0.0	100.0	--	98.6	11:01 (50)
13:05	1	0.0034	100.0	99.7	0.0	100.0	--	99.7	11:01 (100)

14:01	4	0.0137	100.0	100.0	100.0	100.0	100.0	100.0	--
14:02	3	0.0103	100.0	99.0	0.0	100.0	--	99.0	12:01 (67)
14:06	1	0.0034	100.0	99.7	0.0	100.0	--	99.7	12:01 (100)
15:01	13	0.0445	100.0	99.3	100.0	99.3	86.7	100.0	--
15:03	31	0.1062	100.0	99.0	93.5	99.6	96.7	99.2	15:01 (100)
16:02	4	0.0137	100.0	100.0	100.0	100.0	100.0	100.0	--

[1]: the HLA alleles with more than one copy and non-zero sensitivity in the training are listed.[2]: CR – call rate.[3]: ACC – allele accuracy.[4]: the most likely miscalled allele and the proportion of the most likely miscalled allele in all miscalled alleles.

S9. The sensitivity (SEN), specificity (SPE), positive predictive value (PPV) and negative predictive value (NPV) calculated from validation samples for each four-digit HLA allele to 1000g+QUI model.

Allele1	Num. Valid.	Freq. Valid.	CR2 (%)	ACC3 (%)	SEN (%)	SPE (%)	PPV (%)	NPV (%)	Miscall4 (%)
HLA-A									
1:01	7	0.0393	100.0	97.8	42.9	100.0	100.0	97.7	36:01 (75)
1:03	8	0.0449	100.0	100.0	100.0	100.0	100.0	100.0	--
2:01	21	0.1180	100.0	98.9	95.2	99.4	95.2	99.4	... (100)
2:02	3	0.0169	100.0	100.0	100.0	100.0	100.0	100.0	--
2:05	1	0.0056	100.0	99.4	0.0	100.0	--	99.4	02:05 (100)
2:05	1	0.0056	100.0	99.4	100.0	99.4	50.0	100.0	--
2:11	1	0.0056	100.0	99.4	0.0	100.0	--	99.4	... (100)
3:01	17	0.0955	100.0	98.9	100.0	98.8	89.5	100.0	--
11:01	5	0.0281	100.0	100.0	100.0	100.0	100.0	100.0	--
23:01	26	0.1461	100.0	98.3	100.0	98.0	89.7	100.0	--
23:05	3	0.0169	100.0	98.3	0.0	100.0	--	98.3	23:01 (100)
24:02:00	10	0.0562	100.0	100.0	100.0	100.0	100.0	100.0	--
26:01:00	4	0.0225	100.0	94.9	100.0	94.8	30.8	100.0	--
26:08:00	3	0.0169	100.0	98.3	0.0	100.0	--	98.3	26:01 (100)
29:02:00	7	0.0393	100.0	100.0	100.0	100.0	100.0	100.0	--
30:01:00	10	0.0562	100.0	100.0	100.0	100.0	100.0	100.0	--
30:02:00	5	0.0281	100.0	99.4	100.0	99.4	83.3	100.0	--
30:04:00	2	0.0112	100.0	98.9	0.0	100.0	--	98.9	03:01 (100)
31:01:00	3	0.0169	100.0	99.4	100.0	99.4	75.0	100.0	--
32:01:00	1	0.0056	100.0	100.0	100.0	100.0	100.0	100.0	--
33:01:00	3	0.0169	100.0	99.4	66.7	100.0	100.0	99.4	30:02 (50)
33:03:00	2	0.0112	100.0	99.4	100.0	99.4	66.7	100.0	--
34:02:00	4	0.0225	100.0	99.4	75.0	100.0	100.0	99.4	02:01 (100)
36:01:00	3	0.0169	100.0	98.3	100.0	98.3	50.0	100.0	--

66:01:00	7	0.0393	100.0	96.1	0.0	100.0	--	96.1	26:01 (86)
66:03:00	3	0.0169	100.0	100.0	100.0	100.0	100.0	100.0	--
68:01:00	3	0.0169	100.0	100.0	100.0	100.0	100.0	100.0	--
68:02:00	5	0.0281	100.0	100.0	100.0	100.0	100.0	100.0	--
74:01:00	4	0.0225	100.0	98.3	100.0	98.3	57.1	100.0	--
74:10:00	3	0.0169	100.0	98.3	0.0	100.0	--	98.3	74:01 (100)
80:01:00	3	0.0169	100.0	100.0	100.0	100.0	100.0	100.0	--
HLA-B									
7:02	5	0.0431	100.0	100.0	100.0	100.0	100.0	100.0	--
8:01	0	0	--	--	--	--	--	--	--
8:01	6	0.0517	100.0	99.1	83.3	100.0	100.0	99.1	42:01 (100)
13:02	1	0.0086	100.0	99.1	0.0	100.0	--	99.1	14:01 (100)
14:01	0	0	--	--	--	--	--	--	--
14:02	4	0.0345	100.0	99.1	75.0	100.0	100.0	99.1	15:10 (50)
14:03	0	0	--	--	--	--	--	--	--
15:01	1	0.0086	100.0	99.1	100.0	99.1	50.0	100.0	--
15:03	9	0.0776	100.0	99.1	88.9	100.0	100.0	99.1	45:01 (100)
15:04	1	0.0086	100.0	99.1	0.0	100.0	--	99.1	15:01 (100)
15:10	3	0.0259	100.0	99.1	100.0	99.1	75.0	100.0	--
15:16	2	0.0172	100.0	100.0	100.0	100.0	100.0	100.0	--
15:17	1	0.0086	100.0	100.0	100.0	100.0	100.0	100.0	--
15:18	0	0	--	--	--	--	--	--	--
15:20	1	0.0086	100.0	100.0	100.0	100.0	100.0	100.0	--
18:01	3	0.0259	100.0	99.1	66.7	100.0	100.0	99.1	44:02 (50)
35:01:00	7	0.0603	100.0	96.6	100.0	96.3	60.9	100.0	--
35:02:00	2	0.0172	100.0	99.1	50.0	100.0	100.0	99.1	35:01 (100)
35:03:00	5	0.0431	100.0	99.1	80.0	100.0	100.0	99.1	38:01 (100)
35:05:00	1	0.0086	100.0	99.1	0.0	100.0	--	99.1	35:01 (100)
35:20:00	0	0	--	--	--	--	--	--	--
37:01:00	2	0.0172	100.0	100.0	100.0	100.0	100.0	100.0	--
38:01:00	3	0.0259	100.0	98.3	100.0	98.2	60.0	100.0	--
39:05:00	1	0.0086	100.0	99.1	100.0	99.1	50.0	100.0	--
39:06:00	0	0	--	--	--	--	--	--	--
39:10:00	1	0.0086	100.0	99.1	0.0	100.0	--	99.1	38:01 (100)
40:01:00	0	0	--	--	--	--	--	--	--
40:04:00	0	0	--	--	--	--	--	--	--
41:01:00	4	0.0345	100.0	99.1	75.0	100.0	100.0	99.1	15:10 (50)
41:02:00	0	0	--	--	--	--	--	--	--
42:01:00	8	0.0690	100.0	97.4	100.0	97.2	76.2	100.0	--
44:02:00	4	0.0345	100.0	99.1	100.0	99.1	88.9	100.0	--

44:03:00	6	0.0517	100.0	100.0	100.0	100.0	100.0	100.0	--
45:01:00	11	0.0948	100.0	99.1	100.0	99.0	91.7	100.0	--
48:02:00	0	0	--	--	--	--	--	--	--
49:01:00	3	0.0259	100.0	100.0	100.0	100.0	100.0	100.0	--
50:01:00	0	0	--	--	--	--	--	--	--
50:02:00	0	0	--	--	--	--	--	--	--
51:01:00	6	0.0517	100.0	96.6	50.0	99.1	75.0	97.3	52:01 (67)
52:01:00	4	0.0345	100.0	96.6	50.0	98.2	44.4	98.2	51:01 (50)
53:01:00	4	0.0345	100.0	98.3	50.0	100.0	100.0	98.2	35:01 (100)
57:01:00	2	0.0172	100.0	100.0	100.0	100.0	100.0	100.0	--
57:03:00	1	0.0086	100.0	100.0	100.0	100.0	100.0	100.0	--
58:01:00	3	0.0259	100.0	100.0	100.0	100.0	100.0	100.0	--
81:01:00	1	0.0086	100.0	100.0	100.0	100.0	100.0	100.0	--
HLA-C									
1:02	3	0.0169	100.0	100.0	100.0	100.0	100.0	100.0	--
2:02	5	0.0281	100.0	99.4	80.0	100.0	100.0	99.4	02:10 (100)
2:10	18	0.1011	100.0	99.4	100.0	99.4	94.7	100.0	--
3:02	1	0.0056	100.0	100.0	100.0	100.0	100.0	100.0	--
3:03	3	0.0169	100.0	100.0	100.0	100.0	100.0	100.0	--
3:04	20	0.1124	100.0	98.9	100.0	98.7	90.9	100.0	--
4:01	24	0.1348	100.0	100.0	100.0	100.0	100.0	100.0	--
5:01	6	0.0337	100.0	100.0	100.0	100.0	100.0	100.0	--
6:02	5	0.0281	100.0	100.0	100.0	100.0	100.0	100.0	--
7:01	11	0.0618	100.0	100.0	100.0	100.0	100.0	100.0	--
7:02	8	0.0449	100.0	100.0	100.0	100.0	100.0	100.0	--
8:02	10	0.0562	100.0	99.4	90.0	100.0	100.0	99.4	03:04 (100)
8:04	2	0.0112	100.0	100.0	100.0	100.0	100.0	100.0	--
12:02	1	0.0056	100.0	100.0	100.0	100.0	100.0	100.0	--
12:03	8	0.0449	100.0	100.0	100.0	100.0	100.0	100.0	--
14:02	4	0.0225	100.0	100.0	100.0	100.0	100.0	100.0	--
15:02	3	0.0169	100.0	98.3	100.0	98.3	50.0	100.0	--
15:39	2	0.0112	100.0	98.9	0.0	100.0	--	98.9	15:02 (100)
16:01	15	0.0843	100.0	100.0	100.0	100.0	100.0	100.0	--
16:02	1	0.0056	100.0	99.4	0.0	100.0	--	99.4	15:02 (100)
17:01	27	0.1517	100.0	99.4	96.3	100.0	100.0	99.3	03:04 (100)
18:01	1	0.0056	100.0	100.0	100.0	100.0	100.0	100.0	--
HLA-DRB1									
1:01	9	0.0506	100.0	100.0	100.0	100.0	100.0	100.0	--
1:02	13	0.0730	100.0	100.0	100.0	100.0	100.0	100.0	--
3:01	6	0.0337	100.0	98.3	83.3	98.8	71.4	99.4	11:02 (100)
3:02	12	0.0674	100.0	100.0	100.0	100.0	100.0	100.0	--
4:01	2	0.0112	100.0	98.9	100.0	98.9	50.0	100.0	--

4:02	1	0.0056	100.0	100.0	100.0	100.0	100.0	100.0	--
4:04	1	0.0056	100.0	100.0	100.0	100.0	100.0	100.0	--
4:11	1	0.0056	100.0	99.4	0.0	100.0	--	99.4	... (100)
7:01	21	0.1180	100.0	98.3	85.7	100.0	100.0	98.1	04:01 (33)
8:01	2	0.0112	100.0	100.0	100.0	100.0	100.0	100.0	--
8:02	5	0.0281	100.0	98.9	100.0	98.8	71.4	100.0	--
8:04	14	0.0787	100.0	98.9	92.9	99.4	92.9	99.4	08:02 (100)
9:01	1	0.0056	100.0	99.4	0.0	100.0	--	99.4	04:01 (100)
10:01	2	0.0112	100.0	100.0	100.0	100.0	100.0	100.0	--
11:01	15	0.0843	100.0	95.5	100.0	95.1	65.2	100.0	--
11:02	5	0.0281	100.0	98.3	80.0	98.8	66.7	99.4	08:04 (100)
11:04	6	0.0337	100.0	96.6	0.0	100.0	--	96.6	11:01 (100)
12:01	9	0.0506	100.0	100.0	100.0	100.0	100.0	100.0	--
12:02	2	0.0112	100.0	98.9	0.0	100.0	--	98.9	03:01 (100)
13:01	10	0.0562	100.0	100.0	100.0	100.0	100.0	100.0	--
13:02	4	0.0225	100.0	100.0	100.0	100.0	100.0	100.0	--
13:04	3	0.0169	100.0	98.3	0.0	100.0	--	98.3	11:01 (67)
14:01	3	0.0169	100.0	100.0	100.0	100.0	100.0	100.0	--
14:02	1	0.0056	100.0	99.4	100.0	99.4	50.0	100.0	--
14:06	1	0.0056	100.0	99.4	0.0	100.0	--	99.4	14:02 (100)
15:01	7	0.0393	100.0	100.0	100.0	100.0	100.0	100.0	--
15:03	20	0.1124	100.0	99.4	100.0	99.4	95.2	100.0	--
16:02	2	0.0112	100.0	100.0	100.0	100.0	100.0	100.0	--

[1]: the HLA alleles with more than one copy and non-zero sensitivity in the training are listed.[2]: CR – call rate.[3]: ACC – allele accuracy.[4]: the most likely miscalled allele and the proportion of the most likely miscalled allele in all miscalled alleles

S10. The sensitivity (SEN), specificity (SPE), positive predictive value (PPV) and negative predictive value (NPV) calculated from validation samples for each four-digit HLA allele to UW model.

Allele[1]	Num. Valid.	Freq. Valid.	CR[2] (%)	ACC[3] (%)	SEN (%)	SPE (%)	PPV (%)	NPV (%)	Miscall[4] (%)
HLA-A									
1:01	13	0.0445	100.0	98.6	69.2	100.0	100.0	98.6	36:01 (100)
1:02	1	0.0034	100.0	100.0	100.0	100.0	100.0	100.0	--
1:03	11	0.0377	100.0	96.2	0.0	100.0	--	96.2	11:01 (82)
2:01	40	0.1370	100.0	99.7	100.0	99.6	97.6	100.0	--
2:02	4	0.0137	100.0	100.0	100.0	100.0	100.0	100.0	--
2:05	2	0.0068	100.0	99.7	100.0	99.7	80.0	100.0	--

2:05	1	0.0034	100.0	99.7	0.0	100.0	--	99.7	02:05 (50)
2:11	1	0.0034	100.0	99.7	0.0	100.0	--	99.7	... (100)
3:01	25	0.0856	100.0	98.6	100.0	98.5	84.7	100.0	--
11:01	9	0.0308	100.0	96.9	100.0	96.8	48.6	100.0	--
23:01	38	0.1301	100.0	98.3	100.0	98.0	88.4	100.0	--
23:05	5	0.0171	100.0	98.3	0.0	100.0	--	98.3	23:01 (100)
24:02:00	18	0.0616	100.0	99.7	100.0	99.6	94.7	100.0	--
24:03:00	1	0.0034	100.0	99.7	0.0	100.0	--	99.7	24:02 (100)
25:01:00	1	0.0034	100.0	100.0	100.0	100.0	100.0	100.0	--
26:01:00	5	0.0171	100.0	97.6	100.0	97.6	41.7	100.0	--
26:08:00	3	0.0103	100.0	99.0	0.0	100.0	--	99.0	26:01 (100)
29:02:00	7	0.0240	100.0	100.0	100.0	100.0	100.0	100.0	--
30:01:00	16	0.0548	100.0	100.0	100.0	100.0	100.0	100.0	--
30:02:00	13	0.0445	100.0	99.7	100.0	99.6	92.9	100.0	--
30:04:00	2	0.0068	100.0	99.3	0.0	100.0	--	99.3	03:01 (100)
31:01:00	7	0.0240	100.0	99.3	85.7	99.6	85.7	99.6	33:03 (100)
32:01:00	2	0.0068	100.0	99.7	100.0	99.7	80.0	100.0	--
33:01:00	5	0.0171	100.0	99.7	80.0	100.0	100.0	99.7	30:02 (50)
33:03:00	5	0.0171	100.0	99.7	100.0	99.7	83.3	100.0	--
34:02:00	5	0.0171	100.0	99.7	80.0	100.0	100.0	99.7	02:01 (100)
36:01:00	6	0.0205	100.0	96.9	100.0	96.9	40.0	100.0	--
66:01:00	7	0.0240	100.0	98.3	28.6	100.0	100.0	98.3	26:01 (80)
66:03:00	5	0.0171	100.0	98.3	0.0	100.0	--	98.3	36:01 (100)
68:01:00	5	0.0171	100.0	100.0	100.0	100.0	100.0	100.0	--
68:02:00	11	0.0377	100.0	100.0	100.0	100.0	100.0	100.0	--
74:01:00	10	0.0342	100.0	98.3	90.0	98.6	69.2	99.6	03:01 (50)
74:10:00	4	0.0137	100.0	98.6	0.0	100.0	--	98.6	74:01 (100)
80:01:00	4	0.0137	100.0	100.0	100.0	100.0	100.0	100.0	--
HLA-B									
7:02	17	0.0582	100.0	94.2	94.1	94.2	50.0	99.6	40:01 (50)
8:01	19	0.0651	100.0	95.9	36.8	100.0	100.0	95.8	40:01 (45)
8:01	1	0.0034	100.0	99.7	0.0	100.0	--	99.7	40:01 (100)
13:02	4	0.0137	100.0	99.0	25.0	100.0	100.0	99.0	15:10 (67)
14:01	1	0.0034	100.0	99.7	100.0	99.7	50.0	100.0	--
14:02	10	0.0342	100.0	99.3	90.0	99.6	94.7	99.6	15:10 (50)
14:03	1	0.0034	100.0	99.7	0.0	100.0	--	99.7	14:02 (50)
15:01	4	0.0137	100.0	97.9	100.0	97.9	38.1	100.0	--

15:03	17	0.0582	100.0	95.5	29.4	99.6	83.3	95.8	07:02 (92)
15:04	2	0.0068	100.0	99.3	0.0	100.0	--	99.3	15:01 (100)
15:10	6	0.0205	100.0	97.6	50.0	98.6	42.9	98.9	15:01 (50)
15:16	5	0.0171	100.0	99.3	80.0	99.7	88.9	99.7	35:01 (100)
15:17	3	0.0103	100.0	100.0	100.0	100.0	100.0	100.0	--
15:18	1	0.0034	100.0	100.0	100.0	100.0	100.0	100.0	--
15:20	3	0.0103	100.0	99.0	0.0	100.0	--	99.0	15:01 (100)
18:01	7	0.0240	100.0	99.0	85.7	99.3	73.3	99.6	42:01 (50)
35:01:00	17	0.0582	100.0	94.9	82.4	95.6	52.8	98.9	35:02 (100)
35:02:00	6	0.0205	100.0	98.6	100.0	98.6	60.0	100.0	--
35:03:00	8	0.0274	100.0	97.6	37.5	99.3	60.0	98.3	35:01 (80)
35:05:00	2	0.0068	100.0	99.3	0.0	100.0	--	99.3	35:01 (100)
35:20:00	1	0.0034	100.0	99.7	0.0	100.0	--	99.7	35:01 (50)
37:01:00	5	0.0171	100.0	98.3	0.0	100.0	--	98.3	07:02 (100)
38:01:00	6	0.0205	100.0	100.0	100.0	100.0	100.0	100.0	--
39:05:00	3	0.0103	100.0	99.0	0.0	100.0	--	99.0	35:01 (50)
39:06:00	1	0.0034	100.0	99.7	0.0	100.0	--	99.7	... (100)
39:10:00	2	0.0068	100.0	99.3	0.0	100.0	--	99.3	... (50)
40:01:00	1	0.0034	100.0	97.6	100.0	97.6	11.8	100.0	--
40:04:00	1	0.0034	100.0	99.7	0.0	100.0	--	99.7	... (100)
41:01:00	8	0.0274	100.0	97.9	75.0	98.6	60.0	99.3	15:10 (25)
41:02:00	1	0.0034	100.0	99.7	0.0	100.0	--	99.7	41:01 (100)
42:01:00	19	0.0651	100.0	96.2	84.2	97.1	64.6	98.9	15:10 (33)
44:02:00	12	0.0411	100.0	99.7	100.0	99.6	92.3	100.0	--
44:03:00	16	0.0548	100.0	98.3	100.0	98.2	77.1	100.0	--
45:01:00	28	0.0959	100.0	95.2	50.0	100.0	100.0	95.0	42:01 (48)
48:02:00	1	0.0034	100.0	100.0	100.0	100.0	100.0	100.0	--
49:01:00	8	0.0274	100.0	99.3	87.5	99.6	93.3	99.6	44:03 (100)
50:01:00	2	0.0068	100.0	99.3	100.0	99.3	44.4	100.0	--
50:02:00	1	0.0034	100.0	99.7	0.0	100.0	--	99.7	44:03 (100)
51:01:00	11	0.0377	100.0	97.3	63.6	98.6	63.6	98.6	52:01 (50)
52:01:00	7	0.0240	100.0	96.6	28.6	98.2	30.8	98.2	51:01 (80)
53:01:00	9	0.0308	100.0	99.3	77.8	100.0	100.0	99.3	35:01 (50)
57:01:00	4	0.0137	100.0	100.0	100.0	100.0	100.0	100.0	--
57:03:00	2	0.0068	100.0	99.7	100.0	99.7	80.0	100.0	--
58:01:00	6	0.0205	100.0	99.3	66.7	100.0	100.0	99.3	15:03 (50)

81:01:00	3	0.0103	100.0	99.7	66.7	100.0	100.0	99.7	14:01 (100)
HLA-C									
1:02	4	0.0139	100.0	100.0	100.0	100.0	100.0	100.0	--
2:02	8	0.0278	100.0	96.2	87.5	96.4	41.2	99.6	07:02 (100)
2:10	22	0.0764	100.0	94.1	22.7	100.0	100.0	94.0	02:02 (59)
3:02	3	0.0104	100.0	100.0	100.0	100.0	100.0	100.0	--
3:03	4	0.0139	100.0	97.9	100.0	97.9	40.0	100.0	--
3:04	30	0.1042	100.0	96.5	76.7	98.8	88.5	97.3	03:03 (71)
4:01	46	0.1597	100.0	98.6	100.0	98.3	92.0	100.0	--
4:13	1	0.0035	100.0	99.7	0.0	100.0	--	99.7	04:01 (100)
5:01	14	0.0486	100.0	100.0	100.0	100.0	100.0	100.0	--
6:02	9	0.0312	100.0	100.0	100.0	100.0	100.0	100.0	--
7:01	23	0.0799	100.0	99.7	95.7	100.0	100.0	99.6	07:02 (50)
7:02	17	0.0590	100.0	97.9	100.0	97.8	72.3	100.0	--
7:04	2	0.0069	100.0	100.0	100.0	100.0	100.0	100.0	--
8:02	13	0.0451	100.0	99.3	92.3	99.6	92.3	99.6	03:04 (100)
8:04	3	0.0104	100.0	99.7	66.7	100.0	100.0	99.7	03:04 (50)
12:02	2	0.0069	100.0	100.0	100.0	100.0	100.0	100.0	--
12:03	10	0.0347	100.0	100.0	100.0	100.0	100.0	100.0	--
14:02	5	0.0174	100.0	100.0	100.0	100.0	100.0	100.0	--
15:02	5	0.0174	100.0	98.3	100.0	98.2	50.0	100.0	--
15:04	1	0.0035	100.0	99.7	0.0	100.0	--	99.7	15:02 (100)
15:39	3	0.0104	100.0	99.0	0.0	100.0	--	99.0	15:02 (100)
16:01	21	0.0729	100.0	98.6	90.5	99.3	90.5	99.3	03:03 (50)
16:02	1	0.0035	100.0	99.7	0.0	100.0	--	99.7	15:02 (100)
17:01	39	0.1354	100.0	99.3	94.9	100.0	100.0	99.2	03:04 (50)
18:01	2	0.0069	100.0	100.0	100.0	100.0	100.0	100.0	--
HLA-DRB1									
1:01	14	0.0479	100.0	100.0	100.0	100.0	100.0	100.0	--
1:02	23	0.0788	100.0	99.7	95.7	100.0	100.0	99.6	... (100)
3:01	11	0.0377	100.0	97.9	81.8	98.6	66.7	99.3	12:01 (50)
3:02	17	0.0582	100.0	100.0	100.0	100.0	100.0	100.0	--
4:01	4	0.0137	100.0	99.0	75.0	99.3	60.0	99.7	07:01 (50)
4:02	3	0.0103	100.0	100.0	100.0	100.0	100.0	100.0	--
4:04	2	0.0068	100.0	99.7	50.0	100.0	100.0	99.7	14:01 (50)
4:05	1	0.0034	100.0	99.7	100.0	99.7	66.7	100.0	--
4:11	3	0.0103	100.0	99.0	0.0	100.0	--	99.0	... (67)
7:01	34	0.1164	100.0	98.6	91.2	99.6	98.4	98.8	10:01 (33)

8:01	3	0.0103	100.0	100.0	100.0	100.0	100.0	100.0	--
8:02	9	0.0308	100.0	99.3	100.0	99.3	85.7	100.0	--
8:04	25	0.0856	100.0	96.2	60.0	99.6	93.8	96.4	11:01 (30)
9:01	1	0.0034	100.0	99.7	0.0	100.0	--	99.7	04:01 (100)
10:01	3	0.0103	100.0	99.7	100.0	99.7	75.0	100.0	--
11:01	27	0.0925	100.0	94.2	92.6	94.3	61.7	99.2	03:01 (50)
11:02	9	0.0308	100.0	97.6	33.3	99.6	75.0	97.9	14:01 (83)
11:04	7	0.0240	100.0	97.6	0.0	100.0	--	97.6	11:01 (100)
12:01	11	0.0377	100.0	95.9	27.3	98.6	46.2	97.2	13:01 (100)
12:02	2	0.0068	100.0	98.3	0.0	99.0	0.0	99.3	03:01 (100)
13:01	15	0.0514	100.0	96.2	86.7	96.8	59.1	99.3	15:01 (100)
13:02	7	0.0240	100.0	99.3	85.7	99.6	85.7	99.6	13:01 (100)
13:04	4	0.0137	100.0	98.6	0.0	100.0	--	98.6	11:01 (25)
13:05	1	0.0034	100.0	99.7	0.0	100.0	--	99.7	11:01 (100)
14:01	4	0.0137	100.0	97.9	75.0	98.3	35.3	99.6	03:01 (100)
14:02	3	0.0103	100.0	100.0	100.0	100.0	100.0	100.0	--
14:06	1	0.0034	100.0	100.0	100.0	100.0	100.0	100.0	--
15:01	13	0.0445	100.0	96.9	100.0	96.8	55.3	100.0	--
15:03	31	0.1062	100.0	97.6	80.6	99.6	96.2	97.7	15:01 (100)
16:02	4	0.0137	100.0	99.0	25.0	100.0	100.0	99.0	11:01 (100)

[1]: the HLA alleles with more than one copy and non-zero sensitivity in the training are listed.[2]: CR – call rate.[3]: ACC – allele accuracy.[4]: the most likely miscalled allele and the proportion of the most likely miscalled allele in all miscalled alleles.

S11. Characteristics of HLA-B alleles with low sensitivity, among imputations models.

HLA-B	Geographic region	NCHR	1000g		UW		1000g+QUI	
			SEN	PPV	SEN	PPV	SEN	PPV
14:03	AFR	1	0	0	0	0	0	0
15:04	NAM	2	0	0	0	0	0	0
15:20	NAM	3	0	0	0	0	100	100
35:05	NAM	2	0	0	0	0	0	0
35:20	NAM/MDE	1	0	0	0	0	0	0
37:01	EAS	5	0	0	0	0	100	100
39:06	NAM	1	100	100	0	0	100	100
39:10	AFR	2	0	0	0	0	0	0
40:04	NAM	1	0	0	0	0	0	0
41:02	AFR/EUR	1	0	0	0	0	0	0
50:02	AFR	1	0	0	0	0	0	0
57:03	AFR	2	0	0	100	100	100	100

52:01	COSM	7	28.6	100	28.6	98.2	50	98.2
35:03	EUR	8	37.5	75	37.5	99.3	80	100