Table S2. Number of known and total predictions for each HLA allele determined by employing anchor motifs only, the phylo-HMMwD and phylo-HMMwD+M. Predictions which overlapped with epitopes documented in Yusim et al. (2009) were regarded as known. Under all three methods, predicted epitopes were allowed to be between 8 and 11 amino acids in length, satisfying the anchor motifs given in supplementary table S3. Although flanking states were used to identify epitopes under the phylo-HMMwD+M, they were not considered as part of the predicted epitopes in order to ensure comparability between methods.

	Anchors Only		Phylo-HMMwD		Phylo-HMMwD+M	
HLA allele	Known	Total	Known	Total	Known	Total
A*02	11	12	0	0	2	2
A*23	1	6	1	4	1	4
A*29	1	10	0	3	1	3
A*30	3	12	2	6	2	5
A*3001	3	12	1	3	1	3
A*3002	2	4	1	5	1	3
A*68	3	12	1	8	0	4
A*6802	3	10	1	8	0	3
A*0301	5	11	2	2	2	3
B*0702	5	12	0	0	1	3
B*0801	6	11	3	5	2	4
B*15	5	12	1	5	1	3
B*1503	2	3	0	1	0	1
B*1510	2	7	1	6	2	3
B*42	3	5	1	2	3	4
B*4201	3	5	1	2	3	4
B*44	3	8	2	3	1	2
B*4403	3	8	0	1	1	2
B*57	7	7	3	6	3	3
B*5801	6	7	1	1	3	3
B*5802	1	1	2	7	1	1
B*81	1	5	1	4	1	1
C*02	0	9	0	0	0	1
C*0202	0	9	0	0	0	2
C*04	1	13	2	10	1	5
C*06	1	3	0	5	0	6
C*0602	1	3	0	4	0	5
C*07	2	15	0	7	1	10
C*08	3	12	1	6	0	2
C*17	0	9	0	2	0	0
C*1701	0	9	0	4	0	0
C*18	2	13	1	5	0	6
Total[a]	63	191	21	81	24	65

[a] For HLA alleles reported at two- and four-digit resolution, only the number of predictions for the two-digit allele was used to calculate the totals. For HLA alleles with only a single anchor residue, that is HLA-B*0801, HLA-C*06 and HLA-C*0602, the union of the predictions inferred with anchor motifs only encompassed nearly the entire genomic region under consideration. Consequently, these predictions were also excluded from the totals.