

表 1. LOAO 交差検証テストに用いた各データセットにおけるアレルタイプ別配列数

アレルタイプ	Dataset		
	BD2016 <sup>※1</sup>	IEDB2022 <sup>※2</sup>	PDB-HB
HLA-DR	25	17	7
HLA-DQ	20	2	1
HLA-DP	9	0	0

※1 BD2016 の学習にはペプチド数が 10 以上のアレルに対して LOAO を実行

※2 IEDB2022 の学習にはペプチド数が 1000 以上のアレルに対して LOAO を実行

※3 PDB-HB の学習には BD2016 に含まれる同一のアレルに対して LOAO を実行

表 2. LOAO 交差検証テストに用いた各データセットにおけるアレルタイプ別の正例と負例ペプチド数

アレルタイプ	BD2016 <sup>※1</sup>		IEDB2022 (500 [nM]) <sup>※1</sup>		IEDB2022 (1000 [nM]) <sup>※2</sup>	
	正例	負例	正例	負例	正例	負例
HLA-DR	87,363	40,756	22,177	32,195	26,647	26,825
HLA-DQ	15,564	5,135	3,753	9,294	7,978	5,067
HLA-DP	28,081	9,098	801	3,771	3,433	1,139

※1 正例 :  $IC_{50} < 500$  [nM]、負例 :  $IC_{50} \geq 500$  [nM]

※2 正例 :  $IC_{50} < 1000$  [nM]、負例 :  $IC_{50} \geq 1000$  [nM]

図1. DeepSeqPanIIのモデルアーキテクチャ (出典: Liu et al. (2022) IEEE/ACM Transactions on Computational Biology and Bioinformatics)

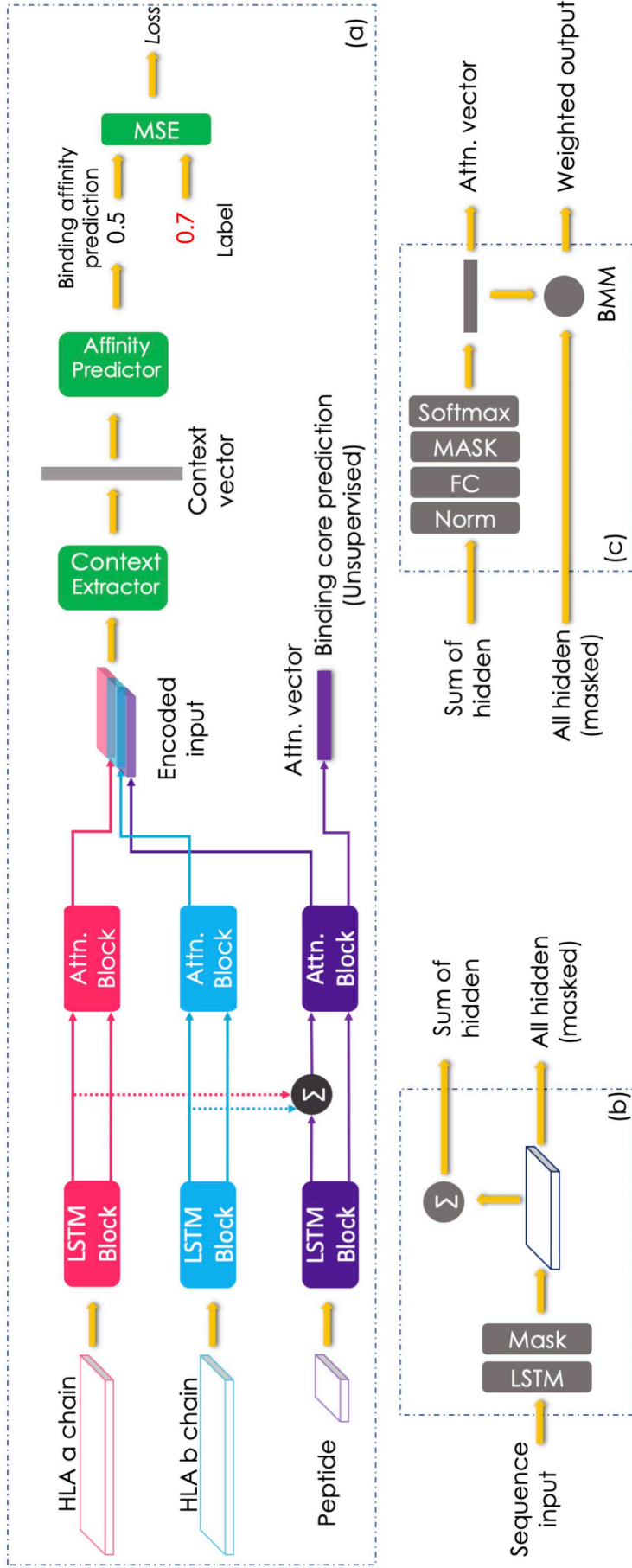


図2. 「IEDB2022データセット」構築フロー

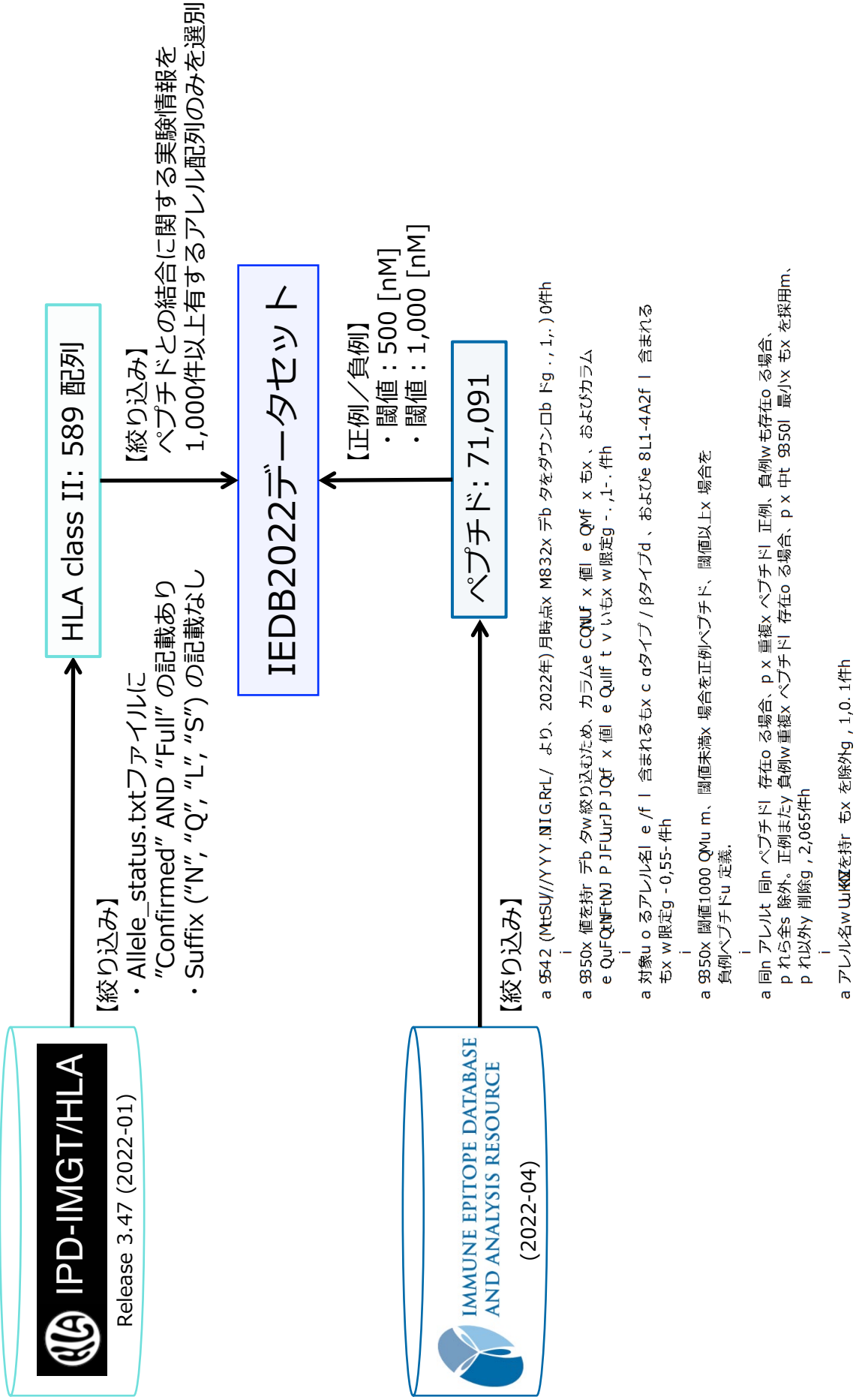


図3. 「PDB-HBデータセット」構築フロー

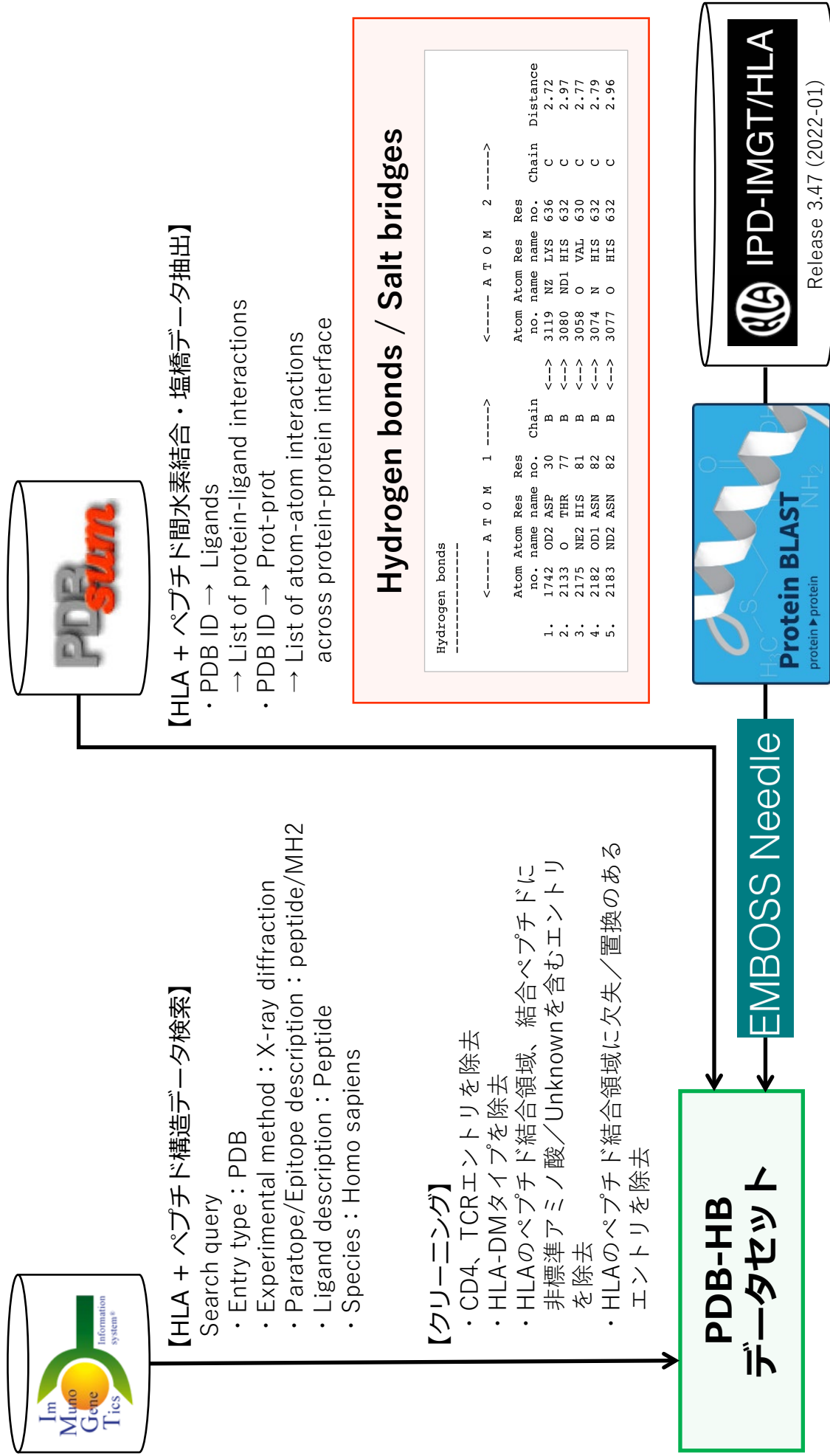


図4. 「BD2016データセット」に対するLOAO交差検証テストの結果（結合ペプチド数10以上のアレル：54）



図5. 「IEDB2022データセット」に対するLOAO交差検証テストの結果（結合ペプチド数1000以上のアレル：19）

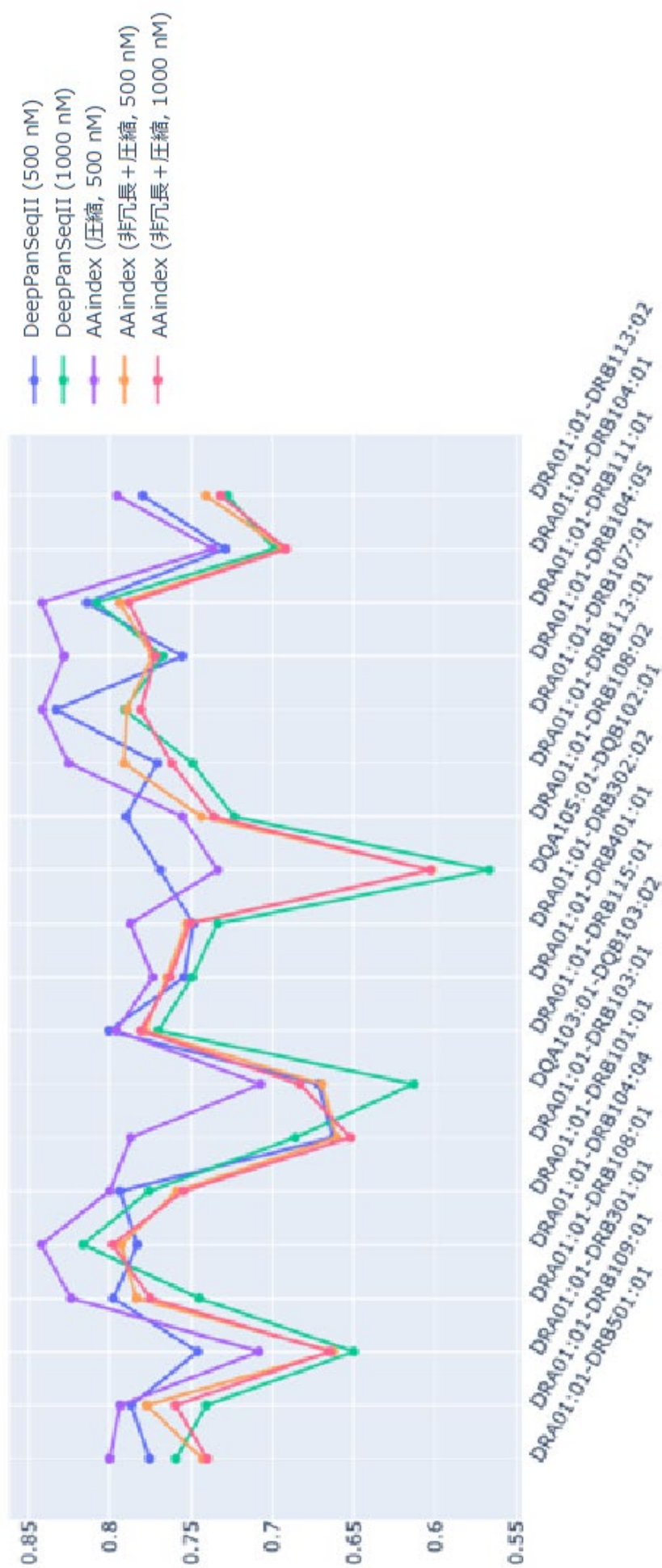


図6. 「PDB-HBデータセット」に対するLOAO交差検証テストの結果（結合ペプチド数1000以上のアレル：8）

