

1 **Title page**

2 *Original paper*

3 **Title**

4 Associations of functional HLA class I groups with HIV viral load in a heterogeneous cohort.

5 **Running head**

6 Functional HLA clusters and HIV-1 viral load

7 **Authors**

8 Adrian G. ZUCCO<sup>1</sup>, Marc BENNEDBÆK<sup>2</sup>, Christina EKENBERG<sup>1</sup>, Migle GABRIELAITE<sup>3</sup>, Preston LEUNG<sup>1</sup>, Mark N.  
9 POLIZZOTTO<sup>4</sup>, Virginia KAN<sup>5</sup>, Daniel D. MURRAY<sup>1</sup>, Jens D. LUNDGREN<sup>1</sup> and Cameron R. MACPHERSON<sup>1</sup> for  
10 the INSIGHT START study group.

11 **Affiliations**

12 <sup>1</sup>PERSIMUNE Center of Excellence, Rigshospitalet, Copenhagen, Denmark

13 <sup>2</sup>Virus Research and Development Laboratory, Virus and Microbiological Special Diagnostics, Statens Serum  
14 Institut, Copenhagen, Denmark

15 <sup>3</sup>Center for Genomic Medicine, Copenhagen University Hospital, Copenhagen, Denmark.

16 <sup>4</sup>Clinical Hub for Interventional Research, College of Health and Medicine, The Australian National University,  
17 Canberra, Australia

18 <sup>5</sup>George Washington University, Veterans Affairs Medical Center, Washington D.C, U.S.A.

19 **Corresponding author**

20 Adrian Gabriel Zucco, MSc., PhD.

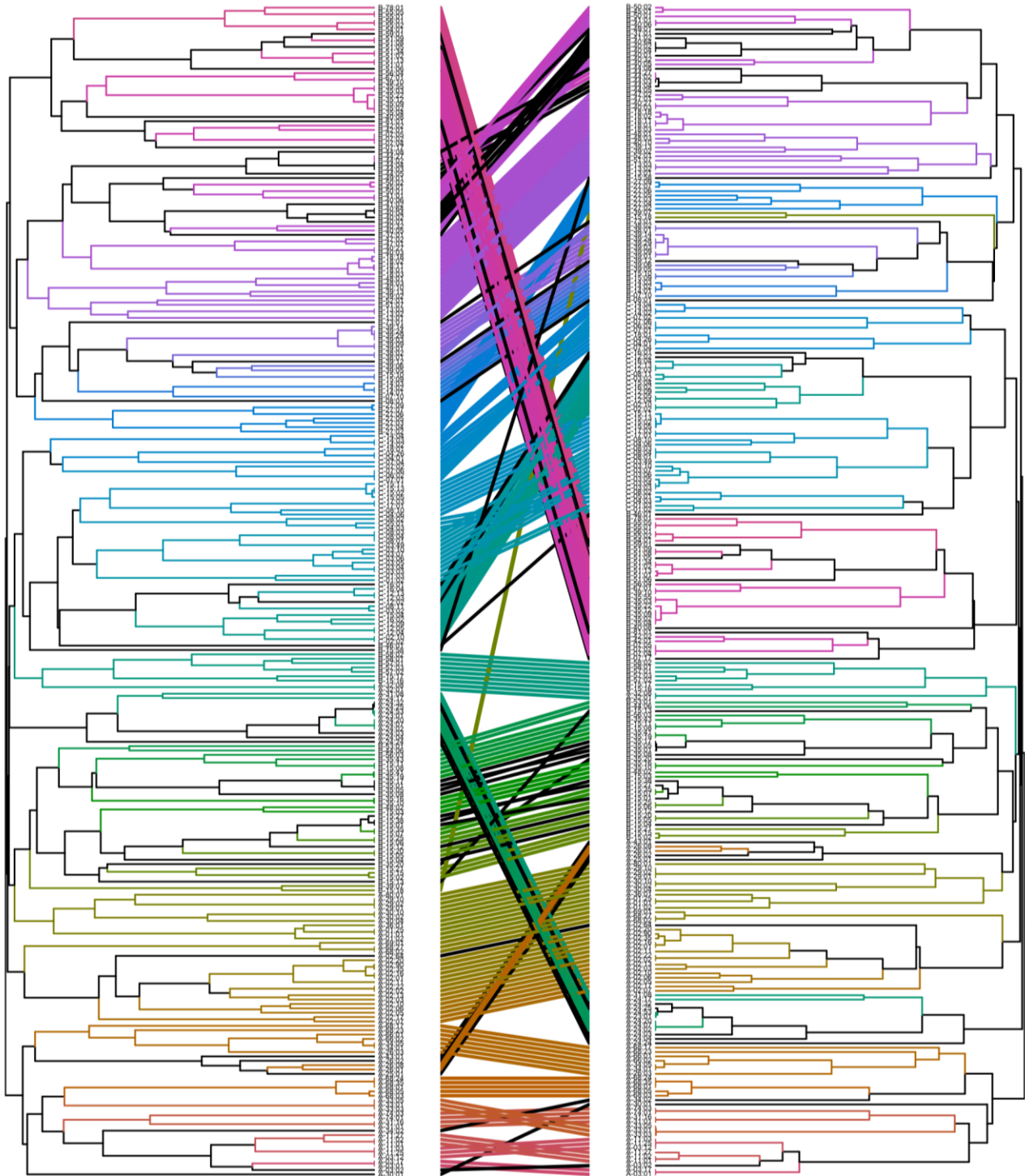
21 Tel: +45 35 45 57 75

22 mail: [adrian.gabriel.zucco@regionh.dk](mailto:adrian.gabriel.zucco@regionh.dk)

23 Rigshospitalet, Copenhagen University Hospital

24 Centre of Excellence for Health, Immunity and Infections (CHIP) & PERSIMUNE

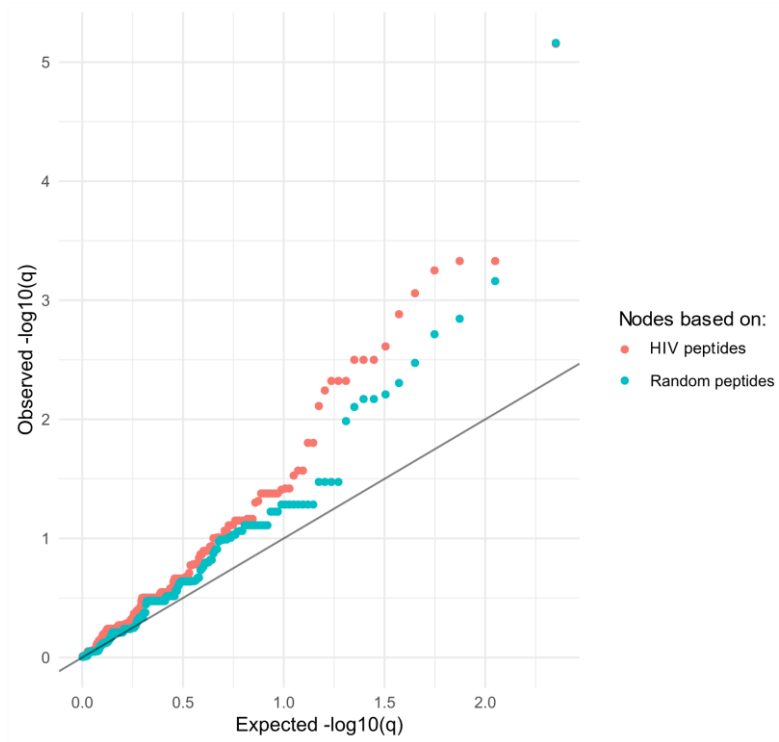
25 Blegdamsvej 9, DK-2100 Copenhagen Ø, Denmark



26

27 **Figure S1. Tanglegram of consensus clustering from predicted immunopeptidomes based on**  
 28 **random versus HIV-specific peptides.**

29 Two different peptide sets were used for consensus clustering based on predicted immunopeptidomes to  
 30 268 HLA class I alleles. On the left, a dendrogram generated from  $5 \times 10^5$  random peptides is compared to a  
 31 dendrogram generated from 173,792 HIV peptides. Black branches and lines connecting both dendrograms  
 32 indicate differences in clustering among both dendrograms.

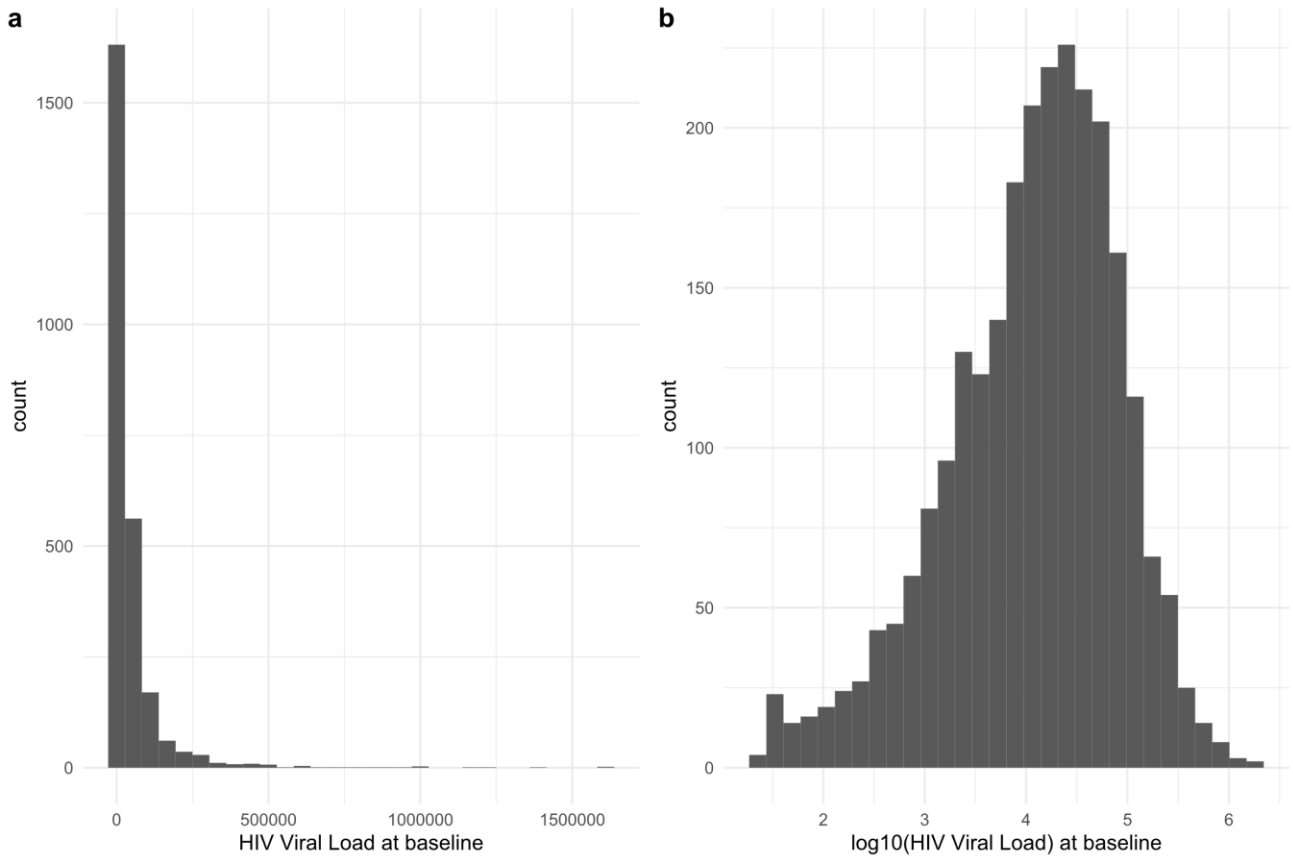


33

34 **Figure S2. Q-Q plot of observed versus theoretical q-values from associations to HIV-VL from HLA**  
 35 **functional nodes**

36 HLA functional nodes were generated by consensus clustering of predicted HIV-specific immune peptidomes  
 37 (red) and unspecific predicted immunopeptidomes from random peptides (blue).

38



39

40 **Figure S3. Histograms of HIV viral load**

41 Histogram of HIV viral load (a) and log<sub>10</sub>-transformed HIV viral load (b) measured once in participants of the  
42 START study at entry.

43

Alleles in the tested node	Adjustment by sex, country, and race		Adjustment by sex and PC1-4	
	Estimate	q-value	Estimate	q-value
B-57:03   B-57:02   B-57:01   B-58:01	-0.2502	7.0217E-06	-0.2570	3.6756E-06
C-08:01   C-08:04	-0.2863	4.1949E-02	-0.2600	7.7905E-02
B-44:05   B-44:08   B-44:04   B-44:03   B-44:02   B-44:27	0.1483	3.1637E-03	0.1549	2.0194E-03
B-56:03   B-35:43   B-15:08   B-15:11   B-35:08   B-35:19   B-35:41   B-35:01   B-35:17   B-35:05   B-35:20   B-35:16   B-35:10   B-15:13   B-53:01   B-44:06	0.1324	4.8579E-02	0.1188	8.1062E-02

44 **Table S1. Estimates and q-values of a sensitivity analysis comparing the use of self-reported race**  
45 **and country versus Principal Components for adjustment of population stratification in**  
46 **association testing of relevant HLA functional nodes to log<sub>10</sub>-transformed HIV VL**

47 Associations to log<sub>10</sub>-transformed HIV VL with each HLA functional node in the consensus tree were tested  
48 and adjusted by sex, country and self-reported race or sex and the first four principal components computed  
49 by Principal Component Analysis (PCA) excluding the HLA region. The estimate corresponds to the  $\beta$   
50 parameter of the fitted linear models and q-values refer to the adjusted p-values for multiple testing using a  
51 Benjamini-Hochberg procedure.