

A composite immune signature parallels disease progression across T1D subjects

Cate Speake¹, Samuel O. Skinner¹, Dror Berel², Elizabeth Whalen^{1@}, Matthew J. Dufort¹, William Chad Young², Jared M. Odegard^{1*}, Anne M. Pesenacker^{3†}, Frans K. Gorus⁴, Eddie A. James¹, Megan K. Levings³, Peter S. Linsley¹, Eitan M. Akirav^{5,6}, Alberto Pugliese⁷, Martin J. Hessner⁸, Gerald T. Nepom^{1,9}, Raphael Gottardo², S. Alice Long¹

¹ Benaroya Research Institute at Virginia Mason, Seattle, WA

² Vaccines and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, WA

³ University of British Columbia BC Children's Hospital Research Institute, Vancouver, BC, Canada

⁴ Medical School and University Hospital (UZ Brussel), Vrije Universiteit Brussel, Brussels, Belgium

⁵ Research Institute, Islet Biology, New York University Winthrop Hospital, Mineola, New York

⁶ Stony Brook University School of Medicine, Stony Brook, New York.

⁷ Diabetes Research Institute, Department of Medicine, Division of Diabetes Endocrinology and Metabolism, and Department of Microbiology and Immunology, Miller School of Medicine, University of Miami, Miami, FL

⁸ Medical College of Wisconsin, Milwaukee, WI

⁹ Immune Tolerance Network, Bethesda, MD

@ Present affiliation: Celgene, Inc.

* Present affiliation: Gilead Sciences, Inc., Seattle, WA

† Present affiliation: University College London, London, UK

Corresponding Author:

Cate Speake, PhD
Diabetes Clinical Research Program
Benaroya Research Institute at Virginia Mason
1201 9th Ave
Seattle, WA 98101
Ph: (206) 287-1078
cspeake@benaroyaresearch.org

SUPPLEMENTAL METHODS

RV144 Validation

We used data from RV144, a previously published HIV vaccine efficacy trial (1) to compare the results of applying Elastic Net analysis on all available data versus performing a down-selection step prior to variable selection. The dataset we used includes the six primary immune assay variables used in the original study (1), the polyfunctional T-cell response identified in (2), and RNA-seq data published in (3), all measured on a set of 172 individuals. We built a multivariate model based on each of these datasets that predicts HIV infection post vaccination. Like the recent onset T1D dataset, the potential covariate pool in the RV144 dataset is dominated by RNAseq data, with 47,323 measurements per individual. When applying Elastic Net naively to this dataset, only RNAseq variables were selected. In contrast, we performed down-selection by obtaining univariate p-values for each variable in glm and then adjusting the p-values for multiple testing to control FDR. None of the RNAseq variables passed this step, and the previously described primary immune correlates (1) and COMPASS polyfunctionality score (2) were not lost in the volume of RNAseq data. Using the DIFAcTO tool identified the same three important variables previously shown to be predictive of HIV infection risk, namely IgG antibodies binding to V1V1, Env-specific IgA antibodies, and COMPASS polyfunctionality score.

Figure S1

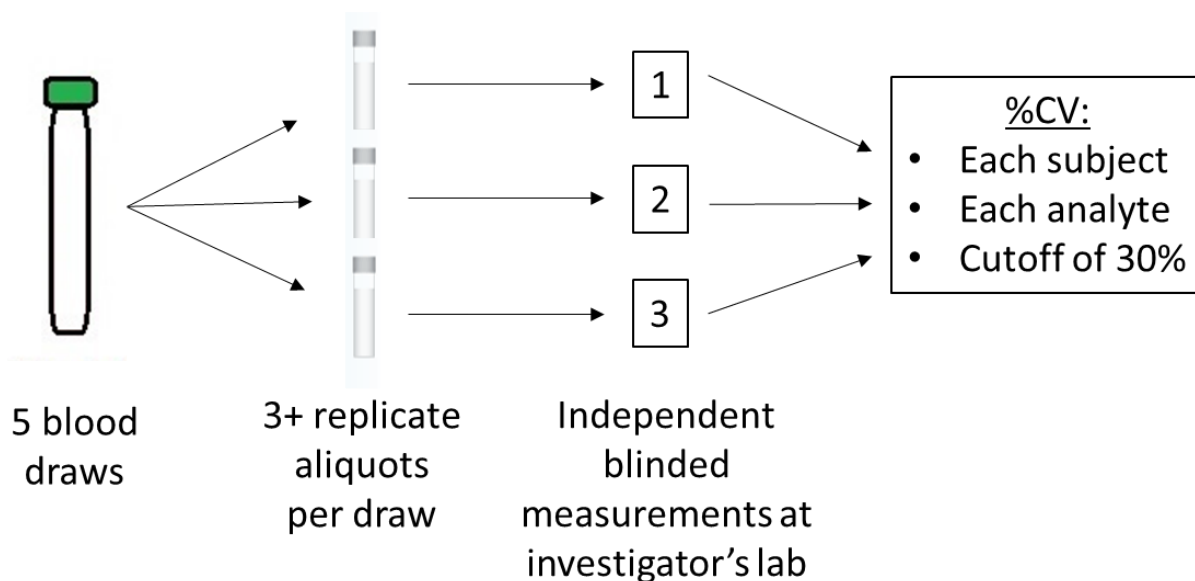


Figure S1: Replicate testing process for this study. Typically, 5 subjects were included in each replicate testing experiment. Each subject's sample was split into 3 aliquots; these were blinded and tested at the participating laboratory. The data were returned to BRI for analysis, where unblinding and coefficient of variation (% CV) calculations took place. % CV is calculated per subject, and then a mean % CV is calculated. CV cutoff for this study was 30%.

Figure S2

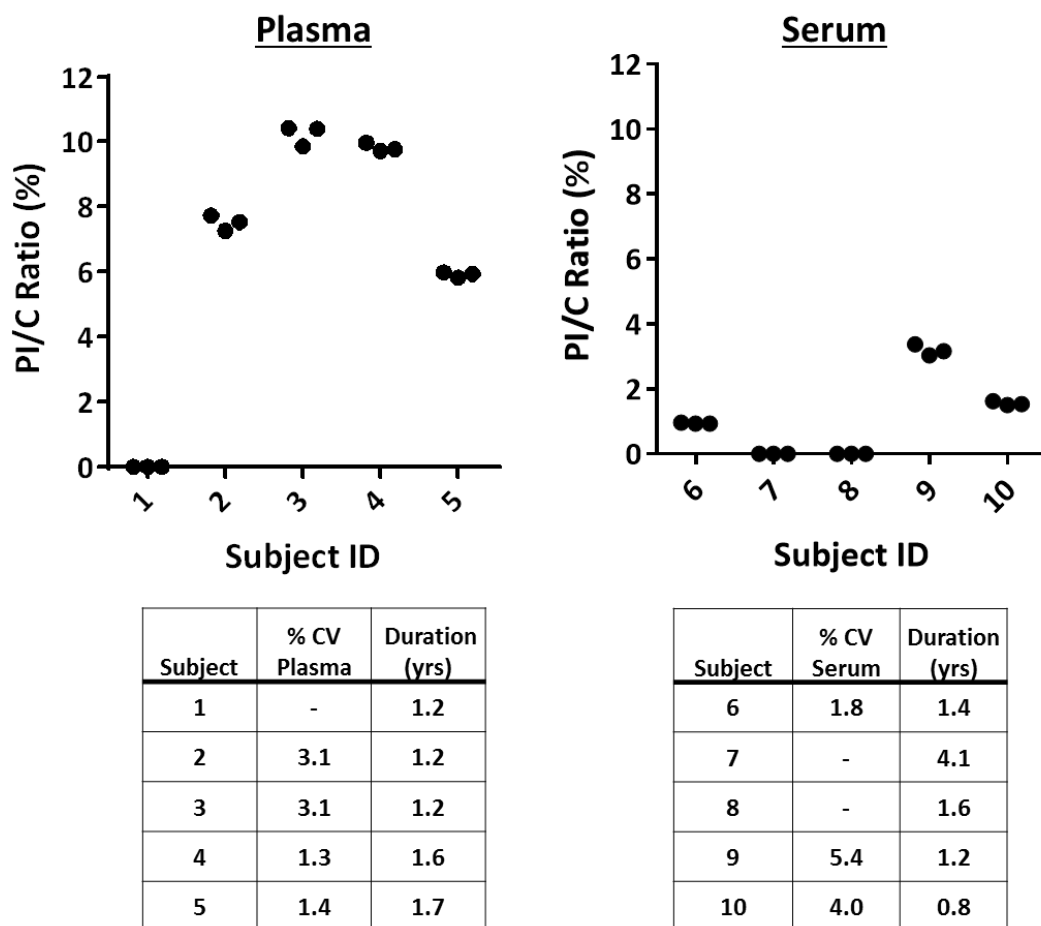


Figure S2: Blinded replicate testing data: proinsulin/C-peptide ratio assay. Each dot represents the proinsulin/C-peptide (PI/C) ratio for 1 aliquot from 1 subject; the data are grouped by subject.

Coefficient of variation (% CV) for each subject derived from the values in the plots are listed in the table below. % CV was 0 for 3 subjects with no residual C-peptide secretion (subjects 1, 7, and 8). Different subjects were used for the plasma (left) and serum (right) studies.

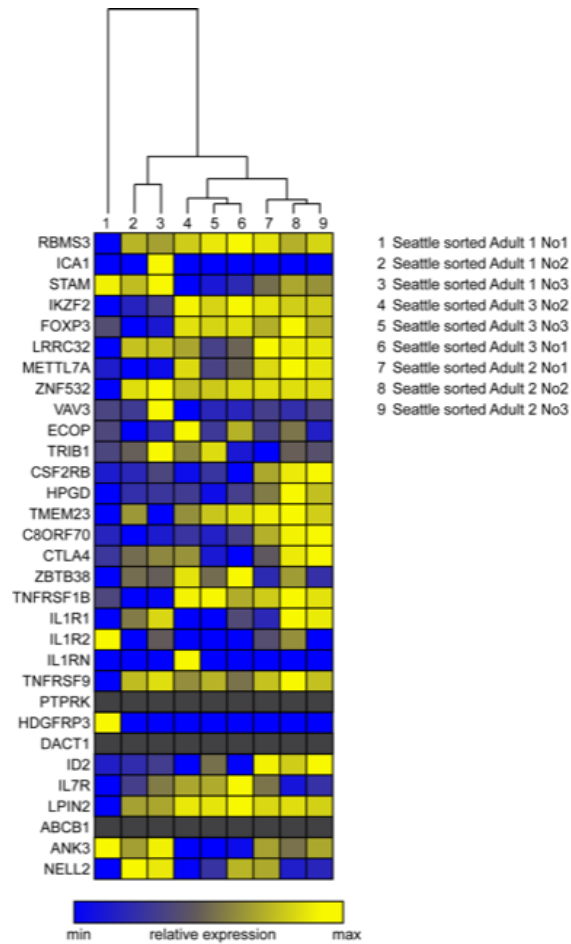
Figure S3

Figure S3: Blinded replicate testing data for Treg transcriptional signature assay. Treg transcriptional results cluster by subject. Each row represents one transcript measured by Nanostring assay; each column represents one aliquot. Legend indicates the subject ID (Adult 1-3) and aliquot ID (No 1-3). Analysis used Pearson average clustering, expression coloring is relative according to row min-max.

Figure S4

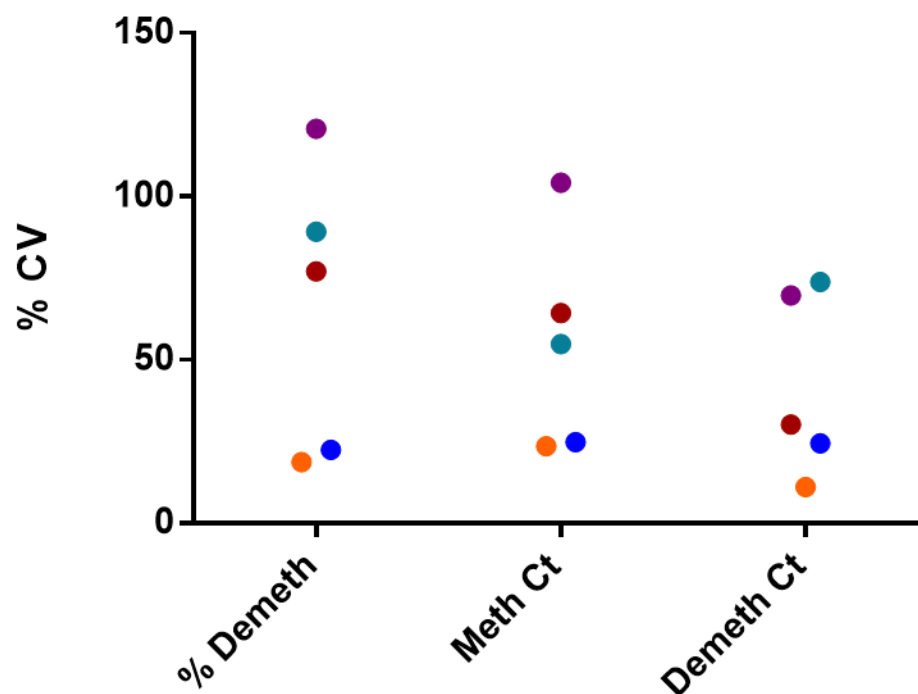


Figure S4: Blinded replicate testing data for demethylated insulin assay. Each dot represents the coefficient of variation (CV) value expressed as a percentage for 1 new onset subject; the CV is derived from the results for 3 replicate aliquots. Dots are colored by subject. Methylated INS gene cycle threshold (Meth Ct), Demethylated INS gene cycle threshold (Demeth Ct), and % of total counts that are demethylated INS gene detected (% Demeth) are included.

Figure S5

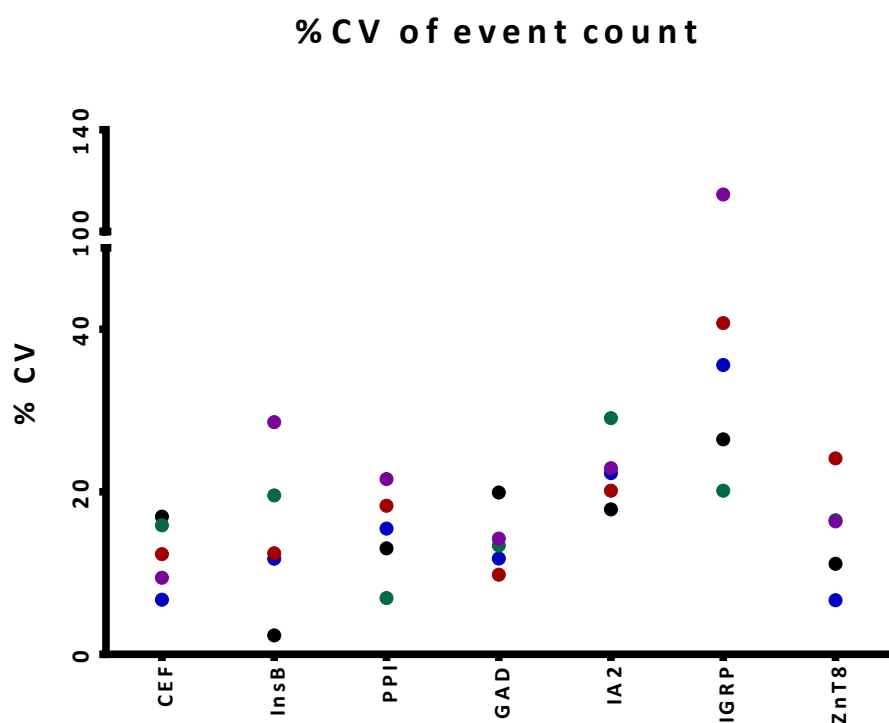


Figure S5: Blinded replicate testing data for antigen specific CD8 T cell qDot assay. Each dot represents the coefficient of variation (CV) value expressed as a percentage for 1 subject; the CV is derived from the event count results (cells/million) for 3 replicate aliquots from 5 individual subjects. Dots are colored by subject. X axis is the antigen tested in qDot assay. All antigens had mean CV <30% except for IGRP.

Figure S6

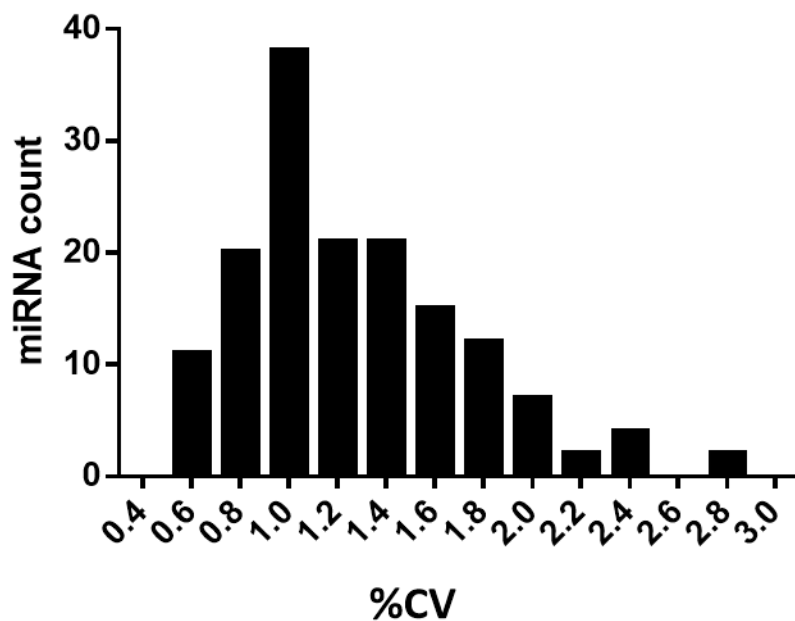


Figure S6: Blinded replicate testing data for miRNA assay (Exiqon). The number of miRNA (miRNA count) at a specific mean coefficient of variation (%CV) value, range 0 – 100, was determined by testing 3 replicate aliquots per subject from 5 T1D subjects. %CV was calculated for each subject, and then a mean %CV calculated for the 5 subjects. Plot includes the 153 miRNA that were detectable in 5/5 subjects (of 752 miRNA measured). Detectable is defined as cycle threshold (Ct) value ≤ 37 , as recommended by the manufacturer (Exiqon).

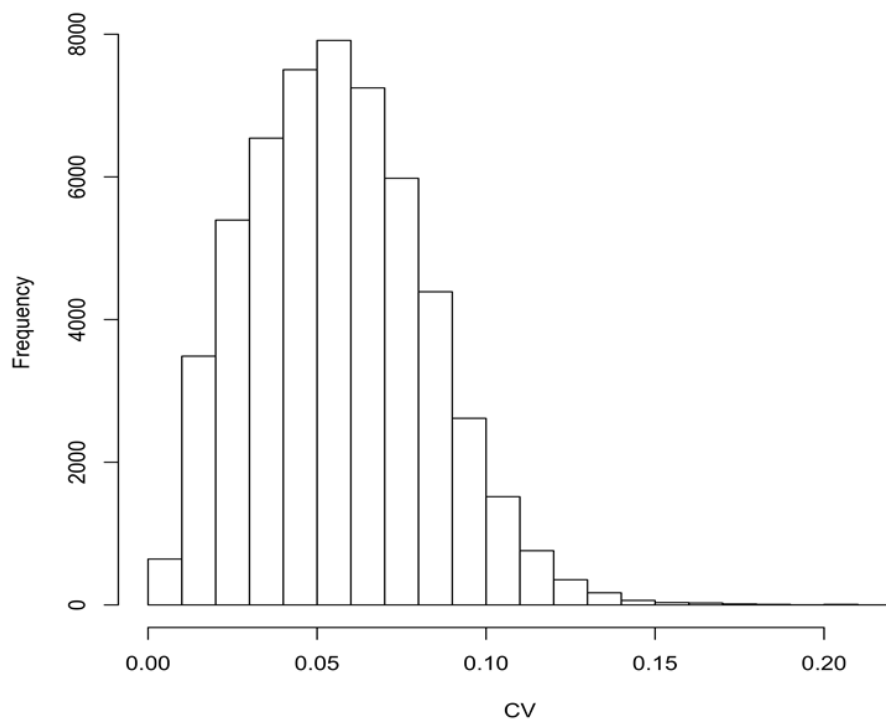
Figure S7

Figure S7: Blinded replicate testing data for transcriptional signature of serum exposure assay. The number of transcripts (Y-axis) in a mean coefficient of variation (CV) bin, maximum range 0-1, was determined by testing 3 replicate aliquots per subject from 5 T1D subjects. CV was calculated for each subject, and then a mean CV calculated for the 5 subjects. These data are not expressed as a percentage. Plot includes all probes present on the Affymetrix U137Plus2 chip (54675 probes). All transcripts/probes on this microarray had a CV < 0.3; CV range for all transcripts/probes was 0.0006 to 0.2128.

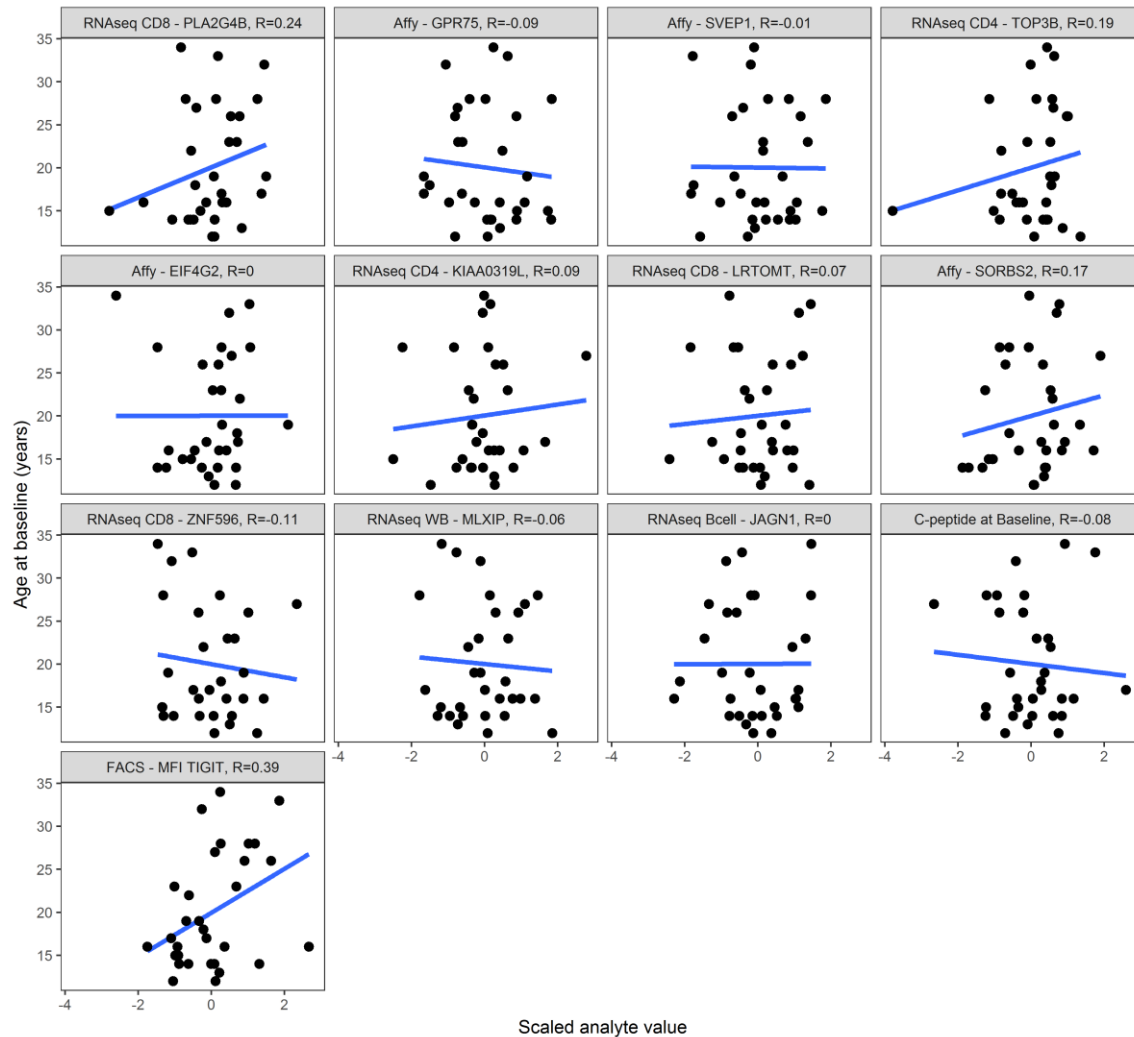


Figure S8: Individual correlations between each selected analyte and age at enrollment. Immune markers were measured at trial enrollment (within 90 days of diagnosis, n=30 subjects), and Y axis indicates subject age at enrollment. Each mini-plot uses the scaled value for the analyte on the X-axis. Pearson correlation values are listed at the top of each mini-plot; mini-plots are ordered by absolute correlation value with C-peptide decline to aid comparison with Figure 5. Regression lines in blue. Note that in this dataset, only one immune parameter has a reasonable correlation with age (MFI TIGIT, $r=0.39$). Assays and analyte names are truncated; full names can be found in Table 3. Affy indicates the transcriptional response to T1D serum assay as this is conducted on the Affymetrix platform.

Table S1: Prediction of C-peptide decline using selected analytes. * ‘Baseline’ model includes C-peptide at diagnosis. ‘Full’ model includes all 17 identified analytes. ‘Maintained’ model includes only the 12 analytes that were found to be robust to tool settings. † Cross-validated RMSE calculated by separating the data into five folds and for each fold training on the other four-fifths of the data and predicting the held-out fold. This was repeated 1000 times to get a robust estimate of the RMSE as well as a 90% interval showing the variability due to different partitions of the data.

Model*	Adjusted R ²	Cross-validated RMSE (90% interval)†
Baseline	0.071	0.00170 (0.00152, 0.00191)
Full	0.979	0.00078 (0.00053, 0.00122)
Maintained	0.981	0.00037 (0.00030, 0.00046)

SUPPLEMENTAL MATERIAL REFERENCES

1. Haynes BF, Gilbert PB, McElrath MJ, Zolla-Pazner S, Tomaras GD, Alam SM, et al. Immune-correlates analysis of an HIV-1 vaccine efficacy trial. *N Engl J Med*. 2012;366(14):1275-86.
2. Lin L, Finak G, Ushey K, Seshadri C, Hawn TR, Frahm N, et al. COMPASS identifies T-cell subsets correlated with clinical outcomes. *Nat Biotechnol*. 2015;33(6):610-6.
3. Fourati S, Ribeiro SP, Blasco Tavares Pereira Lopes F, Talla A, Lefebvre F, Cameron M, et al. Integrated systems approach defines the antiviral pathways conferring protection by the RV144 HIV vaccine. *Nat Commun*. 2019;10(1):863.