

Supplemental online-only material

Ghalwash et al. Data-Driven Phenotyping of Presymptomatic Type 1 Diabetes Using Longitudinal Autoantibody Profiles

Table of contents

Supplementary Methods:

- | | |
|--|----------|
| 1) Profile Similarity Score | (Page 2) |
| 2) Dynamic Time Warping (DTW) Alignment | (Page 5) |
| 3) Hierarchical / Agglomerative Clustering | (Page 6) |
| 4) Number of Clusters from Hierarchical Clustering | (Page 6) |

Supplementary Figure S1: The process of computing the normalized similarity score between two subjects A and B. (Page 8)

Supplementary Figure S2: The time-decay weight for different sigma. (Page 9)

Supplementary Figure S3: Estimating the number of clusters by using the proportion of ambiguously clustered pairs (PAC) method. (Page 10)

Supplementary Figure S4: Longitudinal patterns of islet autoantibodies (IAA, GADA, IA-2A) in 11 out of 18 subclusters of children with distinct dynamics of islet autoimmunity. (Page 11)

Supplementary Table S1: The number of samples analyzed for each islet autoantibody, the percentage of positivity, and the proposed weight for positive match relative to negative one of each autoantibody. (Page 13)

Supplementary Table S2: Characteristics of children positive for islet autoantibodies (IAb) in the 8/18 clusters with more than 10 individuals. (Page 14)

References: (Page 15)

Supplementary Methods

1) Profile Similarity Score (PSS)

To cluster subjects a notion of similarity between subjects needs to be defined. We used subject-level sequences of autoantibodies, i.e. individual autoantibody profiles across time to compute the similarity score. A similarity score of 0 indicates that the two subjects are dissimilar with respect to their IAb sequences while a score of 1 indicates that the two sequences of IAb are the same. To operationalize the above, each subject was encoded as a binary matrix where the columns were IAb types, the rows were visits (profiles), and the entries were whether each IAb was positive or negative.

Endesfelder et al.¹ computed similarity between two IAb profiles by simply counting the number of matches; e.g. the similarity between profile A1 and profile B1 is 2 because the two profiles match in IAA and GADA, but differ in IA-2A (Figure S1). Two limitations exist with this approach. First, it does not account for the imbalance in IAb positivity among the subjects, where different types of IAb have different prevalence of positivity and therefore different risks^{2,3}. Second, it is time agnostic as it does not account for the time gap between the two distinct profiles. Timing is important as it has been previously shown that the age at which the IAb is detected is a predictor of progression rate to type 1 diabetes^{4,5}. Therefore, in this work we addressed the above-mentioned limitations as follows:

1a) Account for varying prevalence in autoantibody positivity

Different types of IAb carry different risks of progression to type 1 diabetes. Therefore, we assigned a weight to an IAb match based on its prevalence in the data. The justification for this is that a positive IAb match between two subjects is rarer than a negative IAb match, particularly at

similar clinical age. Therefore, if the IAb is rarely observed as positive then we assign it a higher weight to emphasize that the positive match between two profiles for that IAb is unlikely to happen. In order to account for the differences in the percentage of positivity, we assigned a data-driven weight for positive matches as $\log\left(\frac{p+n}{p}\right)$ where p and n are the number of positive and negative IAb counts, respectively, and we assign 1 for negative matches. Thus, the weighting accounts for variability of positivity within the same IAb and also across IAb (Table S1). So, instead of computing the number of matches between two profiles to derive a similarity score, we weighted each match differently – a) based on the type of IAb; and b) whether it is positive or negative. We call this score the Profile Similarity Score (PSS). An example of PSS is shown in Figure S1(c).

Computing the similarity score

The score between two profiles is based on the weights of each IAb and whether it is a positive or a negative match. Let's assume that w_i is the weight for the IAb $i = \{IAA, GADA, IA-2A\}$ which are based on Table S1. To compute the score between two profiles $A = [a_1, a_2, a_3]$ and $B = [b_1, b_2, b_3]$ where a_i and b_i are either 1 for positive IAb $i = \{IAA, GADA, IA-2A\}$ and 0 for negative IAb as shown in Figure S1(a), we compute the intermediates scores s_i and m_i between a_i and b_i as:

$$s_i = \begin{cases} w_i & \text{if } a_i = b_i = 1 (\text{positive match}) \\ 1 & \text{if } a_i = b_i = 0 (\text{negative match}) \\ 0 & \text{mismatch} \end{cases}, m_i = \begin{cases} w_i & \text{if } a_i = b_i = 1 (\text{positive match}) \\ 1 & \text{otherwise} \end{cases}$$

Then the profile similarity score is computed as:

$$PSS = \left(\frac{\sum_{i=1}^3 s_i}{\sum_{i=1}^3 m_i} \right)$$

For example, the score between the two profiles $A1 = [0, 1, 0]$ and $B1 = [0, 1, 1]$ (as shown in Figure S1) is $s_1 = 1, s_2 = 3.7, s_3 = 0, m_1 = 1, m_2 = 3.7, \text{ and } m_3 = 1$. The profile similarity score $PSS = \left(\frac{1+3.7+0}{1+3.7+1}\right) = 0.82$. The time weight tw between these two profiles $tw = e^{\left(\frac{-(3.2-2.7)^2}{2\sigma^2}\right)} = 0.993$. The time-aware distance score TPDS is computed as $(1 - PSS) * (1 - tw) = 0.001$. Using the TPDS matrix we apply the dynamic time warping algorithm to find the best alignment path that minimizes the weighted distances in the TPDS matrix. Using the optimal path, the normalized similarity score is computed as $NSS = 1 - \frac{\sum_{a,b \in PATH} TPDS[a,b]}{\sum_{a,b \in PATH} tw[a,b]} = 1 - \frac{0.001+0.003+0.002+0}{0.01+0.02+0.02+0.16} = 0.971$.

1b) Time-aware Similarity

Two IAb profiles that are *closer* in terms of time (i.e. the age of the two subjects) should be considered more similar than two profiles that occur further apart, given that the profiles are otherwise similar. To address this temporal awareness issue, we incorporated the clinical age of each profile occurrence in the similarity score, i.e. the similarity between two profiles is weighted based on the time gap between the profiles. Let us assume that the time gap between two profiles is d years. We defined the weight as an exponential function of time decay, i.e. $e^{-d^2/2\sigma^2}$ where σ is a parameter defined as below. For a pair of profiles that are very far from each other (i.e. $d \gg 0$), the time-decay weight becomes very small indicating that we should put less weight on how similar the two IAb profiles are. For a pair of profiles that occurred around the same time (i.e. same age) we put more weight on their similarity. Note that we still computed the similarity score between the two IAb profiles as detailed above, it was just that the time-decay weight further modified the influence of the duration gap and the clinical age at which measurements were done in the

longitudinal IAb profile similarity score. An example of this is shown in Figure S1(d). We set $\sigma = 4$ in our analysis (see below the Computing the time decay factor section for details of this parameter justification) indicating that if the time gap is more than 12 years, the time decay weight should be close to zero to reduce the influence on the IAb profile similarity.

The Time-aware Profile Distance Score (TPDS) between two IAb profiles was then computed as in equation (1) below; an example is shown in Figure S1(e).

$$TPDS = e^{\left(\frac{-d^2}{2\sigma^2}\right)} * PSS \quad (1)$$

Computing the time decay factor

The time decay weight function depends on the parameter σ which controls the rate of decay of the influence of the time gap on the autoantibody profiles similarity as shown in Figure S2. In order to choose a proper value for σ , we analyzed the combined dataset to look at the correlation between the time gap between every pair of visits and the absolute difference in the autoantibody counts. We found that the majority of changes in autoantibody counts occur when the time gap is roughly less than 12 years. Therefore, we choose 12 years as the time gap threshold so that if the time gap is more than 12 years the time decay weight should be close to zero to reduce the influence of the autoantibody profile similarity. Based on that observation, we choose $\sigma = 4$.

2) Dynamic Time Warping (DTW) Alignment:

After computation of the pairwise similarity score between every pair of visits, we applied the dynamic time warping algorithm ⁶ to align the IAb sequences from a pair of subjects based on the computed TPDS distance scores. After finding the best alignment (i.e., the alignment that

maximizes the sum of scores along the path (Figure S1(e)), we computed the normalized sum of scores (NSS) along the alignment.

3) Hierarchical / Agglomerative Clustering

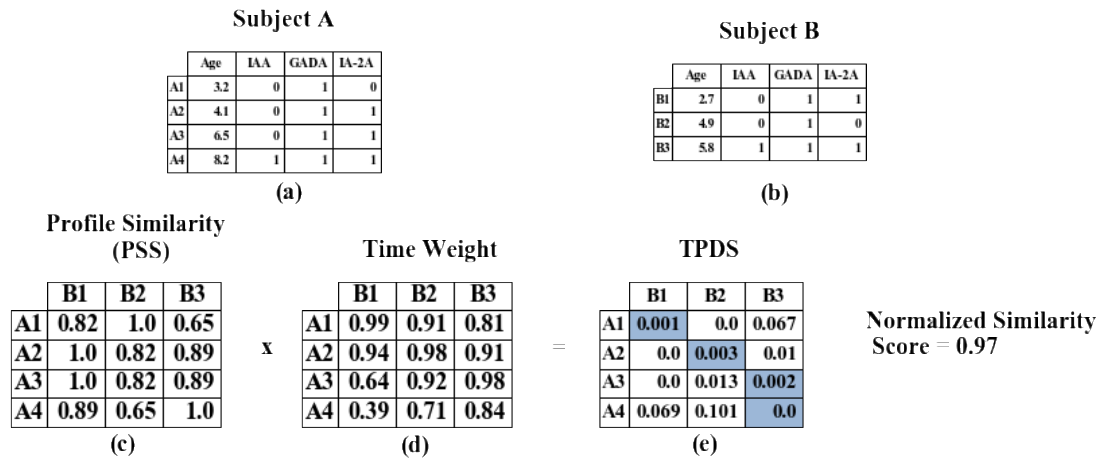
Using the similarity score for every pair of subjects, we computed the distance as $1 - \text{NSS}$ and applied the hierarchical/agglomerative clustering algorithm using average UPGMA (unweighted pair group method with arithmetic mean) linkage on the condensed distance matrix ⁷. This algorithm clusters subjects into a hierarchy of clusters where each subject starts in its own cluster, and as one moves up the hierarchy, pairs of clusters are merged. However, determining the number of clusters from the hierarchy is not a straightforward process.

4) Number of Clusters from Hierarchical Clustering

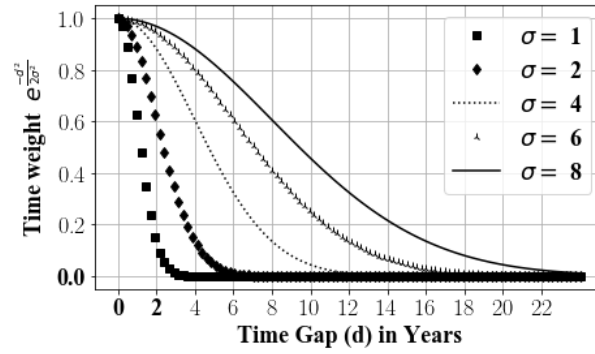
We used a sampling-based method to run random clustering experiments and a recently proposed metric – the proportion of ambiguously clustered pairs (PAC) metric ⁸ - to determine the number of clusters in the cohort. The approach is to randomly sample (without replacement) $N\%$ of the subjects and cluster them into K clusters, repeating the process M times. We chose to randomly sample $N=70\%$ of subjects and repeat the sampling process $M=1000$ times in independent experiments. From each random sampling, we inferred K clusters and by using all M random experiments, we constructed a matrix where each element represents the probability that subject i (row) and subject j (column) belong to the same cluster across the M experiments. This “consensus matrix” is a matrix where entries are estimates of the probability that the two subjects belong to the same cluster in the bootstrap experiments, i.e. the entry ranges from 0 (i.e. the two subjects were never in the same cluster across M runs) to 1 (i.e. the two subjects were always in the same

cluster in all M runs). If the clustering is extremely stable, then a pair of sampled subjects would often cluster together among runs or belong to different clusters in all runs.

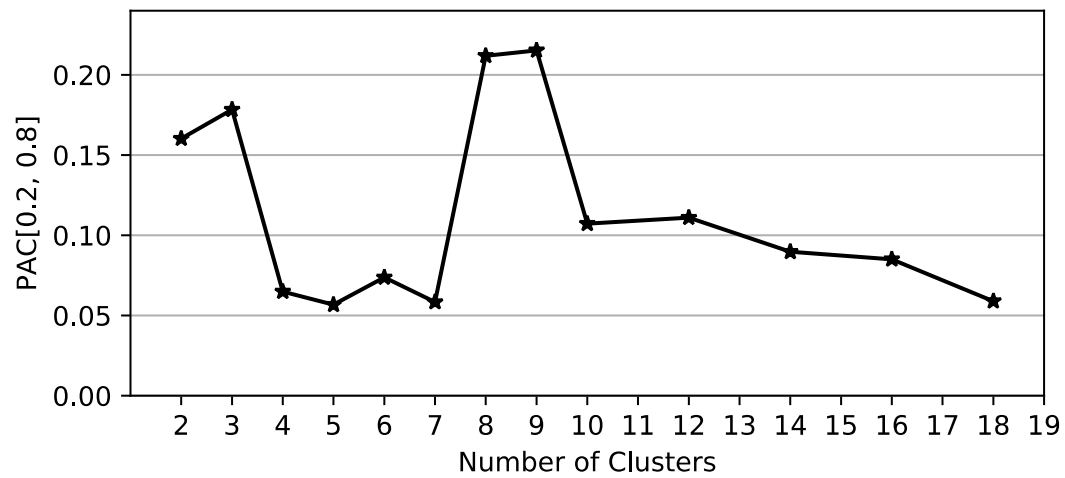
Next, by plotting the cumulative distribution function (CDF) of the entries of the consensus matrix, a new metric, the proportion of ambiguously clustered pairs (PAC) was generated ⁸. Essentially, the lower left portion of the CDF represents subject pairs rarely clustered together, the upper right portion represents those always clustered together, whereas the middle portion represents those with occasional co-assignments in different clustering runs. The clustering is stable if the middle part of the CDF is flat. PAC is defined as the fraction of subject pairs with consensus matrix entries falling in the intermediate sub-interval $(0.2, 0.8) \in [0, 1]$ and can be used as a measure of the stability of the clustering process ⁸. The thresholds 0.2 and 0.8 indicate that the probability that a pair of subjects belongs to the same cluster is either higher than 0.8 or less than 0.2. Lower values of PAC indicate stable clusters, while higher values indicate unstable clusters. We repeated the entire process to compute PAC for different values of K and then plotted K versus PAC to determine the optimal number of clusters as shown in Supplementary Figure S3.



Supplementary Figure S1: The process of computing the normalized similarity score between two subjects A and B using their dynamic islet autoantibody profiles in (a) and (b).

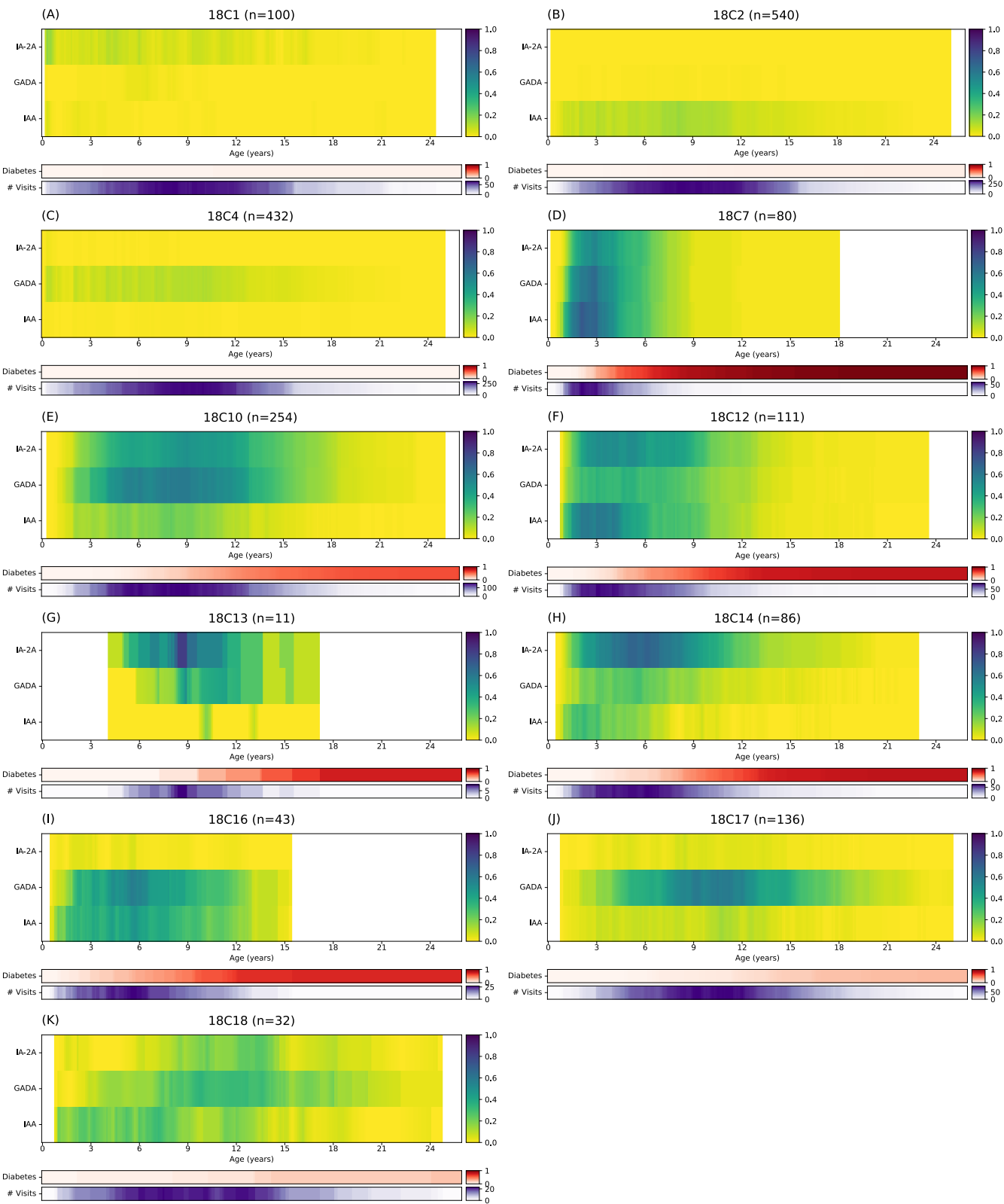


Supplementary Figure S2: The time -decay weight for different sigma. $\sigma = 1$ (faster decay) indicates that the profiles that are more than 2-year apart their autoantibodies profiles are almost negligible, while $\sigma = 8$ still puts non-zero weight on autoantibodies profiles that are 22-year apart.



Supplementary Figure S3: Estimating the number of clusters by using the proportion of ambiguously clustered pairs (PAC) method. Lower value of PAC indicates more stable clusters. 5 and 18 clusters are the most stable clusters.

Supplementary Figure S4.



Supplementary Figure S4: Panels A-K illustrate longitudinal patterns of the three islet autoantibodies (IAA, GADA, IA-2A, shown in the X-axis of each panel) in 11 out of 18 subclusters of children with distinct dynamics of islet autoimmunity (18C1, 18C2, 18C4, 18C7, 18C10, 18C12, 18C13, 18C14, 18C16, 18C17 and 18C18). Green indicates the fraction of positivity for each antibody across all measurements at each age (dark green indicates mostly positive samples, light green indicates that only a small proportion of the samples were positive, and yellow color shows that the samples measured at the corresponding age were negative). The middle section in each panel (Diabetes) shows with red color the cumulative proportion of children progressing to type 1 diabetes. Similarly, the bottom section in each panel (#Visits) shows with purple color the number of measurements collected at each age. Data for the seven subclusters that included less than 10 children is not included.

Subcluster 18C7 was almost identical to cluster 5C2 (n=80, all three IAb persistently positive from age <2 years, 5-year diabetes risk 71.4% and 10-year risk 92.3%, Panel D). Similarly, subcluster 18C16 was almost identical to cluster 5C4 (n=43, persistent IAA and GADA from 2-3 years of age but mostly negative for IA-2A, 5-year diabetes risk 41.2%, Panel I). Subclusters 18C1 (n=100, Panel A), 18C2 (n=540, Panel B) and 18C4 (n=432, panel C) originated from 5C1 and were characterized with positivity for a single, often reverting IAb (IA-2A, IAA or GADA, respectively, and low 5-year and 10-year diabetes risk (1.1% and 3.1%, 2.6% and 5.9%, and 0.3% and 1.3%, respectively). Individuals in the following four subclusters of cluster 5C3 had also relatively high 5-year progression rates to type 1 diabetes: 18C10 (n=254, 5-year risk 26.7%), 18C12 (n=111, 5-year risk 41.5%), 18C13 (n=11, 5-year risk 41.6%), and 18C14 (n=86, 5-year risk 27.5%). These four subclusters were characterized by one or two types of persistent IAb: 18C10 by persistent GADA and IA2A from age 4-6 years but fluctuating or reverting IAA (Panel E), 18C12 by early persistent IAA and IA-2A from age 1-2 years but less frequently GADA (Panel F), 18C13 by persistent GADA and IA-2A from age 6-8 years but mostly negative IAA (Panel G), and 18C14 by persistent IA-2A from age 2.9 years but less frequently positivity for GADA or IAA (Panel H). Individuals in subcluster 18C17 (n=136, 5-year risk 11.7%) resembled those in cluster 5C5, all 136 individuals developing GADA at median age of 5.6 years and rarely having IAA or IA-2A, while children in subcluster 18C18 (n=32, 5-year risk 6.2%) had a pattern of single IAA from age 4 years, and later switching to GADA and/or IA-2A positivity from age 8-9 years (Panels J and K, respectively).

Supplementary Table S1: The number of samples analyzed for each islet autoantibody, the percentage of positivity, and the proposed weight for positive match relative to negative one of each autoantibody. The total number of individual samples was 260,667.

	IAA	GADA	IA-2A
Number of samples	192,707	189,848	189,955
Percentage (number) of positivity	5% (n=9049)	8% (n=14626)	6% (n=10810)
Proposed weight for a positive match	4.4	3.7	4.1

Supplementary Table S2. Characteristics of children positive for islet autoantibodies (IAb) in the 8/18 clusters with more than 10 individuals.

Name of the cluster	18C7	18C10	18C12	18C13	18C14	18C16	18C17	18C18
Number of individuals [males], n	80 [29]	254 [125]	111 [41]	11 [3]	86 [28]	43 [19]	136 [65]	32 [12]
Number of cases with type 1 diabetes, n	77	148	87	8	68	30	34	7
Age at diagnosis, yr, median [IQR]	4.2 [3.0-7.0]	10.4 [7.5-13.5]	7.1 [4.3-9.8]	12.4 [9.7-14.6]	9.0 [6.9-12.1]	7.3 [4.0-9.5]	12.4 [10.2-14.8]	13.1 [9.2-13.7]
HLA Class II group A	31%	31%	30%	18%	28%	47%	24%	16%
HLA Class II group B	48%	47%	53%	36%	59%	30%	46%	59%
HLA Class II group C	11%	7%	6%	36%	3%	12%	12%	16%
HLA Class II group D	10%	15%	11%	9%	9%	12%	18%	9%
Follow-up time, yr, median [IQR]	4.0 [3.0-6.2]	12.4 [9.1-15.2]	8.0 [4.6-10.0]	11.4 [9.2-13.6]	9.1 [6.8-12.1]	7.8 [3.9-11.6]	14.2 [11.4-17.6]	14.6 [10.5-18.1]
*First sample positive for IAA	1.5 (n=80)	5.2 (n=191)	1.6 (n=110)	10.1(n=3)	2.2 (n=68)	2.5 (n=42)	8.0 (n=77)	4.0 (n=31)
*First sample positive for GADA	1.8 (n=80)	4.2 (n=254)	2.2 (n=90)	8.9 (n=9)	3.0 (n=70)	3.0 (n=43)	5.6 (n=136)	8.3 (n=23)
*First sample positive for IA-2A	2.1 (n=80)	6.0 (n=254)	2.5 (n=111)	6.7(n=11)	2.9 (n=86)	4.5 (n=8)	7.6 (n=39)	9.8 (n=19)
*Last sample positive for IAA	4.0 (n=80)	9.8 (n=191)	6.7 (n=110)	10.4 (n=3)	5.0 (n=68)	7.8 (n=42)	11.1 (n=77)	8.0 (n=31)
*Last sample positive for GADA	4.0 (n=80)	12.4 (n=245)	7.3 (n=90)	12.2 (n=9)	6.2 (n=70)	7.8 (n=43)	13.8 (n=136)	5.0 (n=23)
*Last sample positive for IA-2A	4.0 (n=80)	12.4 (n=254)	8.0 (n=111)	11.4 (n=11)	9.1 (n=86)	5.4 (n=8)	9.6 (n=39)	13.9 (n=19)
†Age at seroconversion, yr, median [IQR]	1.4 [1.0-2.0]	4.5 [3.0-7.4]	1.6 [1.1, 3.1]	6.7 [5.2-8.3]	2.8 [1.5-4.5]	2.5 [1.5, 4.0]	6.8 [4.0, 9.1]	5.0 [2.1-9.4]
‡Number of seroconverted individuals, n	80	254	111	11	86	43	131	31
‡IAb profile: IAA only	51.2/0.0	2.8/0.0	51.4/0.0	0.0/0.0	27.9/0.0	48.8/0.0	5.1/0.0	93.8/3.1
‡IAb profile: GADA only	11.2/0.0	51.2/0.4	0.9/0.0	0.0/0.0	7.0/0.0	27.9/4.7	70.6/86.8	3.1/46.9
‡IAb profile: IA-2A only	0.0/0.0	1.6/2.0	9.0/0.0	100/18.2	18.6/96.5	0.0/0.0	0.7/2.9	0.0/34.4
‡IAb profile: IAA+GADA, neg IA-2A	30.0/0.0	19.7/0.0	18.0/0.0	0.0/0.0	3.5/0.0	23.3/95.3	19.9/6.6	3.1/0.0
‡IAb profile: IAA+IA-2A, neg GADA	6.2/1.2	0.8/0.4	12.6/61.3	0.0/0.0	14.0/1.2	0.0/0.0	0.0/0.7	0.0/3.1
‡IAb profile: GADA+IA-2A, neg IAA	0.0/0.0	11.4/71.7	1.8/20.7	0.0/81.8	18.6/2.3	0.0/0.0	2.9/2.9	0.0/12.5
‡IAb profile IAA+GADA+IA-2A	1.2/98.8	12.6/25.6	6.3/18.0	0.0/0.0	10.5/0.0	0.0/0.0	0.7/0.0	0.0/0.0
5-yr risk of type 1 diabetes, % (95% CI)	71.4 (61.3-80.9)	26.7 (21.6-32.9)	41.5 (32.9-51.3)	41.6 (17.7-77.3)	27.5(19.2-38.5)	41.2 (27.9-57.8)	11.7 (7.1-18.9)	6.2 (1.6-22.7)
10-yr risk of type 1 diabetes, % (95% CI)	92.3 (84.6-97.0)	59.4 (52.5-66.3)	76.8 (67.8-84.9)	100 (100-100) [§]	76.7(66.5-85.7)	72.5 (57.5-85.8)	25.6 (18.0-35.6)	21.7 (10.2-42.5)

*Median age (yr) and number of individuals positive for each IAb

†Seroconversion was defined as the first of the two consecutive visits with positivity for the same type of IAb.

‡IAb profile, seven mutually exclusive possibilities, in the first/last positive sample (%)

§The 10-year risk is 100 as all diagnosed individuals were diagnosed with stage 3 type 1 diabetes before age 10

References

1. Endesfelder D, Hagen M, Winkler C, et al. A novel approach for the analysis of longitudinal profiles reveals delayed progression to type 1 diabetes in a subgroup of multiple-islet-autoantibody-positive children. *Diabetologia*. 2016;59(10):2172-2180. doi:10.1007/s00125-016-4050-0
2. Kukko M, Kimpimäki T, Korhonen S, et al. Dynamics of Diabetes-Associated Autoantibodies in Young Children with Human Leukocyte Antigen-Conferred Risk of Type 1 Diabetes Recruited from the General Population. *The Journal of Clinical Endocrinology & Metabolism*. 2005;90(5):2712-2717. doi:10.1210/jc.2004-1371
3. Köhler M, Beyerlein A, Vehik K, et al. Joint modeling of longitudinal autoantibody patterns and progression to type 1 diabetes: results from the TEDDY study. *Acta Diabetol*. 2017;54(11):1009-1017. doi:10.1007/s00592-017-1033-7
4. Gorus FK, Balti EV, Messaaoui A, et al. Twenty-Year Progression Rate to Clinical Onset According to Autoantibody Profile, Age, and HLA-DQ Genotype in a Registry-Based Group of Children and Adults With a First-Degree Relative With Type 1 Diabetes. *Diabetes Care*. 2017;40(8):1065-1072. doi:10.2337/dc16-2228
5. Bauer W, Veijola R, Lempainen J, et al. Age at Seroconversion, HLA Genotype, and Specificity of Autoantibodies in Progression of Islet Autoimmunity in Childhood. *J Clin Endocrinol Metab*. 2019;104(10):4521-4530. doi:10.1210/jc.2019-00421
6. Sakoe H, Chiba S. Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*. 1978;26(1):43-49. doi:10.1109/TASSP.1978.1163055
7. Sokal Robert, Michener C. A statistical method for evaluating systematic relationships | BibSonomy. 1958;XXX38VIII(22). Accessed January 9, 2019. <https://www.bibsonomy.org/bibtex/2e98b6932b71117001ddc4a22b5538532/stephane.guindon?layout=plain>
8. Şenbabaoğlu Y, Michailidis G, Li JZ. Critical limitations of consensus clustering in class discovery. *Scientific Reports*. 2014;4:6207. doi:10.1038/srep06207