Identification of a unique transcriptomic signature associated with islet amyloidosis

Marko Barovic\textsuperscript{1,2,3}, Klaus Steinmeyer\textsuperscript{4}, Nicole Kipke\textsuperscript{1,2,3}, Eyke Schöniger\textsuperscript{1,2,3}, Daniela Friedland\textsuperscript{1,2,3}, Flavia Marzetta\textsuperscript{5}, Almuth Forberger\textsuperscript{6}, Gustavo Baretton\textsuperscript{6,7}, Jürgen Weitz\textsuperscript{8}, Daniela Aust\textsuperscript{6,7}, Mark Ibberson\textsuperscript{5}, Marius Distler\textsuperscript{8}, Anke M Schulte*\textsuperscript{4}, Michele Solimena\textsuperscript{1,2,3#}

\textsuperscript{1}Department of Molecular Diabetology, University Hospital and Faculty of Medicine, Technische Universität Dresden, Dresden, Germany.
\textsuperscript{2}Paul Langerhans Institute Dresden (PLID), Helmholtz Center Munich, University Hospital and Faculty of Medicine, TU Dresden, Dresden, Germany.
\textsuperscript{3}German Center for Diabetes Research (DZD e.V.), Neuherberg, Germany.
\textsuperscript{4}Sanofi Aventis Deutschland GmbH, Diabetes Research, Industriepark Höchst, Frankfurt am Main, Germany.
\textsuperscript{5}Vital-IT Group, SIB Swiss Institute for Bioinformatics, Lausanne, Switzerland.
\textsuperscript{6}Department of Pathology, Medical Faculty, University Hospital Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany.
\textsuperscript{7}NCT Biobank Dresden, University Hospital Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany.
\textsuperscript{8}Department of Visceral, Thoracic and Vascular Surgery, University Hospital Carl Gustav Carus, Medical Faculty, Technische Universität Dresden, Dresden, Germany.

*current contact address: anke.schulte1@yahoo.com
#corresponding author: Michele.Solimena@uniklinikum-dresden.de

Word count: 3750
Abstract

Aims: This cross-sectional study aims to identify potential transcriptomic changes conveyed by presence of amyloid deposits in islets from pancreatic tissue obtained from metabolically profiled living donors.

Methods: After establishing Thioflavin S as the most sensitive approach to detect islet amyloid plaques, we utilized RNA sequencing data obtained from laser capture microdissected islets to define transcriptomic effects of this pathological entity. The RNA sequencing data was used to identify differentially expressed genes by linear modeling. Further analyses included functional enrichment analysis of KEGG and Hallmark gene sets as well as a weighted gene correlation network analysis.

Results: Eleven differentially expressed genes were identified in islets affected by amyloidosis. Enrichment analyses pointed to signatures related to protein aggregation diseases, energy metabolism and inflammatory response. A gene co-expression module was identified that correlated to islet amyloidosis.

Conclusion: Although the influence of underlying Type 2 diabetes could not be entirely excluded, this study presents a valuable insight into the biology of islet amyloidosis, particularly providing hints into the potential relationship between islet amyloid deposition and structural and functional proteins involved in insulin secretion.

Keywords: pancreatic islet amyloidosis, transcriptomics, protein aggregation, Type 2 diabetes, metabolically phenotyped living donors
Research in context

What is already known about this subject?

- Islet amyloidosis is the only histological marker of Type 2 diabetes in the pancreas
- Individuals not suffering from Type 2 diabetes can also be affected by islet amyloidosis
- The clinicopathological significance of this phenomenon is still unclear

What is the key question?

- Does the islet transcriptome of individuals with islet amyloidosis provide explanations for the onset of this phenomenon and its pathophysiological value?

What are the new findings?

- Islet transcriptomes of affected subjects exhibit only limited transcriptomic differences compared to unaffected ones.
- Structural and functional proteins involved in insulin secretion machinery may be involved in the pathophysiological sequence of amyloid formation
Introduction

Type 2 diabetes mellitus (T2D) is a metabolic disorder characterized by chronic hyperglycemia in the absence of immune-mediated beta cell destruction. The main disease burden is defined through its numerous complications that occur over several years, and negatively affects healthcare systems across the world. This disease is principally characterized by two pathological features: beta cell failure and the resistance of peripheral tissues to the action of insulin. The body of research focusing on beta cell failure as the *conditio sine qua non* for the onset of T2D spans across decades and scientific fields, seeking the mechanisms causing the demise of the beta cell secretory machinery and thus the decompensation of glucose homeostasis.

Diagnostic tests from peripheral blood have been validated and well established in the clinical practice. These provide valuable, easily accessible and therapeutically crucially important information on the metabolic state of the patient. On the other hand, morphological changes in the islets of Langerhans and more specifically the insulin-producing beta cells which could potentially inform research advances in the field of beta cell biology are yet to be discovered, with a notable exception of islet amyloidosis. Hyaline deposits in islets of Langerhans have been identified in human pancreatic tissue sections essentially since the dawn of modern histopathology. (1) This phenomenon was, however, not viewed as particularly significant because of its occurrence also in the tissue of normoglycemic subjects, although in a smaller extent compared to patients with T2D. (2)

Despite the noted inter-species conservation, the minute differences in the peptide sequence of IAPP across species proved to be sufficient to convey a disposition of the human and some other vertebrate variants to fibrillate and form aggregates. These, in turn, are
considered to be the precursors of islet amyloid plaques, by serving as deposition nuclei. It remains disputed whether these nuclei come to be in the intercellular space or within the cell, which would explain the cytotoxic effect of amyloid plaques on beta cells given that the extracellular amyloid deposits are considered to be inert. The most recent understanding of the chain of events proposes that the fibrillation indeed starts within the beta cell leading to cell death and release of the seeding aggregates into the intercellular space where they grow by further deposition to form the extracellular plaques nowadays most commonly understood as islet amyloidosis. (3) Notably, the IAPP of the most frequently used experimental animals in diabetes research – mouse and rat – does not have the sequence motifs that confer amyloidogenic properties. For this reason, several transgenic model organisms have been established in order to study islet amyloidosis on a molecular level, each with their own shortcomings in recapitulating this phenomenon in humans.

In this study, we profiled the pancreatic tissue biobank of the LIving DOnor PAncreatic COhort (LIDOPACO) initiative for the presence of islet amyloidosis after establishing a sensitive methodology for its detection by microscopy. This data was then integrated with the extensive clinical data for each individual donor, thereby enabling the establishment of a database on human islet amyloidosis with unprecedented detail. Molecular patterns associated with islet amyloidosis were pursued by means of the analysis of islet RNA sequencing data available for all the donors in this cohort.
Material and methods

Cohort

Based on tissue availability, a total of 121 patients enrolled in the LIDOPACO program were included in this study in accordance with the approval of the Ethical Committee of the Technische Universität Dresden. After obtaining informed consent, medical history and clinical laboratory data relevant for metabolic status was gathered either from existing preoperative preparatory examinations or from peripheral blood samples collected with the goal of completing the metabolic profiling as part of the LIDOPACO (e.g. oral glucose tolerance test (OGTT)). According to this data, patients were assigned either to the non-diabetic (ND), impaired glucose tolerance (IGT), pancreoprive diabetes (T3cD) or T2D categories according to the criteria defined by the American Diabetes Association (ADA). The ND group included 18 patients, the IGT group 39 patients classified as pre-diabetic (including patients with impaired fasting glucose) and the T3cD and T2D groups 32 patients each. Patients who underwent neoadjuvant chemotherapy, as well as those with pancreatic neoplasms of endocrine origin were excluded from the study.

Pancreatic tissue procurement

Immediately following the surgical procedures of partial or total pancreatectomy conducted for various indications at the Clinic for Visceral, Thoracic and Vascular Surgery of the University Hospital Carl Gustav Carus, surgical pathological samples were collected and processed by a pathologist. Fragments of healthy tissue were prepared and either snap frozen in liquid nitrogen and thereafter stored at -80° or fixed in formalin and embedded in paraffin.
Detection of islet amyloid plaques

Islet amyloid plaques were detected in 2 μm thin sections of formalin fixed paraffin embedded (FFPE) tissue using histochemical labeling by either Congo Red or Thioflavin S. Slides were stained with Congo Red according to the standard protocols of the Institute for Pathology of the University Hospital Carl Gustav Carus and evaluated for presence of amyloidosis in a binary fashion (present/not present) by a pathologist.

Staining with Thioflavin S was performed by incubating the slides for 10 minutes in the staining solution (0.5% Thioflavin S w/v in distilled water) followed by washing steps. Immediately afterwards, the slides were subjected to an automated immunofluorescence protocol on the Roche Ventana Discovery system targeting synaptophysin (Rabbit anti-Synaptophysin (ThermoFisher, MA5-14532, final dilution 1:30); goat anti-rabbit AF750 (Invitrogen, A21039; final dilution 1:150)) without signal amplification and with DAPI as nuclear counterstain. Whole slide images were acquired on a Zeiss AxioScan slide scanner at 20x magnification in the appropriate fluorescence channels.

Across all slide scans, at least 200 islets were manually segmented per sample based on the synaptophysin expression, with the examiner blinded for the Thioflavin S signal (in total 28344 islets). These were saved as regions of interest (ROIs) and used as delimiter for binary detection of a signal in the Thioflavin S channel (present/not present) in each ROI. Accounting for lack of specificity, we set the threshold for classification into the amyloid positive group to be >1 rather than >0 affected islets. Considering that the aimed for number of assessed islets per donor was 200, this corresponds to a threshold of 0.5%, which is widely regarded in statistics as an acceptable error margin. Substantiating our approach of assessing only a small portion of all islets in the whole gland for practicality reasons is the previous finding of homogeneous distribution of
amyloidosis affected islets in transgenic mice. (4) All image analysis was done using Fiji version 1.53.

**Transcriptomics**

**Laser Capture Microdissection**

Pancreatic islets for RNA sequencing were collected using Laser Capture Microdissection (LCM) as previously described. (5–7) Briefly, snap frozen tissue was sectioned in a cryostat into 20μm thin sections and mounted on Zeiss PEN Membrane Slides. After drying and in strictly RNAse-free conditions, these slides were processed on a Zeiss Palm MicroBeam system. A total islet area of 1x10^6 μm^2 was excised based on the islet lipofuscin autofluorescence for each patient. Complete mRNA was isolated using Arcturus PicoPure RNA isolation kit and its concentration was determined using Qubit RNA high sensitivity platform. Before proceeding to the next steps, RNA integrity was determined using an Agilent Bioanalyzer 2100.

**RNA sequencing**

Complementary DNA (cDNA) libraries were constructed using the Illumina Smart-seq2 kit selecting only for the Poly(A)^+ strands. After enrichment, purification and fragmentation steps, the libraries were subjected to single ended sequencing at the DRESDEN-Concept Genome Center on the Illumina NovaSeq 500 platform with 76bp read length and target depth of at least 35 million reads per library.

Low quality sequences and adapters were filtered using Cutadapt (ver. 1.8), and fastq_screen (ver. 0.11.1) was used for removal of ribosomal sequences. Low complexity reads were subsequently removed with reaper (ver. 15-065) and the remaining reads aligned to the
Genome Reference Consortium Human Build 38 patch release 12 index using STAR aligner (ver. 2.5.3a). Read counts per gene locus were summarized using htsq-count (ver. 0.9.1).

Samples were grossly assessed for presence of exocrine contamination by examining the expression levels of genes typical for the exocrine tissue (PLIP, PRSS1, AMY2A, AMY2B). This and further quality assurance steps (analysis of the absolute number of detected genes, gene body coverage and cumulative gene diversity) were used for a more granular quality assurance.

For downstream analyses, the dataset was filtered for minimal expression retaining only the genes with at least 1 count per million reads (CPM) in at least 10% of the samples.

**Differential expression and enrichment analysis**

Bioconductor package DESeq2 functions calcNormFactors and vst were used for library size normalization and variance stabilizing transformation respectively. Exploratory multidimensional scaling plots were created from this data using the R function plotMDS (limma).

Analysis of differential gene expression was performed using the bioconductor package limma. In an attempt to isolate the effects of islet amyloidosis from the effects of the disease itself, diabetes status was included as covariate in the linear model, together with age, gender, BMI and sequencing batch. Genes with Benjamini-Hochberg multiple comparison corrected $p$ value lower than 0.05 and absolute fold change higher than 1.2 were considered differentially expressed.

Functional enrichment analysis of Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways was conducted using Gene Set Enrichment Analysis using the gseKEGG function of the R package clusterProfiler (ver. 3.10.1) on unfiltered gene lists ranked by decreasing differential expression adjusted $p$ value in an unsigned manner. Gene sets were filtered to include only those sets which contained >100 and <500 genes, and the enrichment scores were normalized by gene set size. Results were filtered using a cutoff of 0.01 on the BH-adjusted $p$
value and term redundancy was reduced using the function `simplify` from the same package and using enrichment maps generated by the `emapplot` function from the R package enrichplot (ver. 1.2.0).

Enrichment of The Broad Institute Molecular Signatures Database v7.4 hallmark gene sets was assessed by competitive gene set test accounting for the inter-gene correlation, using the function `camera` from the limma Bioconductor package.

**Weighted Gene Correlation Network Analysis**

The Weighted Gene Correlation Network Analysis (WGCNA) was performed using the R package WGCNA (ver. 1.68). A gene coexpression network was constructed from normalized, batch-corrected (`removeBatchEffect` function from the package limma) and vst transformed expression counts. An unsigned adjacency matrix was constructed using Pearson correlation of pairwise complete observations. Soft threshold parameter was optimized using the `pickSoftThreshold` function and the best threshold ($\alpha=6$) was visually determined following developers’ recommendations. A topological overlap matrix was derived from the adjacency matrix and converted to distances, which were hierarchically clustered using average linkage clustering. Finally, dynamic tree cut using hybrid method and parameters `minClusterSize=20` and `deepSplit=2` identified individual gene coexpression modules. Similar modules were in turn merged based on an eigengene distance threshold of 0.15.

Spearman correlation was used to relate the module eigengenes to selected clinical traits and to the qualitative islet amyloidosis status based on the Thioflavin S staining. A heatmap of the module-trait correlations was produced using `labeledHeatmap` function from the WGCNA package, depicting the Spearman correlation coefficients and corresponding $p$ values calculated using `cor` and `corPvalueStudent` functions from the same package. In the module with the
strongest correlation to islet amyloidosis, genes with the highest correlation to the trait and the module eigengene were selected as module hub genes.

## Results

The cohort used in this study consisted of 121 surgical pancreatic tissue donors participating in the LIDOPACO programme. The patients were included into this study based on the availability of tissue for the necessary investigations. Eighteen of the patients were classified as ND through the preoperative metabolic profiling procedure, 39 were in the IGT group while the T3cD and T2D groups consisted of 32 subjects each (Fig. 1A). In terms of basic demographic/clinical measures, the ND subjects were found to have a significantly lower age and BMI. Expectedly, these subjects also exhibited a lower HbA1c and glycaemia at 2h point of OGTT values compared to the other groups (Fig. 1B). Malignant tumors were markedly more common in the T2D and T3cD groups than in the IGT group, with an even bigger gap when compared to the ND group (Table 1).

At first, we aimed to determine the most sensitive method of amyloid plaque detection. To this end, we employed two historically well-established methods for the detection of amyloid plaques – Congo Red and Thioflavin S. The Congo Red staining evaluated by a pathologist revealed only 6 amyloid positive samples. At odds with this finding, a custom semi-automated pipeline applied on the manually segmented fluorescence microphotographs of 28 344 pancreatic islets stained with Thioflavin S, identified islet amyloidosis in 5/18 (27.8%) ND subjects, 16/39 (41%) in IGT, 14/32 (43.7%) in T3cD and 25/32 (78.1%) in T2D subjects (Figure 2A). The islet amyloidosis assignment based on this method was therefore used for all downstream analyses. Representative images of amyloid positive and amyloid negative islets are shown in Figure 2B.
Assessment of the basic metabolically relevant clinical variables in the subjects with and without islet amyloidosis unsurprisingly revealed a significantly higher HbA1c levels (Figure 3A). This is consistent with the growing percentage of amyloidosis affected subjects in the disease categories progressing towards overt T2D. Interestingly, no difference was found in age, a demographic parameter typically associated with onset of amyloidosis in general. Same was true for BMI and glycaemia at 2h point of OGTT (Figure 3A).

Although a metric of limited value due to the nature and the scope of examined pancreatic islets, the distribution of the absolute number of amyloidosis affected islets pointed to a heavily left-skewed distribution. Most patients exhibited the affection of less than 50 out of the assessed minimally 200 islets and only few showing a broadly reaching islet amyloidosis (Figure 3B). The same parameter when observed in the two diabetes groups pointed out the mentioned distribution and further illustrated the presumed etiological, and possibly perduration, difference between the metabolic disorder of IGT (persistent) and T3cD (acute or sudden) (Figure 3C). The ND patient with an extensive islet amyloidosis could likely be considered an outlier.

Batch-corrected and normalized RNA sequencing data was used for data exploration considering amyloidosis status as well as other clinical (diabetes status, gender) and technical characteristics (sequencing batch). The multidimensional scaling plots only revealed a distinct separation of two patient groups, which was driven primarily by gender differences in gene expression. Based on this analysis, the batch effect appeared to be sufficiently corrected for. (Figure 4A-D).

In an effort to pinpoint the transcriptomic changes correlated with islet amyloidosis itself rather than the metabolic disorder (prediabetes or diabetes), the information on diabetes status was included as a covariate in the linear fitting model to test for differential expression between amyloid positive or negative islets. Based on a cutoff of $p$ value <0.01 and an absolute fold change
1.2, a total of 11 genes were found to be differentially expressed (Figure 4E, Table 2). Intriguingly, a long non-coding RNA was the hit with the highest fold change, while two other genes with a known relation to beta cell biology were detected as downregulated in this analysis (PTPRN (p = 0.03, logFC = -0.399) and FURIN (p = 0.04, logFC = -0.416)). None of these 11 differentially expressed genes had also been identified as differentially regulated in T2D donors in our previous study on a superset of the cohort presented here, supporting the validity of our approach to isolate the effects of amyloidosis from the effects of the disease/hyperglycaemic disorder.

To gain insights into the observed transcriptional changes, a gene set enrichment analysis of KEGG pathways was performed on the differentially expressed gene list ranked by adjusted p value and results were filtered using cut-offs for normalized enrichment score of 1.5 and for adjusted p value of 0.01. Oxidative phosphorylation, RNA transport and cytoskeleton related terms were enriched in amyloid positive islets, similarly to our previous study of transcriptomic changes in islets of T2D (Table 3). The KEGG term oxidative phosphorylation was not only interesting for its role in the islet cell energy flux, but also for a significant overlap with the terms Parkinson’s disease, Alzheimer’s disease and Huntington’s disease.

Testing for higher ranking of genes belonging to The Broad Institute Molecular Signatures Database v7.4 hallmark gene sets returned 15 hallmark gene sets overrepresented in amyloid positive islets with a false discovery rate of 0.05 (Table 4). Interestingly, and partially at contrast with the analysis of the KEGG gene sets, this analysis identified several gene sets primarily related to inflammation and immune response. Still, the oxidative phosphorylation gene set was also identified in this analysis.

Exploring the transcriptomic data in pursuit of co-expressed genes with correlation to islet amyloidosis, a WGCNA was performed (Figure 4F). The pink module was identified with a weak,
although significant correlation to islet amyloidosis \( (r = 0.3) \), which included 20 genes recognized as hub genes for their strong correlation to the module eigengene. Table 5 lists these genes, out of which four were previously reported in relation to T2D and pancreatic beta cells: ARDB1, FXYD3, F13A1 and FXYD2 \((8–11)\). None of the module hub genes, including the four mentioned ones, have been previously associated with islet amyloidosis.

**Discussion**

The depth and extent of the data collected in the LIDOPACO studies was a prerequisite for this analysis, unique in its pursuit to uncover transcriptomic signatures correlated with islet amyloidosis. These investigations demonstrate how technologies relatively recently introduced into scientific practice can be utilized for extracting new knowledge regarding phenomena that are well known in the broader medical science field since many decades.

Few studies exist that directly compared the sensitivity of the two most commonly used techniques for detection of amyloid plaques - Thioflavin S and Congo Red staining.\((12)\) The data we gathered in this study unequivocally show the superior sensitivity of Thioflavin S staining for detection of islet amyloid plaques - a procedure that is routinely used for this purpose in pathology. We hypothesize that this effect is largely conveyed by a higher contrast ratio in fluorescence-based imaging, enabling easier detection of small deposits of amyloid material. Enhanced contrast also facilitates the automated microscopy detection of signals in islet regions of interest. While Thioflavin S staining is highly sensitive for the detection of islet amyloid plaques, it comes with the cost of a loss in specificity, for which we tried to compensate by using a more stringent detection threshold.

There are several historical reports of islet amyloidosis also affecting non-diabetic individuals with varying records of its prevalence.\((1,2,13–16)\) While the differences among the
previously published investigations and also with the present study could have a biological
background based in the cohort characteristics, we argue that this variability most likely stems
from non-standardized methodologies for amyloid detection, as illustrated above.

Our data on the clinicopathological aspects of islet amyloidosis aligns with previous
knowledge and support the notion that this alteration is positively correlated with the progression
of hyperglycemia and T2D. To date there is no other detailed human clinical data available that
is presented in view of pancreatic islet. The extent of islet amyloidosis among LIDOPACO donors
with IGT and T3cD as presented in Figure 3C is in line with the presumed different temporal
dynamics of these two entities, i.e. a chronic but slowly progressive versus an acute, but more
severe deterioration of glucose homeostasis and islet function.

The unsupervised exploratory data analysis revealed no clear grouping of samples along
the amyloidosis positivity line along with an incidental finding of a remarkable clustering according
to gender. The following differential expression analysis was performed in accordance to the
primary aim - pursuit of transcriptomic demarcations of islet amyloidosis. This entailed a complex
linear modeling with multiple covariates to ensure the highest possible degree of specificity, this
time with a potential sacrifice in sensitivity. Endorsing the specificity of these differentially
expressed genes is the fact that none of them has been identified in the differential expression
analyses on the superset of the cohort used in this study we previously published.(7) The list of
differentially expressed genes, albeit short, provided valuable cues in relation to existing
knowledge and for future endeavors. Interestingly, PTPRN (ICA512) and FURIN, genes closely
associated with vesicle mediated exocytosis, were found to be downregulated in islets of subjects
with islet amyloidosis. The case of PTPRN (ICA512) is especially intriguing given the high
propensity of the insulin secretory granule luminal N-terminal RESP18 homology domain
(ICA512-RESP18HD), to aggregate at neutral pH, i.e. in the condition encountered by granule
cargoes upon their release in the extracellular milieu.(17,18) This remarkable feature sets
ICA512-RESP18HD apart from other granule released peptides, which are instead soluble at
neutral pH, while forming condensates at pH ≥ 6.3, as in the lumen of the trans-Golgi network and the secretory granules. Notably, ICA512-RESP18HD contains information for the sorting of peptides into insulin secretory granules and it is cleaved from the rest of the PTPRN/ICA512 ectodomain by proprotein convertases, such as FURIN. Furthermore, our data indicates that secretion of ICA512-RESP18HD from glucose-stimulated insulinoma cells is coupled with the cleavage of its C-terminal intrinsic disordered domain, which accounts for its peculiar condensing properties. In the future it would interesting to investigate whether one or both these cleavages of ICA512-RESP18HD are perturbed in conditions correlated with the formation of islet amyloidosis.

Higher level analyses primarily through Gene Set Enrichment Analysis revealed some interesting patterns. Protein misfolding/ER stress signatures align with previously published studies in transgenic mice (19), but are challenged by other evidence in the same experimental model.(20) Related enriched terms referring to Alzheimer’s, Parkinson’s and Huntington’s disease are of high interest given that pathological protein aggregation is a common denominator for both amyloidosis and these neurodegenerative diseases. In addition, patients with T2D are at a higher risk to develop Alzheimer’s and/or Parkinson’s disease.(21,22) This is further supported by recent evidence of direct interaction of amyloid ß and IAPP in vitro.(23)

Inflammation and cytokine signaling related signatures found both in KEGG and Hallmark functional enrichment analyses do find support in the literature (24), despite the disputed importance of these processes in the pathogenesis of T2D. Our data agrees also with other conclusions of that specific study, with special emphasis on mitochondrial function and protein synthesis.

Finally, it must be noted that despite our best efforts, the extensive effects of T2D on islet cell biology could not be entirely eliminated. This limitation is revealed by the overlap with our previous findings where T2D was the focus, especially in the domain of integrative and overview analyses.(7) Nevertheless, the gene co-expression module correlating with islet amyloidosis identified in this paper might be a meaningful source of inspiration for potential future research.
Acknowledgements

We would like to thank Juha Torkko from Solimena lab, Mario Ermacora (Universidad Nacional de Quilmes, Argentina) and Steven Kahn (VA Puget Sound Health Care System and University of Washington, Seattle, WA, USA) for valuable insights, suggestions and discussion, and Katja Pfriem for administrative assistance.

Funding

This project has received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No 115881 (RHAPSODY). This Joint Undertaking receives support from the European Union’s Horizon 2020 research and innovation program and EFPIA. This work is further supported by the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract number 16.0097-2. Work in the Solimena lab is also supported with funds from the German Ministry of Education and Research to the German Center for Diabetes Research (DZD). The opinions expressed and arguments employed herein do not necessarily reflect the official views of these funding bodies.

Disclosures

Authors disclose no conflicts of interest.

Author contributions
JW and MD patient recruitment, surgery, provision of clinical data; ES, NK and DF tissue collection, processing and data collection; AF, DA, GB pathology and assessment of congo red staining; MB ES and NK patient data management and selection; KS immunofluorescent staining; MB development and application of amyloidosis quantification approach; FM and MB bioinformatic analysis; AS and MS conceptual insights and design, MB and MS writing of the manuscript; all authors have revised the manuscript and approved its final version.

References


Figure captions
Figure 1. Cohort characteristics overview (A) Treemap plot of the cohort composition in view of diabetes status. (B) Basic clinical characteristics of the cohort, separated into diabetes status categories. Boxplot corresponds to the span between the 25th and 75th percentile, with horizontal line marking the median. The whiskers reach to the most extreme data point that is no more than 1.5 times the length of the box away from the box. ND non-diabetic, IGT impaired glucose tolerance, T3cD Type 3c diabetes, T2D Type 2 diabetes.

Figure 2. Islet amyloidosis detection (A) Stacked bar chart of proportions between amyloid positive and amyloid negative samples between the two amyloid staining approaches. (B) Representative images of Congo red and Thioflavin S staining techniques with positive and negative staining examples. Scale bar corresponds to 50 μm.

Figure 3. Clinical and morphological characteristics of islet amyloidosis (A) Boxplots of representative clinical parameters in relation to amyloidosis positivity. Boxplot corresponds to the span between the 25th and 75th percentile, with horizontal line marking the median. The whiskers reach to the most extreme data point that is no more than 1.5 times the length of the box away from the box. ***p<0.001 (B) Bar chart of the distribution of extent of islet amyloidosis in the cohort. (C) Boxplot of the absolute number of affected islets in subjects with islet amyloidosis. Boxplot corresponds to the span between the 25th and 75th percentile, with horizontal line marking the median. The whiskers reach to the most extreme data point that is no more than 1.5 times the length of the box away from the box.

Figure 4. Transcriptomic signatures in islet amyloid affected subjects (A-D) Multidimensional scaling plots for the first four dimensions labelled by diabetes status (A), sequencing batch (B), amyloid status (C), and gender (D). (E) Volcano plot representation of differential expression analysis of amyloid positive vs amyloid negative islets. (F) Correlation heatmap of gene co-expression modules with relevant clinical parameters.

Table 1. Underlying pancreatic disease. Frequencies of underlying pancreatic disorders categorized by diabetes status
Table 2. Differentially expressed genes. Islet genes differentially expressed in islet amyloidosis affected individuals. Only genes with adjusted p value < 0.05 and a fold change > 1.2 are reported.

Table 3. Gene set enrichment analysis - KEGG. Enriched pathways from the Kyoto Encyclopedia of Genes and Genomes in the differential expression analysis.

Table 4. Gene set enrichment analysis - Hallmark. Enriched terms from the Molecular Signatures Database Hallmark dataset.

Table 5. Weighted gene correlation network analysis. Hub genes of the gene co-expression module with the strongest correlation with islet amyloidosis.
Figure 1

A

- T3D
  - 39
- ND
  - 18
- IGT
  - 32
- T2D
  - 32

B

- Age (Years)
  - ND
  - IGT
  - T3D
  - T2D
- BMI
  - ND
  - IGT
  - T3D
  - T2D
- HbA1c (%)
  - ND
  - IGT
  - T3D
  - T2D
- Glycerol at 2h OGTT (mmol/L)
  - ND
  - IGT
  - T3D
  - T2D
Figure 2

A

- **Congo Red**
  - T2D
  - T3D
  - IGT
  - ND

- **Thioflavin S**
  - T2D
  - T3D
  - IGT
  - ND

Proportion

Orange: Amyloid positive
Green: Amyloid negative

B

- **Congo red**
  - Amyloid negative
  - Amyloid positive

- **Thioflavin S**
  - Amyloid negative
  - Amyloid positive
Figure 3

A

![Box plots showing age, BMI, HbA1c, and fasting glucose levels in negative and positive Amyloidosis groups.]

B

![Histogram showing the distribution of positive islets across different number of patients.]

C

![Box plots comparing the number of affected islets in different diabetes statuses: ND, IGT, T3D, and T2D.]

---

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.09.515784; this version posted November 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.
Figure 4
Table 1

<table>
<thead>
<tr>
<th>Diabetes status</th>
<th>Chronic pancreatitis</th>
<th>Benign tumor</th>
<th>Malignant tumor</th>
</tr>
</thead>
<tbody>
<tr>
<td>ND</td>
<td>4 (22.22%)</td>
<td>5 (27.78%)</td>
<td>9 (50%)</td>
</tr>
<tr>
<td>IGT</td>
<td>6 (15.38%)</td>
<td>9 (23.07%)</td>
<td>24 (61.53%)</td>
</tr>
<tr>
<td>T3cD</td>
<td>5 (15.62%)</td>
<td>3 (9.37%)</td>
<td>24 (75%)</td>
</tr>
<tr>
<td>T2D</td>
<td>6 (18.75%)</td>
<td>3 (9.37%)</td>
<td>23 (71.87%)</td>
</tr>
<tr>
<td>Gene.stable.ID</td>
<td>Gene.name</td>
<td>logFC</td>
<td>AveExpr</td>
</tr>
<tr>
<td>---------------</td>
<td>-----------</td>
<td>-------</td>
<td>---------</td>
</tr>
<tr>
<td>ENSG0000005435</td>
<td>PTPRN</td>
<td>-0.399466</td>
<td>8,2309211</td>
</tr>
<tr>
<td>ENSG0000005511</td>
<td>KCNH2</td>
<td>-0.324531</td>
<td>4,8611198</td>
</tr>
<tr>
<td>ENSG0000007437</td>
<td>ATP2A3</td>
<td>0,4028669</td>
<td>7,2470098</td>
</tr>
<tr>
<td>ENSG0000011492</td>
<td>SLC4A3</td>
<td>0,4747201</td>
<td>2,8537569</td>
</tr>
<tr>
<td>ENSG0000013346</td>
<td>SLC2A11</td>
<td>0,3089391</td>
<td>4,3597096</td>
</tr>
<tr>
<td>ENSG0000013592</td>
<td>DNAJB2</td>
<td>0,3269333</td>
<td>4,3597096</td>
</tr>
<tr>
<td>ENSG0000014056</td>
<td>FURIN</td>
<td>0,4173172</td>
<td>4,3588284</td>
</tr>
<tr>
<td>ENSG0000018146</td>
<td>RAP2B</td>
<td>0,4792213</td>
<td>5,5234012</td>
</tr>
<tr>
<td>ENSG0000019622</td>
<td>SRGAP3</td>
<td>0,3743544</td>
<td>2,9899396</td>
</tr>
<tr>
<td>ENSG0000019655</td>
<td>CACNA1H</td>
<td>0,5075699</td>
<td>4,6221361</td>
</tr>
<tr>
<td>ENSG0000027060</td>
<td>AL353622.1</td>
<td>0,6200652</td>
<td>2,3856817</td>
</tr>
<tr>
<td>ID</td>
<td>Description</td>
<td>setSize</td>
<td>enrichmentScore</td>
</tr>
<tr>
<td>----------</td>
<td>------------------------------------------</td>
<td>---------</td>
<td>-----------------</td>
</tr>
<tr>
<td>hsa04510</td>
<td>Focal adhesion</td>
<td>185</td>
<td>0.3218814</td>
</tr>
<tr>
<td>hsa04510</td>
<td>Regulation of actin cytoskeleton</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hsa04810</td>
<td>Cytokine-cytokine receptor interaction</td>
<td>179</td>
<td>0.3226991</td>
</tr>
<tr>
<td>hsa04060</td>
<td>Phagosome</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hsa04514</td>
<td>Cell adhesion molecules (CAMs)</td>
<td>108</td>
<td>0.3415718</td>
</tr>
<tr>
<td>hsa04142</td>
<td>Lysosome</td>
<td>115</td>
<td>-0.3128041</td>
</tr>
<tr>
<td>hsa04120</td>
<td>Ubiquitin mediated proteolysis</td>
<td>129</td>
<td>0.3147468</td>
</tr>
<tr>
<td>hsa04110</td>
<td>Cell cycle</td>
<td>113</td>
<td>0.324237</td>
</tr>
</tbody>
</table>
Table 4

<table>
<thead>
<tr>
<th></th>
<th>NGenes</th>
<th>Direction</th>
<th>PValue</th>
<th>FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>HALLMARK_MYC_TARGETS_V1</td>
<td>198</td>
<td>Up</td>
<td>4.55E-09</td>
<td>2.27E-07</td>
</tr>
<tr>
<td>HALLMARK_INTERFERON_ALPHA_RESPONSE</td>
<td>97</td>
<td>Up</td>
<td>5.54E-08</td>
<td>1.39E-06</td>
</tr>
<tr>
<td>HALLMARK_INTERFERON_GAMMA_RESPONSE</td>
<td>191</td>
<td>Up</td>
<td>8.94E-08</td>
<td>1.49E-06</td>
</tr>
<tr>
<td>HALLMARK_ALLOGRAFT_REJECTION</td>
<td>161</td>
<td>Up</td>
<td>1.25E-06</td>
<td>1.56E-05</td>
</tr>
<tr>
<td>HALLMARK_COMPLEMENT</td>
<td>175</td>
<td>Up</td>
<td>1.86E-06</td>
<td>1.86E-05</td>
</tr>
<tr>
<td>HALLMARK_OXIDATIVE_PHOSPHORYLATATION</td>
<td>198</td>
<td>Up</td>
<td>1.19E-05</td>
<td>9.68E-05</td>
</tr>
<tr>
<td>HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITIFICATION</td>
<td>193</td>
<td>Up</td>
<td>1.39E-05</td>
<td>9.68E-05</td>
</tr>
<tr>
<td>HALLMARK_E2F_TARGETS</td>
<td>184</td>
<td>Up</td>
<td>1.55E-05</td>
<td>9.68E-05</td>
</tr>
<tr>
<td>HALLMARK_PROTEIN_SECRETION</td>
<td>95</td>
<td>Up</td>
<td>0.00108</td>
<td>0.0060003</td>
</tr>
<tr>
<td>HALLMARK_KRAS_SIGNALING_UP</td>
<td>182</td>
<td>Up</td>
<td>0.0015183</td>
<td>0.0075917</td>
</tr>
<tr>
<td>HALLMARK_UV_RESPONSE_DN</td>
<td>142</td>
<td>Up</td>
<td>0.0024307</td>
<td>0.0110487</td>
</tr>
<tr>
<td>HALLMARK_INFLAMMATORY_RESPONSE</td>
<td>168</td>
<td>Up</td>
<td>0.0041634</td>
<td>0.0173477</td>
</tr>
<tr>
<td>HALLMARK_FATTY_ACID_METABOLISM</td>
<td>143</td>
<td>Up</td>
<td>0.0061761</td>
<td>0.0237544</td>
</tr>
<tr>
<td>HALLMARK_ANDROGEN_RESPONSE</td>
<td>97</td>
<td>Up</td>
<td>0.0092661</td>
<td>0.0330934</td>
</tr>
<tr>
<td>HALLMARK_ADIPGENESIS</td>
<td>191</td>
<td>Up</td>
<td>0.0113278</td>
<td>0.0377592</td>
</tr>
<tr>
<td>EnsemblID</td>
<td>Symbol</td>
<td>Module</td>
<td>Module corr</td>
<td>Amyloid corr</td>
</tr>
<tr>
<td>--------------</td>
<td>--------</td>
<td>--------</td>
<td>-------------</td>
<td>--------------</td>
</tr>
<tr>
<td>ENSG0000004359</td>
<td>ADRB1</td>
<td>pink</td>
<td>0.5892136</td>
<td>0.3392804</td>
</tr>
<tr>
<td>ENSG0000008935</td>
<td>FXYD3</td>
<td>pink</td>
<td>0.5014116</td>
<td>0.3340798</td>
</tr>
<tr>
<td>ENSG0000009993</td>
<td>SERPIND1</td>
<td>pink</td>
<td>0.7548138</td>
<td>0.3037167</td>
</tr>
<tr>
<td>ENSG0000010110</td>
<td>STK4</td>
<td>pink</td>
<td>0.5754323</td>
<td>0.301087</td>
</tr>
<tr>
<td>ENSG0000010792</td>
<td>LARP4B</td>
<td>pink</td>
<td>0.7866032</td>
<td>0.3158398</td>
</tr>
<tr>
<td>ENSG0000011184</td>
<td>GCNT2</td>
<td>pink</td>
<td>0.5965631</td>
<td>0.314767</td>
</tr>
<tr>
<td>ENSG0000011227</td>
<td>BVES</td>
<td>pink</td>
<td>0.7735395</td>
<td>0.3555421</td>
</tr>
<tr>
<td>ENSG0000011526</td>
<td>IFIH1</td>
<td>pink</td>
<td>0.6922266</td>
<td>0.3714096</td>
</tr>
<tr>
<td>ENSG0000011877</td>
<td>ABCG2</td>
<td>pink</td>
<td>0.6046203</td>
<td>0.3495549</td>
</tr>
<tr>
<td>ENSG0000012449</td>
<td>F13A1</td>
<td>pink</td>
<td>0.5555863</td>
<td>0.3293968</td>
</tr>
<tr>
<td>ENSG0000013242</td>
<td>POPDC3</td>
<td>pink</td>
<td>0.9268188</td>
<td>0.3190513</td>
</tr>
<tr>
<td>ENSG0000013550</td>
<td>ACVR1B</td>
<td>pink</td>
<td>0.5486397</td>
<td>-0.3367938</td>
</tr>
<tr>
<td>ENSG0000013773</td>
<td>FXYD2</td>
<td>pink</td>
<td>0.5732391</td>
<td>-0.3402203</td>
</tr>
<tr>
<td>ENSG0000014143</td>
<td>ASXL3</td>
<td>pink</td>
<td>0.8786155</td>
<td>0.3475139</td>
</tr>
<tr>
<td>Genomic Reference</td>
<td>Gene Symbol</td>
<td>Color</td>
<td>log2Ratio</td>
<td>log2FoldChange</td>
</tr>
<tr>
<td>-------------------</td>
<td>-------------</td>
<td>-------</td>
<td>-----------</td>
<td>----------------</td>
</tr>
<tr>
<td>ENSG0000014536</td>
<td>ANK2</td>
<td>pink</td>
<td>0.6798407</td>
<td>0.3079538</td>
</tr>
<tr>
<td>ENSG0000016928</td>
<td>KCNAB1</td>
<td>pink</td>
<td>0.7899757</td>
<td>0.3033225</td>
</tr>
<tr>
<td>ENSG0000017869</td>
<td>KCTD12</td>
<td>pink</td>
<td>0.5872604</td>
<td>0.3225368</td>
</tr>
<tr>
<td>ENSG0000018226</td>
<td>FIGN</td>
<td>pink</td>
<td>0.6225719</td>
<td>0.329814</td>
</tr>
<tr>
<td>ENSG0000018505</td>
<td>SGCZ</td>
<td>pink</td>
<td>0.5829244</td>
<td>0.3119447</td>
</tr>
<tr>
<td>ENSG0000023091</td>
<td>DPP4-DT</td>
<td>pink</td>
<td>0.6969811</td>
<td>0.3110371</td>
</tr>
</tbody>
</table>