JCI insight

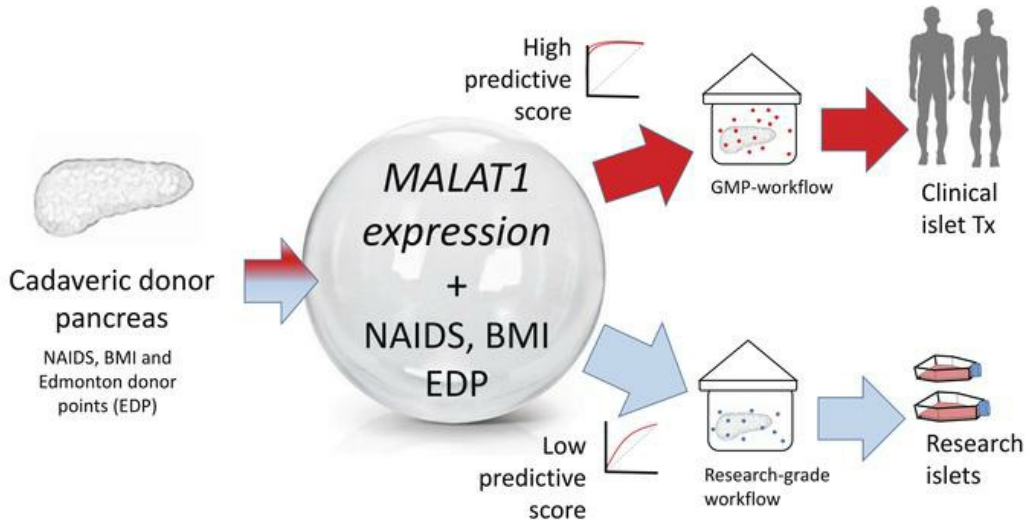
The long noncoding RNA *MALAT1* predicts human pancreatic islet isolation quality

Wilson K.M. Wong, … , Mugdha V. Joglekar, Anandwardhan A. Hardikar

JCI Insight. 2019. https://doi.org/10.1172/jci.insight.129299.

Research In-Press Preview Transplantation

Graphical abstract



The Long Non-coding RNA *MALAT1* Predicts Human Islet Isolation Quality

Wilson K.M. Wong[1], Guozhi Jiang[2], Anja E. Sørensen[3], Yi Vee Chew[4], Cody Lee-Maynard[1], David Liuwantara[4], Lindy Williams[4], Philip O'Connell[4], Louise T. Dalgaard[3], Ronald C. Ma[2], Wayne J. Hawthorne[4,*], Mugdha V. Joglekar[1,*] and Anandwardhan A. Hardikar[1,*]

Running title: The lncRNA *NEAT2*/*MALAT1* predicts islet isolation outcome

1- Diabetes and Islet Biology Group, NHMRC Clinical Trials Centre, The University of Sydney, Level 6, Medical Foundation Building, 92-94 Parramatta Road, Camperdown, NSW 2050, Australia.
2- Department of Medicine and Therapeutics, The Chinese University of Hong Kong, Prince of Wales Hospital, Hong Kong.
3- Department of Science and Environment, Roskilde University, Universitetsvej 1, DK-4000 Roskilde, Denmark
4- Centre for Transplant and Renal Research, Westmead Institute for Medical Research, Sydney University, 176 Hawkesbury Rd., Westmead, NSW 2145, Australia.

* Equal senior authors; address for correspondence:

*Wayne J Hawthorne,
Centre for Transplant and Renal Research, Westmead Institute for Medical Research, Sydney University, 176 Hawkesbury Rd., Westmead, NSW 2145, Australia.
E-mail: wayne.hawthorne@sydney.edu.au

*Mugdha V. Joglekar or *Anandwardhan A. Hardikar
Diabetes and Islet Biology Group, NHMRC Clinical Trials Centre,
The University of Sydney, Level 6, Medical Foundation Building,
92-94 Parramatta Road, Camperdown,
NSW 2050, AUSTRALIA
Phone: +61 2 9562 5071 | Fax: +61 2 9565 1863
E-mails: mugdha.joglekar@sydney.edu.au; anand.hardikar@sydney.edu.au
Web: http://www.isletbiology.com/ | http://www.ctc.usyd.edu.au/

Conflict of interest statement
The authors have declared that no conflict of interest exists.

Total word Count: 7,700
Abstract Word Count: 200
Number of figures and tables: 3 figures, 1 table

Abstract

Human islet isolation is a cost-/resource-intensive program generating islets for cell therapy in Type 1 diabetes. However, only a third of cadaveric pancreas get to clinical transplantation due to low quality/number of islets. There is a need to identify biomarker(s) that predict the quality of islets, prior to initiating their isolation. Here, we sequenced transcriptome from 18 human islet preparations stratified into three groups (Gr.1: Best quality/transplantable islets, Gr.2: Intermediary quality, Gr.3: Inferior quality/non-transplantable islets) based on routine measurements including islet purity/viability. Machine-learning algorithms involving penalized regression analyses identified 10 long-non-coding(lnc)RNAs significantly different across all group-wise comparisons (Gr1VsGr2, Gr2vsGr3, Gr1vsGr3). Two variants of Metastasis-Associated Lung Adenocarcinoma Transcript-1(*MALAT1*) lncRNA were common across all comparisons. We confirmed RNA-seq findings in a "validation set" of 75 human islet preparations. Finally, in 19 pancreas samples, we demonstrate that assessing the levels of *MALAT1* variants alone (ROC curve AUC: 0.83) offers highest specificity in predicting post-isolation islet quality and improves the predictive potential for clinical islet transplantation when combined with Edmonton Donor Points/Body Mass Index(BMI)/North American Islet Donor Score(NAIDS). We present this resource of islet-quality-stratified lncRNA transcriptome data and identify *MALAT1* as a biomarker that significantly enhances current selection methods for clinical (GMP)-grade islet isolation.

Introduction

Type 1 Diabetes (T1D) is currently managed via multiple daily insulin injections/continuous glucose monitors. Transplantation of human cadaveric islets is the only approved/healthcare supported cell therapy for T1D in countries such as Australia, Canada, France, Italy, Switzerland and the United Kingdom. In the USA, although islet auto-transplants (in chronic pancreatitis) is reimbursable, islet allo-transplantation in individuals with T1D is only through clinical trials. Usually, costs of islet isolation from a human cadaveric pancreas are very high (~US$40,000/cadaveric pancreas) due to good manufacturing practice (GMP)-grade reagents/workflows (1) necessary for clinical islet transplantation. Donor characteristics are important in selecting cadaveric pancreas for GMP-grade islet isolation procedures(2). Whilst some studies (3, 4) demonstrated donor body mass index (BMI) as a positive predictor of islet yield, islet viability and insulin secretion, the Collaborative Islet Transplant Registry (CITR) data failed to validate these observations (5). Although older donors tend to yield higher number of islets than younger donors (6, 7), islet function appears to deteriorate with age (8). Therefore, islet isolation centers often face the difficult question if costs/resource-intensive "GMP-grade" ($40,000) or the less pricy (~$700) research grade reagents should be used for an available donor cadaveric pancreas. There is a need to identify biomarkers that can predict the quality and not just higher numbers of islets within donor pancreas, prior to undertaking this resource-, time- and costs-intensive islet isolation procedure. Such a biomarker should not only independently predict the quality of islets prior to their isolation from the donor pancreas, but also add value when combined with existing parameters for donor pancreas selection.

Long non-coding (lnc) RNAs constitute a group of >200 nucleotide long RNA molecules that are not translated into proteins. LncRNAs are associated with different diseases (9, 10), including diabetes (11-14) and are known to orchestrate biologically relevant (15) signalling networks that regulate cell/tissue-specific gene expression (16-18). Recent studies have examined lncRNAs in human pancreatic islets (19-23); one of these landmark studies (22) uncovered human islet-cell lncRNAs that are associated with key pancreatic transcription factors. A recent follow-up study from the same group (19) identified 2,373 β-cell lncRNAs in a pool of 41 human islet preparations. The lncRNA *PDX1-AS1/PLUTO* (19), although expressed at low abundance in human islets, demonstrated a positive correlation with PDX1 repression in islets from Type 2 diabetes (T2D) or impaired glucose tolerance (IGT) donors. These studies describe the potential role of lncRNAs in development of IGT or T2D.

Here, we sought to examine and compare the lncRNA profiles of isolated human islet sample preparations that were stratified into three groups based on their post-isolation islet quality. Whole transcriptome RNA sequencing (RNA-seq) followed by implementation of machine-learning statistical analytical algorithms identified lncRNA candidate biomarkers of islet quality. Next, we validated the expression of these lncRNAs using TaqMan™-based real-time quantitative (q)PCR in 75 additional islet sample preparations. Finally, the potential of these lncRNAs to predict post-isolation islet quality was assessed in 19 donor pancreatic tissue samples. The lncRNA *MALAT1* predicted the quality of islets within the cadaveric pancreas, prior to their isolation. Pancreatic *MALAT1* lncRNA expression alone could predict the suitability of donor pancreas for clinical islet transplantation (ROC curve AUC: 0.83), better than the current BMI-alone criterion (ROC curve AUC: 0.79). A combined mathematical model generated to predict higher islet number (BMI) and better islet quality (*MALAT1* lncRNA expression) delivered higher predictive power (ROC curve AUC: 0.94) that significantly exceeds current practice. When combined with existing pancreas donor scores, such as the North American Islet Donor Score (NAIDS) (24) and/or the Edmonton Donor Points (2), *MALAT1* offered the highest predictive power.

We present here our discovery (n=18), validation (n=75) and prediction (n=19) datasets for this first resource of lncRNA expression in human islets stratified on the basis of their post-isolation islet quality. Since tissue quality prediction is of utmost importance in clinical islet transplantation, which currently is the only approved cell therapy for T1D, our study presents molecular tools to advance selection of cadaveric tissues for stratification into GMP-grade workflows.

Results

LncRNA expression profiles are different across groups with varying islet quality

We had access to 93 pancreatic islets samples and 19 human pancreatic tissue samples from the Westmead Islet Isolation Centre(25) (**Supplementary Figure 1**). Based on the results of standard quality control protocols(26), isolated islet samples from each of the pancreas were stratified into three groups (Group 1=high quality/transplantable, 2=Intermediary quality (good quality but low yield) and 3=inferior quality islet preparations). This stratification is performed

using their islet quality score (**Supplementary Table 1-3)** derived after assessment of four key post-isolation measurements; (i) islet purity, (ii) islet viability, (iii) β-cell viability and (iv) transplantability (**Table 1**). Human islet samples within Group 2 are also of high quality. These islet samples could not be transplanted in individuals with Type 1 diabetes as the number of IEQs from these isolations were fewer than the desired number (4000IEQ/kg of recipient body weight) for clinical transplantation. We randomly selected 18 islet samples from 3 different groups for high-depth RNA-sequencing (rRNA depletion library construction) using the HiSeq4000 platform. All the samples passed post-sequencing Quality Control (QC) and on average 77,127,891 clean-reads were obtained from these samples of the discovery set. A total of 6,983 annotated lncRNAs and 450 novel lncRNAs were identified (**Supplementary Figure 2**). An unsupervised hierarchical cluster analysis performed on the annotated lncRNAs revealed that all (*n*=5) of the Group 1 (transplantable) islet samples clustered closely with the intermediary quality Group 2 islet samples (*n*=6), but not with the inferior quality (*n*=7) Group 3 islet sample preparations (**Figure 1A**). We then compared the levels/abundance (FPKMs) of all lncRNAs expressed between the three groups of these discovery set islet samples. A significantly large number (246-322) of candidates were identified to be differentially expressed (>2-fold, P<0.05) in group-wise comparisons of lncRNAs (**Figure 1B-D**).

Penalized regression analyses identify key lncRNAs associated with islet quality

In order to identify key lncRNAs that possess discriminatory capacity to identify the high quality islet sample preparations, we performed penalized logistic/linear regression (PLR) analyses (27). In order to confirm that the outcome of this regression analysis was not a result of the sampling bias, we implemented resampling validation using bootstrapping workflows. These analyses identified 5-10 lncRNAs in each group-wise comparison (**Supplementary Table 4**). Of these, Metastasis-Associated Lung Adenocarcinoma Transcript 1 (*MALAT1*) was found to be common and significantly different across all comparisons (**Figure 2A-C**). Intriguingly, two (out of eight) splice variants of *MALAT1* (ENST00000616691.1; denoted henceforth as "*MALAT1*-1.1" and ENST00000619449.1; denoted henceforth as "*MALAT1*-9.1") were identified to be common in each bootstraps comparing Group 1 islet samples, while another *MALAT1* variant (ENST00000620902.1/"*MALAT1*-2.1") was identified in comparisons involving Group 3 islet samples (**Figure 2A-C, Supplementary Figure 3**). Receiver Operator Characteristic (ROC) curves for the significantly dysregulated lncRNAs

demonstrated the capacity of *MALAT1* variants and GAS5 lncRNAs to stratify islet samples into the relevant groups (**Figure 2D-F**).

We also performed penalized linear (islet purity, islet viability, beta-cell viability) or logistic (transplant outcome:Yes/No) regression analyses and bootstrapping on lncRNAs for each of these four parameters. Four *MALAT1* splice variants (*MALAT1*-9.1, *MALAT1*-2.1, ENST00000508832.2/"*MALAT1*-2.2" and *MALAT1*-1.1) were identified to be common across individual penalized regression comparisons (**Supplementary Table 5**) between lncRNA levels and each of the four islet quality score criteria (outlined in **Table 1**). In univariate correlation analysis, the same four *MALAT1* variants showed significant positive correlation with post-isolation islet viability, islet purity or β-cell viability index (**Supplementary Table 6**). We also identified other novel and annotated lncRNAs that show significance ($P<0.05$) in univariate analyses for each of the four islet quality score criteria (data not shown). However, these were significantly lower in abundance (at least 10-48 fold less) than the highly abundant *MALAT1* splice variants (*MALAT1*-1.1 and *MALAT1*-9.1)**.** The biological relevance of these low abundance lncRNAs can be the subject of future investigations.

The lncRNA *MALAT1* is a potential biomarker for islet isolation outcome

Since high expression of *MALAT1*-1.1 and *MALAT1*-9.1 was identified in Group 1 islet samples, we decided to confirm our findings in a validation set of 75 human islet samples for which all of the relevant information, to categorize them into three groups (Group 1, n=40; Group 2, n=18; Group 3, n=17), as well as RNA for qPCR-based *MALAT1* measurement were available. Although Group 1 islet samples had significantly higher insulin content than Group 3 islet samples (**Figure 3A**) the levels of insulin gene transcript (**Figure 3B**) were similar. The inferior quality (Group 3) islet samples had significantly lower levels of *MALAT1* gene transcript (assay-907) (**Figure 3C**), but not the other (assay-177) *MALAT1* gene transcript assay (**Figure 3D**). In order to test if these *MALAT1* variants could be biomarkers of post-isolation islet quality, it was essential to measure *MALAT1* prior to islet isolation. We assessed *MALAT1* levels in a set of 19 (pre-isolation) pancreas samples, where a small piece of tissue was stored from the middle of the pancreas at the time of transection for cannulation, usually close to the duct itself prior to islet isolation. The surgical team isolated islets from each of these 19 pancreatic samples (**Supplementary Figure 1, Figure 3E**) and categorized them into Group 1 or Group 3 islet samples (Group 1=transplantable, n=14; Group 3=non-transplantable,

n=5) while the lab team assessed the levels of *MALAT1* transcripts (assay-907 and assay-177) in each of the 19 de-identified pancreas samples. ROC curves, based on the pancreatic levels of the two *MALAT1* variants and/or donor BMI returned an AUC of 0.79 for BMI alone, 0.83 for both *MALAT1* assays (907 + 177) and 0.94 for both *MALAT1* variants+BMI (**Figure 3F,G**). Since islet isolation centers across the world follow several criteria, we tested if *MALAT1* measurements had any added benefit to current donor pancreas selection score. A combination of *MALAT1* measurements with any of the standard scores, such as Edmonton Donor Points or the NAIDS(24), offers the highest predictive power (**Figure 3G**).

The stimulation index for Group 1 islet samples appeared to be higher, although there was no statistically significant difference across the groups (**Supplementary Figure 4A**). Islet samples from the three groups were transplanted under the kidney capsule of diabetic animals and the time for diabetes reversal was monitored to further classify the transplanted animals as "Early" or "Late" responders (**Supplementary Figure 4B**). Higher proportion of mice showed diabetes reversal within the first three weeks post transplantation (Early responders) for Group 1 and 2 islets as compared to Group 3 islet recipients (**Supplementary Figure 4C**).

Discussion

The present study was designed to generate lncRNA profiles of human islet sample preparations that were stratified based on their post-isolation islet quality. Currently, there are no predictive biomarkers of post-isolation islet quality and data from the CITR (5) have failed to validate the potential of clinical characteristics, such as BMI, in predicting the quality of islet sample preparations isolated from each cadaveric human pancreas. There is a need to identify molecular biomarkers that could allow real-time analysis and prediction of islet quality.

Islets were categorized into three groups based on their post-isolation islet quality variables (purity, viability, beta cell viability and transplantability; **Table 1**). The functionality of all isolated human islets was determined using in vitro glucose stimulated insulin secretion (GSIS) assay and in vivo islet transplantation in diabetic mice. The GSIS and in vivo mouse transplant outcomes were included to assess the function of islets from the three different groups (**Supplementary Figure 4A-C**). Intriguingly, majority of the Group 3 islet samples offer a late

response (>3 weeks) for diabetes reversal (**Supplementary Figure 4B,C**). Although 71% of Group 1 islet samples (prediction study set) were transplanted into individuals with Type 1 diabetes, the beta-scores (28) for these are not shown as most of our islet recipients had a second transplant within three months of receiving their first islet transplant. Thus changes in HbA1c over three months post-transplantation would be confounded as a result of the successive transplant. The current study was not designed/powered to predict islet transplantation outcome. This study presents the potential of *MALAT1* variants as a biomarker to help accurately predict the quality of islets before initiating human islet isolation from the pancreas. The prediction of the outcome of islet transplant procedure, is a separate question, which will be the focus of future investigation.

A recent study demonstrated the potential of assessing insulin immuno-positive area in predicting islet isolation outcomes with 89% sensitivity and 76% specificity (29). Our data demonstrate that *MALAT1* expression alone (AUC: 0.83, maximum specificity of 100% and a sensitivity of 80%) or along with Edmonton Donor Points, BMI and/or NAIDS can accurately stratify donor pancreas to group 1 (transplantable) or group 3 (not transplantable) prior to the initiation of islet isolation. In addition to *MALAT1*, lncRNA such as GAS5 (variant: ENST00000450589.5) may add better discriminatory value to identify the impact of glucocorticoids (30) and stratify Group 2 vs Group 3 human islet sample preparations. *MALAT1* was originally identified to be the most abundant lncRNA in the pancreas (31), purified human islets (19), as well as being highly abundant in purified human β-cells (19). Intriguingly, mouse knockout (KO) of *MALAT1* showed no differences in the histology of pancreas/islets (32, 33), and did not affect the proliferation or viability in hepatocellular carcinoma cells (32). It is currently unclear as to why *MALAT1* is associated with high quality (Group 1) islet sample preparations. *MALAT1* is encoded within an active enhancer cluster offering multiple binding sites for pancreatic islet transcription factors (34). Thus, a possible role of *MALAT1* could be in facilitating (pro-)endocrine gene expression within human islets. Apart from the role of *MALAT1* in islet cells, it is intriguing to note that *MALAT1* is regulated by hypoxia (35-37). Indeed, hypoxia can be a major determinant of post-isolation islet quality (38) and therefore the regulation of *MALAT1* under hypoxic conditions can impact on islet survival. Knock down of *MALAT1* leads to apoptosis (39), whereas increase in *MALAT1* expression has been shown to inhibit apoptosis (40). It can therefore be speculated that islets that do not demonstrate the capacity to upregulate *MALAT1* expression following exposure to

hypoxic conditions would be more vulnerable to apoptotic cell death. Therefore, levels of *MALAT1* detected in human pancreas prior to islet isolation may provide a measure of the post-isolation islet quality. Indeed the abundance of the two *MALAT1* variants (*MALAT1*-1.1 and -9.1) negatively correlated with post-isolation islet cell death (**Supplementary Figure 5**).

Our study strengths are (i) the use of an unprejudiced high depth sequencing platform for human islets from three different groups (n=18), (ii) use of sophisticated, unbiased data analytical methods with resampling validation (bootstrapping), (iii) the independent replication of these findings in a separate set (n=75) of human islet samples, and (iv) the contribution of *MALAT1* variants to improving the predictive power of Edmonton Donor Points, BMI and/or NAIDS in classifying islet samples from a set (n=19) of pre-isolation pancreatic tissue samples. The study limitations are the lack of larger number of matching high quality discovery samples (which we have attempted to address via resampling/bootstrapping strategies) and the need to replicate our findings in independent islet isolation center(s). It is essential to note that *MALAT1* variants alone demonstrate better specificity to predict post-isolation quality of islets, than Edmonton Donor Points, BMI or NAIDS. We do not know whether *MALAT1* can predict clinical outcomes or graft functions after transplant, which would be an interesting question to pursue.

The present study identifies lncRNA biomarkers such as *MALAT1*, which could potentially facilitate the prediction of islet isolation quality, within minutes of receiving the cadaveric pancreas. Currently, commercially available kits have been demonstrated (41) to offer the capacity to isolate and directly measure RNA in minutes. Advances in nanotechnology further enable reliable as well as costs- and time-efficient assessment of RNA quantity (42) and can be used to measure biomarker levels (43). Cross-discipline technological advances can be easily used to develop a rapid detection kit, similar to a pregnancy test, for clinical biomarkers (such as MALAT1 variants). Platforms to measure microRNAs and lncRNAs using nanosensor (44) and photonics-based technologies (45) can detect ultra-low (sub-picomolar to attomolar) levels of ncRNAs from cell-lysates. Since MALAT1 is a highly abundant lncRNA in human pancreas, the development of such rapid detection tests for MALAT1 variants is achievable, and is indeed one of our research focus in the coming years. Nanotechnology/photonics-based platforms offer a rapid (2-5 minutes) and low costs (<$1) method that could improve donor selection and enhance the development of tests that ultimately help us to stratify donors with high accuracy, so as to meet the release criteria for clinical islet transplantation. Since the time from cross clamp at retrieval of donor pancreas to initiating islet isolation is around 9.8$\pm$0.179

(mean±SEM) hours (46), the identification of MALAT1 lncRNA variants allows us sufficient window to use this test that could change current practice. Measurement of MALAT1 variants will improve donor selection whilst offering the other (non-selected) pancreas for research-grade isolation workflows. Our study underscores the importance of existing donor pancreas selection criteria (Edmonton Donor Points, BMI and NAIDS) and presents the added predictive power offered by the new lncRNA biomarkers. As discussed above, although existing criteria (such as donor BMI) may be good estimators of islet number, the measurement of *MALAT1* lncRNA variants in the pancreas provides a 'readout' of their quality, significantly increasing the predictive power for selection of donor pancreas in clinical islet transplantation.

Study impact: The present study provides the capacity to make an informed choice for research or clinical islet isolation workflow. Clinical islet isolation costs using GMP-grade facilities and reagents, and involving screening, organ recovery, transportation and islet preparation is estimated around US$ 40,000 per preparation (1). Based on CITR data, ~US$ 97,000,000 would have been spent on 2,421 isolations (47) of which, only 750 preparations (48) were deemed to be suitable for human islet transplantation. The current study provides the means to not only direct the best donor tissues for clinical transplants, but also to make appropriate decisions for the less cost-intensive research grade islet isolation workflows. Our data corroborate previous findings indicating the high presence of *MALAT1* lncRNA in islets. The stratification of islets based on the quality of islet preparations help to understand the role of *MALAT1* lncRNA variants as a biomarkers for predicting islet isolation outcome. We also provide the first report of human islet lncRNA expression profiles in quality stratified set of islet samples to the research community through this publication.

Data availability: We intend to replicate these findings at other islet isolation centers and welcome collaborative or independent assessment of our predictive algorithms. All data from our lncRNA-sequencing studies are uploaded to the Gene Expression Omnibus (GEO) database and are available through the study accession # GSE134068. We welcome future collaborations to analyse and validate these in other centers. Such studies through multiple islet isolation centers would help in confirming the potential of these *MALAT1* lncRNA variants in improving the prediction of pancreatic islet isolation procedure.

Methods

Human Pancreatic Islets

Human tissue samples were obtained through the human research ethics committee (HREC) approvals X16-0289 (previously X12-0176) and the HREC/12/RPAH/282 at the University of Sydney and research consented-tissues were through the National Islet Transplantation Program at the Westmead Hospital, University of Sydney, Australia (AU RED/HREC/15/WMEAD/284). Pancreatic tissue samples were taken from the middle of the pancreas at the time of transection of the pancreas for cannulation, usually close to the duct itself. Islets were isolated following the standard protocol carried out by the National Islet Transplantation Unit at the Westmead Hospital, University of Sydney, Australia (25). Standard set of donor characteristics were recorded by the team and are used in calculation of the islet quality score (Table 1), Edmonton donor points (2, 6) and the North American Islet Donor Score (NAIDS) (24). Islet samples were categorized into three groups based on their islet quality score and presented in this study as Group 1 (high quality), Group 2 (intermediary quality) and Group 3 (inferior quality) islet sample preparations. In order to compare the predictive power of lncRNA biomarkers along with existing clinical and donor characteristics, we also compared the classification of islet quality based on Edmonton donor points as well as NAIDS.

Islet quality Score

Islet quality score (presented in **Supplementary Tables 1, 2, 3**) was mathematically computed as the average score divided by the number of available criteria (**Table 1**) from the scoring of each criteria listed in **Table 1**. The methodology for each of these assessments is provided below:

(i) Post-isolation islet purity

Islet purity was assessed as described elsewhere (49). Briefly, post-isolation islet purity was assessed by using Dithizone stain (3mg/mL). Triplicate aliquots of known volumes are sampled from the final islet cell preparation. The total number of islets counted using standard criteria and percentage of acinar tissue quantified and scored in each of the aliquots (IEQ >4000 IEQ/kg of recipient body weight) as detailed elsewhere (49).

(ii) Post-isolation islet viability

Post-isolation islet viability was determined using fluorescent labelling of cell preparations using DNA-binding dyes, so as to differentiate between live and dead cells. Fluorescein diacetate (FDA; 24µM) and propidium iodide (PI; 750µM) solutions were used to stain 10

replicates of cells with compromised membranes. Westmead Hospital, Australia utilises a cut off of 70% viability as a minimum for release of the product for clinical transplantation. FDA diffuses passively across the cell membrane and is converted to fluorescein by non-specific esterases in the cytoplasm, causing live cells to fluoresce green under a 490nm excitation wavelength. Dead or dying cells have compromised cell membranes and do not show cytoplasmic esterase activity and therefore do not fluoresce green. Counterstaining with PI allows better identification of these dead/dying cells as these cells (nuclei) will take up PI, fluorescing red at 545nm. Triplicate aliquots of a known volume are sampled from the final islet cell preparation and stained with FDA and PI. A total of at least 100 islets are stained and cells counted and quantified as to the percentages of viable (green) versus dead (red) cells. A total accumulated score is then calculated and the mean value taken as the viability score.

(iii) Post-isolation islet β-cell viability index

Flow cytometric assessment was carried out on the human islet cells following an established method developed by Ichii *et al* (50). This method simultaneously determines the β-cell composition, viability and apoptotic cell percentage in enzymatically dispersed single cells from islets. Zinc-binding dye Newport Green (NPG) is used to determine beta cell composition (NPG positive staining), while apoptotic cells were probed using tetramethylrhodamine ethyl ester (TMRE) and the membrane-impermeant 7-aminoactinomycin D (7-AAD) staining (50). Our centre at Westmead Hospital, Australia conducts flow cytometric analysis on islet cells post-culture to determine this beta cell viability index, defined as [% β-cells x % Viable β-cells]/10,000, with indices of 0.5 or higher considered as satisfactory for release for transplantation.

(iv) Transplantability

The suitability of the islet preparation for human transplantation was carried out using the following criteria: The release criteria formally accepted for our program are based on islet count per recipient weight (5,000 – 20,000 IEQ/kg for the first transplant, and 3,000 – 20,000 for following transplants), with purity ≥ 30%, viability ≥ 70%, endotoxin concentration < 5 endotoxin units (EU) per kg recipient weight, and no detectable organisms in a Gram stain prior to transplant, in addition to a glucose stimulation index (ratio of stimulated insulin secretion: basal insulin secretion) > 1 and a beta cell viability index with indices of 0.5 or higher.

Islet insulin content

Insulin content was measured in two replicates for each post-isolation islet preparation. One ml of Azol was added to each replicate prior to sonication. Sonicated aliquots were then mixed with FSA solution and appropriately processed/diluted for total insulin content analysis using manufacturer's method for immunoassay on the Architect (Abbott Diagnostics, USA). Data was normalized to DNA content (Quant-IT Pico Green ds DNA Assay Kit, Life Technologies). Values are presented in total insulin content for the entire islet preparation, after normalizing for amount of DNA.

Human islet transplantation into mice

Athymic mice were rendered diabetic with streptozotocin, before transplantation of between 2000 to 3000 IEQ of (post-culture) islets beneath the kidney capsule. Blood sugar levels were monitored post-transplant to determine success of islet transplant procedure (defined as either a reduction in non-fasting blood glucose levels to under 11 mmol/L (200 mg/dL), or halving in non-fasting blood glucose levels from the diabetic state, in two successive measurements 48 to 72 hours apart). An intraperitoneal glucose tolerance test was conducted 3-4 weeks post-transplant to assess glucose clearance and hence determine if success in islet function was achieved post-transplant.

Human islet stimulation index

The functional capacity of human islets was assessed by measuring insulin secretion in response to glucose. Briefly, (post-culture) islets were incubated at 37°C and 5% CO_2 for one hour in RPMI 1640 with either 2.8mM glucose or 25mM glucose. For each islet preparation, six replicates were exposed to 2.8mM glucose and 25 mM glucose. Supernatants were collected and their insulin content was measured using an Architect (Abbott Diagnostics, USA). The stimulation index was calculated by dividing the insulin content from the 25mM glucose incubation sample by the insulin content in the 2.8mM glucose incubation sample. Higher values indicate a better insulin secretion response to glucose stimulation, and therefore a higher functional capacity (26, 51).

RNA isolation and QC

Total RNA was isolated from 2000 to 5000 islet equivalent's using the manufacturer's TRIzol® RNA Isolation (Thermo Fisher Scientific, Waltham, MA) protocol with minor modification (52). Details and QC of the selected samples (discovery set, n = 18) for RNA-seq are provided in supplemental online material (SOM; **Supplementary Table 7)**. RNA integrity

number (RIN) was assessed for every sample and only those with a RIN value ≥7.4 were taken for discovery analysis.

High-depth RNA-sequencing

High-depth RNA-sequencing (rRNA depletion library construction) was carried out using the HiSeq4000, 150 Paired-end (PE) reads platform (Novogene, Beijing, China) on n = 18 human pancreatic islet preparations (**Supplementary Table 7**). An input of 3μg RNA per a sample was used for sequencing. Epicentre Ribo-zeroTM rRNA Removal Kit (Epicentre, USA) and ethanol precipitation was used to remove the ribosomal RNA and rRNA free residues for each sample respectively. Subsequently, sequencing libraries were generated using the rRNA-depleted RNA with the NEBNext® UltraTM Directional RNA Library Prep Kit for Illumina® (NEB, USA) through following manufacturer's protocol. The clustering of the index-coded samples was performed on a cBot Cluster Generation System using HiSeq PE Cluster Kit cBot-HS (Illumina) following the manufacturer's instructions. After cluster generation, the library preparations were sequenced on an Illumina Hiseq platform and paired-end reads were generated. Fragments Per Kilobase of transcript sequence per Millions base pair sequenced (FPKM) were calculated for both lncRNAs and coding genes in each sample using Cuffdiff (v2.1.1). FPKMs are the means mapped and calculated based on the length of the fragments and read count mapped to this fragment (53). Results are presented in normalized FPKMs

RNA-Seq analysis

In-house perl scripts were used to remove reads containing adapter, reads on containing ploy-N and low quality reads from the raw data (reads) leaving only the clean data (reads). The index of the reference human GRCh38 (hg38) genome (//ftp.ensembl.org/pub/release-82/fasta/homo_sapiens/dna/) and the GTF files (ftp://ftp.ensembl.org/pub/release-82/gtf/homo_sapiens/) were used for annotation (obtained from the genome website) and built using Bowtie v2.0.6. TopHat v2.0.9 was used to align paired-end clean reads to the reference genome. On average 77,127,891 clean-reads were obtained from each of the 18 samples that passed QC in this discovery set. PhyloCSF (phylogenetic codon substitution frequency) (v20121028) was used to identify and distinguish the characteristics to align conserved coding regions (54). Multi-species genome sequence alignments were built and run on phyloCSF with default parameters. Transcripts predicted with coding potential (by all or either of the four tools (CMCI, CPS, Pfam-scan and phyloCSF)) were removed, whilst transcripts without coding potential were identified to be the candidate set of lncRNAs.

Real-time qPCR

High Capacity cDNA reverse transcription kit (Thermo Fisher Scientific, Waltham, MA) was used followed by TaqMan® real-time qPCR for assays. Briefly, synthesis of cDNA from RNA was carried out using the High Capacity cDNA reverse transcription kit (Thermo Fisher Scientific, Waltham, MA). TaqMan® real-time qPCR was performed in 5µl reactions using 96-well plates with 33.3 ng input cDNA with TaqMan® Fast Universal PCR Master Mix (Thermo Fisher Scientific, Waltham, MA). Selected TaqMan® primer/probe gene expression assays are provided in **Supplementary Table 8**. We used two different *MALAT1* gene transcript assays (907 and 177), which span different regions of the human *MALAT1* gene. *MALAT1* gene transcript assay-907 captures the *MALAT1* variants-9.1 and -2.1, whilst *MALAT1* gene transcript assay-177 only covers *MALAT1* variant-1.1. Results were normalized to 18s rRNA values. Real-time quantitative PCR was carried using the ViiA7 platform (Thermo Fisher Scientific, Waltham, MA).

Statistics

All statistical analyses were performed using GraphPad Prism V7 (GraphPad Software), SPSS, R glmnet, penalized or corrplot packages. A two-tailed distribution, with two-sample unequal variance Students *t* test was used to compare specific lncRNAs between two groups for the RNA-seq data. To identify a subset of lncRNA presenting strongest associations with the three categorized groups of islet preparations, L1-penalized logistic/linear regression (Lasso) techniques were used, as described elsewhere (27). The lncRNAs identified by Lasso analyses were confirmed by resampling validation/bootstrapping analyses (for 1000 iterations). During bootstrapping, ~37% of samples were randomly removed and the same number of different (randomly selected) samples duplicated so that the total number of samples remained same in each of the 1000 iteration. Resampling validation was an important part of the process so as to eliminate any sampling bias. A Wilcoxon P-value cut-off of $P>0.5$ was also applied in each iteration. Univariate logistic and linear regression was carried for each comparison to examine the association of each independent variable with outcome. A one-way ANOVA test with multiple comparisons was used for all group comparisons for the qPCR data. R-scripts for the above machine-learning algorithms would be made available following publication.

Study approval

Human tissue samples were obtained through the human research ethics committee (HREC) approvals X16-0289 (previously X12-0176) and the HREC/12/RPAH/282 at the University of Sydney and research consented-tissues were through the National Islet Transplantation Program at the Westmead Hospital, University of Sydney, Australia (AU RED/HREC/15/WMEAD/284).

Animal studies were approved by the animal ethics committee (AEC) (protocol 4198.10.12 and 5146.10.17).

Author contributions:

WKMW: Data analysis, experimentation, write up, revision. GJ, AES, RCM, LTD: Biostatistics, data analysis/validation, interpretations. CLM, YVC, DL, LW, PO'C, WJH: Data analysis, interpretation, islet isolation/banking. AAH, MVJ: study design, guarantors of the work, data analysis, interpretations and manuscript writing and revision. All authors agreed on final revised draft.

Acknowledgements:

Funding: Marian and E.H Flack Trust (WJH; FLACK-2016-17), Juvenile Diabetes Research Foundation (JDRF) Australia T1D CRN (AAH; 4-CDA-2016-228-M-B), JDRF International (MVJ; 3-APF-2016-178-A-N), Australian Research Council (AAH), Sydney Medical Foundation (AAH, MVJ) and the Danish Diabetes Academy funded by the Novo Nordisk Foundation (AAH, LTD, AES). AAH, MVJ, WW and WJH have intellectual property.

References

1. Guignard AP, Oberholzer J, Benhamou PY, Touzet S, Bucher P, Penfornis A, et al. Cost analysis of human islet transplantation for the treatment of type 1 diabetes in the Swiss-French Consortium GRAGIL. *Diabetes Care.* 2004;27(4):895-900.
2. O'Gorman D, Kin T, Murdoch T, Richer B, McGhee-Wilson D, Ryan EA, et al. The standardization of pancreatic donors for islet isolations. *Transplantation.* 2005;80(6):801-6.
3. Brandhorst H, Brandhorst D, Hering BJ, Federlin K, and Bretzel RG. Body mass index of pancreatic donors: a decisive factor for human islet isolation. *Exp Clin Endocrinol Diabetes.* 1995;103 Suppl 2:23-6.
4. Nano R, Clissi B, Melzi R, Calori G, Maffi P, Antonioli B, et al. Islet isolation for allotransplantation: variables associated with successful islet yield and graft function. *Diabetologia.* 2005;48(5):906-12.
5. Balamurugan AN, Naziruddin B, Lockridge A, Tiwari M, Loganathan G, Takita M, et al. Islet product characteristics and factors related to successful human islet transplantation from the Collaborative Islet Transplant Registry (CITR) 1999-2010. *Am J Transplant.* 2014;14(11):2595-606.
6. Lakey JR, Warnock GL, Rajotte RV, Suarez-Alamazor ME, Ao Z, Shapiro AM, et al. Variables in organ donors that affect the recovery of human islets of Langerhans. *Transplantation.* 1996;61(7):1047-53.
7. Toso C, Oberholzer J, Ris F, Triponez F, Bucher P, Demirag A, et al. Factors affecting human islet of Langerhans isolation yields. *Transplant Proc.* 2002;34(3):826-7.
8. Niclauss N, Bosco D, Morel P, Demuylder-Mischler S, Brault C, Milliat-Guittard L, et al. Influence of donor age on islet isolation and transplantation outcome. *Transplantation.* 2011;91(3):360-6.
9. de Almeida RA, Fraczek MG, Parker S, Delneri D, and O'Keefe RT. Non-coding RNAs and disease: the classical ncRNAs make a comeback. *Biochem Soc Trans.* 2016;44(4):1073-8.
10. Anastasiadou E, Jacob LS, and Slack FJ. Non-coding RNA networks in cancer. *Nat Rev Cancer.* 2018;18(1):5-18.
11. Guay C, Jacovetti C, Nesca V, Motterle A, Tugay K, and Regazzi R. Emerging roles of non-coding RNAs in pancreatic beta-cell function and dysfunction. *Diabetes Obes Metab.* 2012;14 Suppl 3:12-21.
12. Bunch H. Gene regulation of mammalian long non-coding RNA. *Mol Genet Genomics.* 2018;293(1):1-15.
13. Nie L, Wu HJ, Hsu JM, Chang SS, Labaff AM, Li CW, et al. Long non-coding RNAs: versatile master regulators of gene expression and crucial players in cancer. *Am J Transl Res.* 2012;4(2):127-50.
14. Mirza AH, Kaur S, and Pociot F. Long non-coding RNAs as novel players in beta cell function and type 1 diabetes. *Hum Genomics.* 2017;11(1):17.
15. Pelechano V, and Steinmetz LM. Gene regulation by antisense transcription. *Nat Rev Genet.* 2013;14(12):880-93.
16. Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F, et al. GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res.* 2012;22(9):1760-74.
17. Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, et al. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.* 2011;25(18):1915-27.
18. Mattick JS, and Makunin IV. Non-coding RNA. *Hum Mol Genet.* 2006;15 Spec No 1:R17-29.

19. Akerman I, Tu Z, Beucher A, Rolando DMY, Sauty-Colace C, Benazra M, et al. Human Pancreatic beta Cell lncRNAs Control Cell-Specific Regulatory Networks. *Cell Metab.* 2017;25(2):400-11.
20. Benner C, van der Meulen T, Caceres E, Tigyi K, Donaldson CJ, and Huising MO. The transcriptional landscape of mouse beta cells compared to human beta cells reveals notable species differences in long non-coding RNA and protein-coding gene expression. *BMC Genomics.* 2014;15:620.
21. Li B, Bi CL, Lang N, Li YZ, Xu C, Zhang YQ, et al. RNA-seq methods for identifying differentially expressed gene in human pancreatic islet cells treated with pro-inflammatory cytokines. *Mol Biol Rep.* 2014;41(4):1917-25.
22. Moran I, Akerman I, van de Bunt M, Xie R, Benazra M, Nammo T, et al. Human beta cell transcriptome analysis uncovers lncRNAs that are tissue-specific, dynamically regulated, and abnormally expressed in type 2 diabetes. *Cell Metab.* 2012;16(4):435-48.
23. Singer RA, and Sussel L. Islet Long Noncoding RNAs: A Playbook for Discovery and Characterization. *Diabetes.* 2018;67(8):1461-70.
24. Wang LJ, Kin T, O'Gorman D, Shapiro AMJ, Naziruddin B, Takita M, et al. A Multicenter Study: North American Islet Donor Score in Donor Pancreas Selection for Human Islet Isolation for Transplantation. *Cell Transplant.* 2016;25(8):1515-23.
25. O'Connell PJ, Holmes-Walker DJ, Goodman D, Hawthorne WJ, Loudovaris T, Gunton JE, et al. Multicenter Australian trial of islet transplantation: improving accessibility and outcomes. *Am J Transplant.* 2013;13(7):1850-8.
26. Ricordi C. Quantitative and qualitative standards for islet isolation assessment in humans and large mammals. *Pancreas.* 1991;6(2):242-4.
27. Goeman JJ. L1 penalized estimation in the Cox proportional hazards model. *Biom J.* 2010;52(1):70-84.
28. Ryan EA, Paty BW, Senior PA, Lakey JR, Bigam D, and Shapiro AM. Beta-score: an assessment of beta-cell function after islet transplantation. *Diabetes Care.* 2005;28(2):343-7.
29. Berkova Z, Saudek F, Girman P, Zacharovova K, Kriz J, Fabryova E, et al. Combining Donor Characteristics with Immunohistological Data Improves the Prediction of Islet Isolation Success. *J Diabetes Res.* 2016;2016:4214328.
30. Lucafo M, Di Silvestre A, Romano M, Avian A, Antonelli R, Martelossi S, et al. Role of the Long Non-Coding RNA Growth Arrest-Specific 5 in Glucocorticoid Response in Children with Inflammatory Bowel Disease. *Basic Clin Pharmacol Toxicol.* 2018;122(1):87-93.
31. Ji P, Diederichs S, Wang W, Boing S, Metzger R, Schneider PM, et al. MALAT-1, a novel noncoding RNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer. *Oncogene.* 2003;22(39):8031-41.
32. Eissmann M, Gutschner T, Hammerle M, Gunther S, Caudron-Herger M, Gross M, et al. Loss of the abundant nuclear non-coding RNA MALAT1 is compatible with life and development. *Rna Biol.* 2012;9(8):1076-87.
33. Nakagawa S, Ip JY, Shioi G, Tripathi V, Zong XY, Hirose T, et al. Malat1 is not an essential component of nuclear speckles in mice. *Rna.* 2012;18(8):1487-99.
34. Pasquali L, Gaulton KJ, Rodriguez-Segui SA, Mularoni L, Miguel-Escalada I, Akerman I, et al. Pancreatic islet enhancer clusters enriched in type 2 diabetes risk-associated variants. *Nat Genet.* 2014;46(2):136-43.
35. Brock M, Schuoler C, Leuenberger C, Buhlmann C, Haider TJ, Vogel J, et al. Analysis of hypoxia-induced noncoding RNAs reveals metastasis-associated lung adenocarcinoma transcript 1 as an important regulator of vascular smooth muscle cell proliferation. *Exp Biol Med (Maywood).* 2017;242(5):487-96.
36. Lelli A, Nolan KA, Santambrogio S, Goncalves AF, Schonenberger MJ, Guinot A, et al. Induction of long noncoding RNA MALAT1 in hypoxic mice. *Hypoxia (Auckl).* 2015;3:45-52.

37. Salle-Lefort S, Miard S, Nolin MA, Boivin L, Pare ME, Debigare R, et al. Hypoxia upregulates Malat1 expression through a CaMKK/AMPK/HIF-1alpha axis. *Int J Oncol.* 2016;49(4):1731-6.
38. Lau J, Henriksnas J, Svensson J, and Carlsson PO. Oxygenation of islets and its role in transplantation. *Curr Opin Organ Transplant.* 2009;14(6):688-93.
39. Zhao M, Wang S, Li Q, Ji Q, Guo P, and Liu X. MALAT1: A long non-coding RNA highly associated with human cancers. *Oncol Lett.* 2018;16(1):19-26.
40. Xin JW, and Jiang YG. Long noncoding RNA MALAT1 inhibits apoptosis induced by oxygen-glucose deprivation and reoxygenation in human brain microvascular endothelial cells. *Exp Ther Med.* 2017;13(4):1225-34.
41. Van Peer G, Mestdagh P, and Vandesompele J. Accurate RT-qPCR gene expression analysis on cell culture lysates. *Sci Rep.* 2012;2:222.
42. Dimov IK, Lu R, Lee EP, Seita J, Sahoo D, Park SM, et al. Discriminating cellular heterogeneity using microwell-based RNA cytometry. *Nat Commun.* 2014;5:3451.
43. Sharma TK, Ramanathan R, Weerathunge P, Mohammadtaheri M, Daima HK, Shukla R, et al. Aptamer-mediated 'turn-off/turn-on' nanozyme activity of gold nanoparticles for kanamycin detection. *Chem Commun (Camb).* 2014;50(100):15856-9.
44. Joshi GK, Deitz-McElyea S, Liyanage T, Lawrence K, Mali S, Sardar R, et al. Label-Free Nanoplasmonic-Based Short Noncoding RNA Sensing at Attomolar Concentrations Allows for Quantitative and Highly Specific Assay of MicroRNA-10b in Biological Fluids and Circulating Exosomes. *ACS Nano.* 2015;9(11):11075-89.
45. Vaisocherova H, Sipova H, Visova I, Bockova M, Springer T, Ermini ML, et al. Rapid and sensitive detection of multiple microRNAs in cell lysate by low-fouling surface plasmon resonance biosensor. *Biosens Bioelectron.* 2015;70:226-31.
46. Andres A, Kin T, O'Gorman D, Livingstone S, Bigam D, Kneteman N, et al. Clinical islet isolation and transplantation outcomes with deceased cardiac death donors are similar to neurological determination of death donors. *Transpl Int.* 2016;29(1):34-40.
47. Investigators CITRC. CITR Coordinating Center; 2016.
48. Bruni A MM, Shapiro J.A.M. In: Ledbetter D.J JPRV ed. *Endocrine Surgery in Children*. Springer, Berlin, Heidelberg; 2018:181-96.
49. Hawthorne WJ. Necessities for a Clinical Islet Program. *Adv Exp Med Biol.* 2016;938:67-88.
50. Ichii H, Inverardi L, Pileggi A, Molano RD, Cabrera O, Caicedo A, et al. A novel method for the assessment of cellular composition and beta-cell viability in human islet preparations. *Am J Transplant.* 2005;5(7):1635-45.
51. Ricordi C, Lacy PE, Finke EH, Olack BJ, and Scharp DW. Automated method for isolation of human pancreatic islets. *Diabetes.* 1988;37(4):413-20.
52. Joglekar MV, and Hardikar AA. Isolation, expansion, and characterization of human islet-derived progenitor cells. *Methods Mol Biol.* 2012;879:351-66.
53. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol.* 2010;28(5):511-5.
54. Lin MF, Jungreis I, and Kellis M. PhyloCSF: a comparative genomics method to distinguish protein coding and non-coding regions. *Bioinformatics.* 2011;27(13):i275-82.

Figures and figure legends:

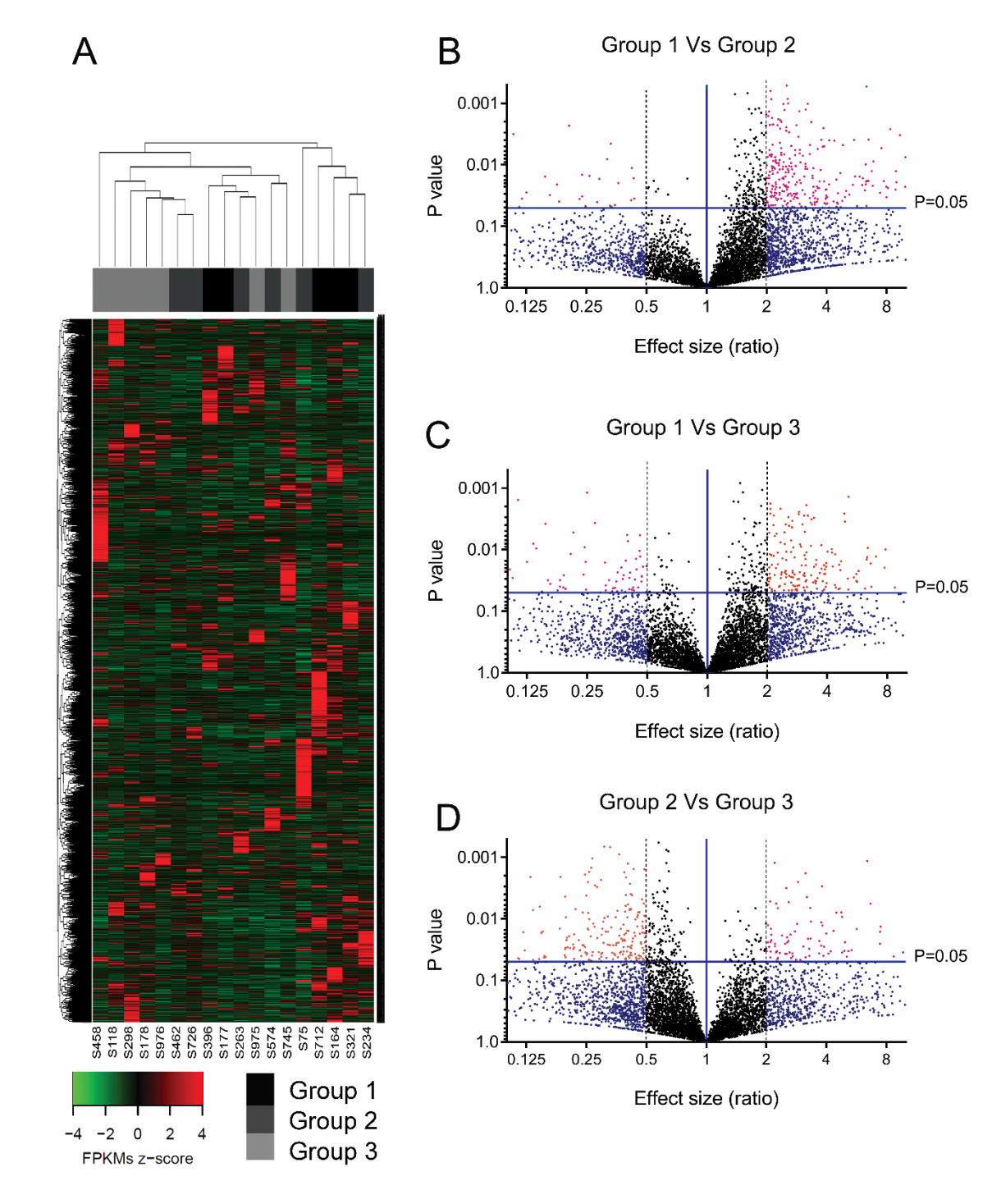


Figure 1: (A) Unsupervised hierarchical euclidean (complete) cluster heatmap of the annotated (6721) lncRNAs in *n*=18 human islets, categorised into three groups. **(B-D)** Volcano plot of annotated lncRNAs in the human islets categorized Group 1 vs Group 2 (**B**), Group 1 vs Group 3 (**C**) and Group 2 vs Group 3 (**D**). The effect size differences are depicted on the X-axis while the −log10 P-value is depicted on the Y-axis. The horizontal blue line represents the significant P-value=0.05, while the vertical blue line represents no difference.

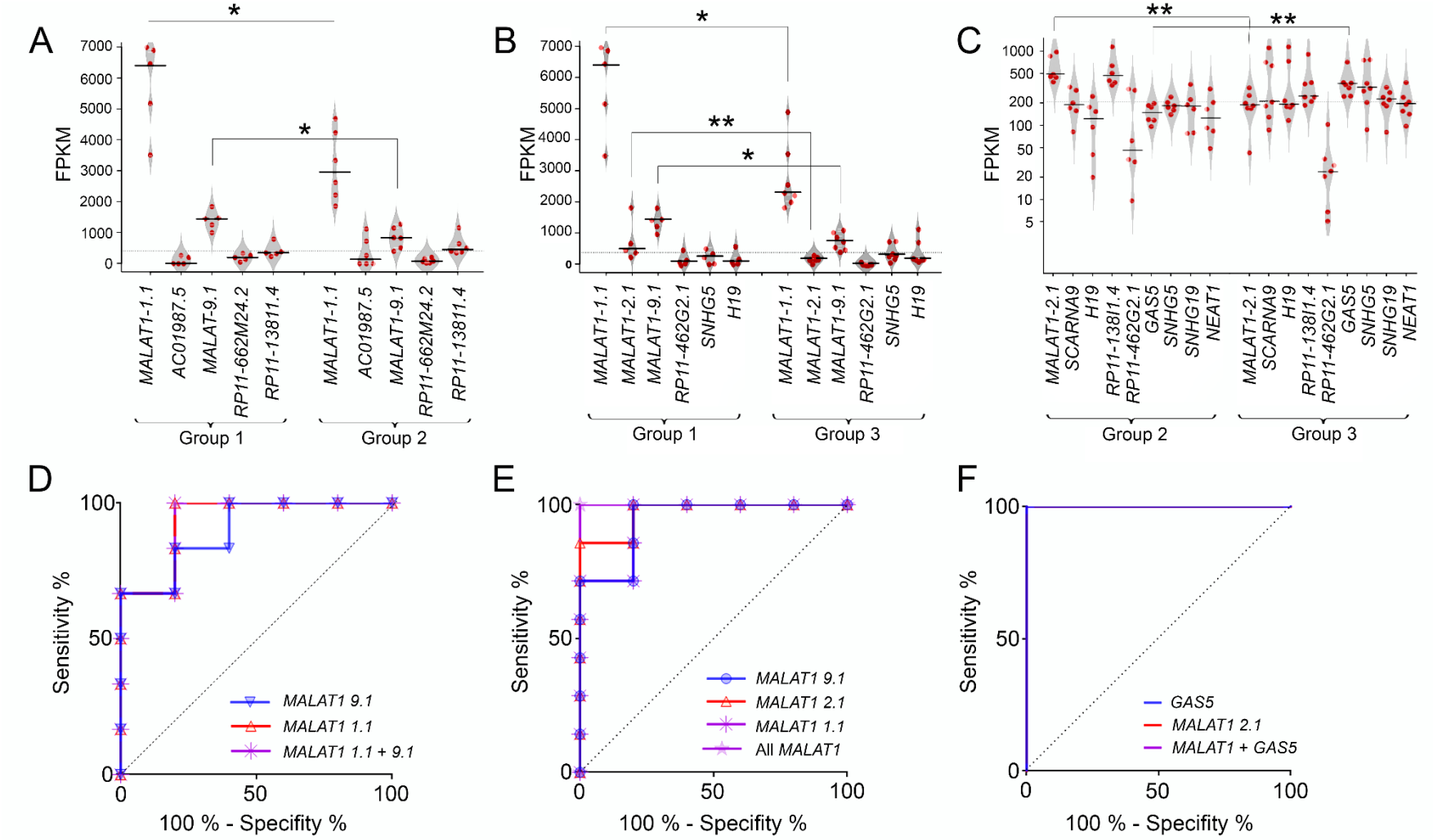
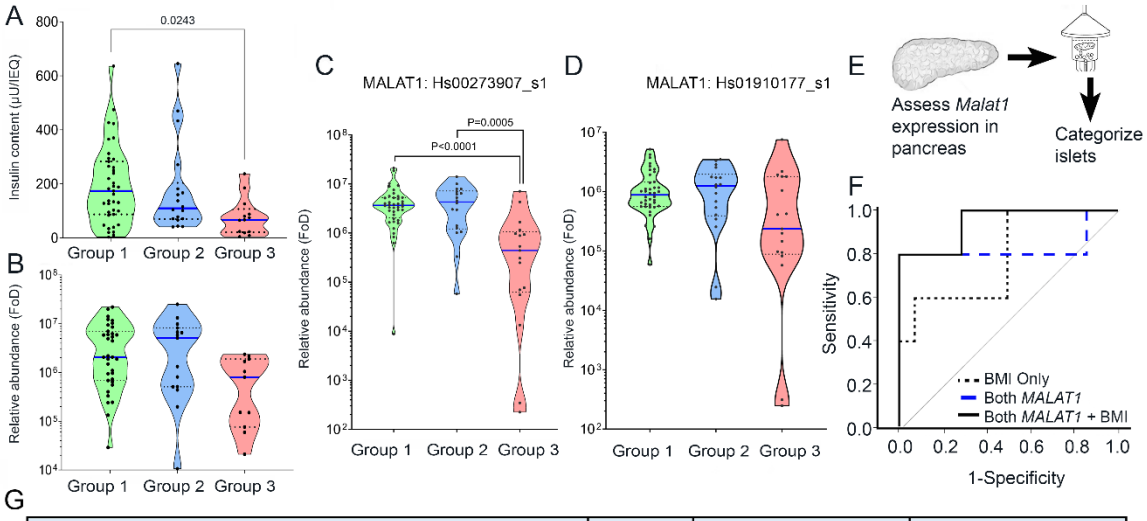


Figure 2: Penalized regression analysis and bootstrapping was performed on the (6,983) annotated lncRNAs and the identified lncRNAs across all group-wise comparisons (n=18 human islet samples, categorised into three groups) are presented in **A-C; A**) Group 1 vs Group 2; **B**) Group 1 vs Group 3 and **C**) Group 2 vs Group 3. (see **Supplementary Table 4** for details; unequal variance Students t test was used. Note: The lncRNA

AC010987.5 was present at very low FPKMs and hence not shown in Panel C). A * represents P<0.05 and ** represents P<0.01. Each red dot presents individual islet preparation, the horizontal line represents mean while the polygons represent the estimated density of data (scatter plot). Two-tailed distribution, with two-sample unequal variance Students *t* test was used to identify the difference for each lncRNA between each group-wise comparison. ROC curve: **D)** Group 1 vs Group 2: *MALAT1*-1.1, *MALAT1*-9.1 and both *MALAT1*variants (9.1+1.1); **E)** Group 1 vs 3: *MALAT1*-1.1, *MALAT1*-2.1, *MALAT1*-9.1 and all of the above three *MALAT1*(1.1+2.1+9.1); (**F**) Group 2 vs 3: GAS5 (ENST00000450589.5), *MALAT1*-2.1 and both; GAS5+*MALAT1*-2.1.



	AUC	Specificity	Sensitivity
Edmonton Donor Points	0.41	0.79	0.40
Donor BMI kg/m2	0.79	0.93	0.60
Both MALAT1 + Edmonton Donor Points	0.83	1.00	0.80
Both MALAT1	0.83	1.00	0.80
NAIDS	0.84	0.86	0.80
Both MALAT1 + NAIDS	0.93	1.00	0.80
Edmonton Donor Points + NAIDS + BMI	0.93	0.79	1.00
Both MALAT1 + BMI	0.94	1.00	0.80
Both MALAT1 + BMI + NAIDS	0.96	1.00	0.80
Both MALAT1 + BMI + Edmonton Donor Points	1.00	1.00	1.00
Both MALAT1 + BMI + Edmonton Donor Points + NAIDS	1.00	1.00	1.00

Figure 3: The three categorized human islet sample groups in the validation set (*n*=75) were assessed for **A**) insulin content levels **B**) (pro-)insulin transcript levels, **C,D**) *MALAT1* lncRNA expression (using qPCR primer/probe assay Hs_00273907 and Hs_01910177 respectively) targeting the two different variants of *MALAT1* lncRNA. To test the difference between the three categorised groups, non-parametric one-way ANOVA was used and significant P-values, adjusted for multiple comparisons are reported. Panel A-D: Each dot in the violin plot represents a different islet sample preparation. Horizontal solid blue line represents median for each group, the horizontal dotted line represents quartiles, while the polygons represent the density of individual data points and extend to min/max values. **E**) *MALAT1* lncRNA qPCR was carried prior to islet isolation and **F,G**) ROC curve for *MALAT1* lncRNA and/or donor characteristics (BMI, Edmonton Donor Points or NAIDS) of pancreas tissue samples (*n* = 19) to stratify post-isolation islet quality (Group 1 vs Group 3).

Table

Criteria assessed			
Post-isolation islet purity (%)	**Post-isolation islet viability (%)**	**Post-isolation islet β-cell viability index**	**Transplantability: Transplantable (Yes/1), Not transplantable (No/3)**
Score of 1 = 81-100	Score of 1 = 95-100%	Score of 1 = 0.5-1.0	Score of 1 = YES
Score of 2 = 61-80	Score of 2 = 85-94%	Score of 2 = 0.2-0.49	No Score
Score of 3 = ≤60	Score of 3 = ≤84	Score of 3 = <0.2	Score of 3 = NO

Table 1: Categorization of islet sample preparations into three groups was carried out by the surgical/islet isolation team based on the donor characteristics, procurement characteristics and isolation characteristics. The isolation characteristics represent islet quality, which is derived through measurement of post-isolation islet viability, purity, β-cell viability and transplantability. The transplantability is based on donor characteristics (the Edmonton donor points) and the availability of a sufficient number (4,000 IEQs/kg of recipient body weight) of isolated islet yield). The categorization of islets into three groups is a novel aspect of this paper and the isolation team adapted the following procedure for categorization. A measurement range of the four isolation characteristics were used to assign a criterion score (1, 2, 3) to each islet preparation and this is presented in the respective columns. For islet isolation outcome, islets were either scored as 1 (Transplantable) or 3 (not transplantable). Criteria scores were used to calculate a "weighted islet quality score" using the following formula: Weighted islet quality score = A / C, wherein "A" = Average of the criteria scores, and "C" = Number of available criteria. For example, an islet sample preparation may have scores of 1 (if post-isolation islet purity is 85%), a score of 2 (if post-isolation islet viability is 89%), a score of 1 (if post-isolation islet β-cell viability index is 0.75) and a score of 1 (if that islet preparation was suitable for clinical transplantation). For this sample the value of A in the above formula would be = ((1+2+1+1)/4) = 1.25 and C = 4. The islet quality score therefore would calculate to be =1.25/4= 0.3125. A weighted score system was used as some islet preparations did not have measurements for all of the four isolation characteristics. In this case, the islet isolation team compared procurement and donor characteristics (including well-known contributing factors such as cold/warm ischaemia: see **Supplementary Table 1-3**) to assign a final category score to the islet preparation. For example, an islet sample preparation may have scores of 1 (if post-isolation islet purity is 85%), a score of 2 (if post-isolation islet viability is 89%), no data for post-isolation islet β-cell viability index and a score of 3 (if that islet preparation was unsuitable for clinical transplantation). For this sample the value of A in the above formula would be = ((1+2+3)/3) = 2 and C =

3. The islet quality score therefore would calculate to be =2/3= 0.667. These islet scores were then rounded off to second decimal and classified into respective groups as per pre-defined cut-off values: Group 1 islet samples had an islet quality score $\leq$0.55, Group 2 islet samples had an islet quality score >0.55 but $\leq$0.85 whilst Group 3 islet samples had an islet quality score >0.85. All of the scoring and categorization of islets into Group 1, Group 2 or Group 3 was carried independent of the measurements of lncRNA in these samples. The molecular biomarkers team was blinded to this information.