

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (<i>n</i>) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided <i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>
☐	☒ A description of all covariates tested
☒	☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐ For null hypothesis testing, the test statistic (e.g. <i>F</i> , <i>t</i> , <i>r</i>) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted <i>Give P values as exact values whenever suitable.</i>
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's <i>d</i> , Pearson's <i>r</i>), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection	Sequencing libraries were evaluated for quality using the Agilent TapeStation and quantified with a Qubit 2.0 Fluorometer. Pooled libraries were further quantified using qPCR before being loaded onto an Illumina NovaSeq 6000 S2 flow cell. Sequencing was performed in line with 10X Genomics' guidelines. Raw sequence data (.bcl files) were converted to fastq files and demultiplexed using CellRanger mkfastq. Subsequent UMI and cell barcode deconvolution, as well as genome mapping, were performed using the CellRanger software pipeline (version 3.0.0), resulting in digital gene expression matrices and Cloupe files. This methodology ensured the isolation of viable single cells from pancreatic slice cultures, suitable for high-quality scRNA-seq analysis. The approach effectively balanced cell viability and yield across the three experimental groups, facilitating subsequent genomic analyses. Integration and dimensionality reduction. Following quality control, the Seurat R package (v4.1.1) was used to process the data. Reciprocal PCA (RPCA) was employed for the integration of scRNA-seq datasets. The RPCA integration method projects each dataset into the other's PCA space and confines the anchors by the same mutual neighborhood constraint. Post-integration, the NormalizeData() function was used to "LogNormalize" the count data. A shared nearest neighbor (SNN) modularity optimization-based clustering algorithm and the Uniform Manifold Approximation and Projection (UMAP) algorithm were utilized to project all cells into two-dimensional coordinates for visualization.
Data analysis	To identify the cell types within clusters (UMAP), we utilized the FindAllMarkers() function in the Seurat package with default settings (min.pct = 0.25, logfc.threshold = 0.25). The most significantly expressed genes in each cluster were used for cell type identification. The PanglaoDB database and the Enrichr tool were used to assign cell identities accurately. Gene Ontology (GO) analyses were performed using Enrichr. Sub-clustering of UMAP clusters. We subsetting the Endoc and H2 clusters from the UMAP of Alloxan+THR, excluding all other clusters, and performed re-clustering using Seurat to obtain four new clusters: H2.1, H2.2, H2.3, and Endoc. Further sub-setting and re-clustering of the Endoc cluster resulted in three additional clusters: 1, 2, and 3. Two sub-groups within cluster 1 were identified manually by creating a ggplot2-based scatter plot (e.g., using DimPlot() or FeaturePlot()) and selecting cells interactively with CellSelector().

RNA Velocity Analysis. RNA velocity analysis was performed using the scVelo package. Initially, the gene expression matrix was preprocessed by filtering genes and cells based on minimum expression thresholds and normalizing the data with the `scv.pp.filter_and_normalize` function. Moments necessary for RNA velocity estimation were then computed using principal components and nearest neighbors through `scv.pp.moments`. Subsequent steps for RNA velocity computation included:

1. Velocity Computation: RNA velocities were estimated using the dynamical model implemented in scVelo (`scv.tl.velocity`).

2. Velocity Graph Construction: A velocity graph, representing potential future states of cells based on RNA velocity, was constructed (`scv.tl.velocity_graph`).

Latent time, which represents the cell's position along the inferred developmental trajectory, was computed using `scv.tl.latent_time`. This step incorporates RNA velocities to provide a temporal ordering of cells. For dimensionality reduction and visualization of the high-dimensional single-cell RNA-seq data, UMAP was employed. The UMAP coordinates were computed using the scanpy package (`sc.tl.umap`). To infer connectivity between cell clusters and visualize the overall structure of cell state transitions, the PAGA algorithm implemented in scVelo was used. PAGA computation was based on pre-identified cell clusters (Seurat clusters) using `scv.tl.paga`. The PAGA transition confidence matrix was then adjusted to reflect transitions based on latent time, ensuring that transitions followed the direction of increasing latent time and thereby maintaining biological relevance to the inferred cell trajectories.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

Data availability statement. The datasets generated throughout this study are available as an NCBI BioProject #PRJNA1147257, have been deposited in NCBI's Gene Expression Omnibus, and are accessible through GEO Series accession number GSE274591 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE274591>). All code used to generate the results can be found on GitHub at JDBLab/Pancreatic-beta-cell-regeneration-by-THR-123. The R (Seurat) and Python (scVelo and Velocyto) scripts are available at JDBLab/Pancreatic-beta-cell-regeneration-by-THR-123 to analyze data for single-cell RNA seq and RNA Velocity analysis respectively. The code/Rscript files for the analyses reported in this paper are available in a GitHub coding repository maintained by the laboratories of the corresponding authors (<https://github.com/JDBLab/Pancreatic-beta-cell-regeneration-by-THR-123>).

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender	n/a
Reporting on race, ethnicity, or other socially relevant groupings	n/a
Population characteristics	n/a
Recruitment	n/a
Ethics oversight	n/a

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size for murine experiments was determined by Power Analysis (alpha error 5%; Power = 80%, inc. 20-30% attrition)
Data exclusions	The only data excluded from the analyses were from Milliplex plasma cytokine concentrations (Fig. 1): When specific datapoints are missing (be it between different cytokines or between each experimental group), it is because they were below the curve range (OOR<). This means that the analyte concentration was too low to be accurately quantified within the assay's standard curve. According to Sigma-Aldrich's guidelines, only data points within the Lower Limit of Quantification (LLOQ)-Upper Limit of Quantification (ULOQ) range are considered valid

for analysis. We report the missing data points as out of range in the figure legend.

Replication	Consistency and reproducibility of our findings on the effect of THR-123 was ensured by the conduct of multiple sets of confirmatory in vivo experiments as described in the manuscript.
Randomization	Randomization of mice to each experimental group was made prior to the development of diabetes (i.e., prior to alloxan treatment and initiation of THR-123 injections) or, in settings where diabetes was induced before the initiation of treatment, by random group generation followed by manual adjustment to ensure similar starting glycemic averages in each group.
Blinding	Due to animals being housed in clearly labeled cages, blinding was not done for data collection. Also, animals in control (saline) groups tended to develop more severe diabetes symptoms than those treated with THR-123, requiring more supportive treatment and therefore making it impractical for a blind design.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

Antibodies and probes

All from ThermoFisher Scientific:

qPCR:

Assay ID: BMR1A : Mm00477650_m1, Catalog # 4453320,

Beta Actin: Mm02619580_g1. catalog #: 4331182

UCP1: nMm01244861_m1, catalog # 4331182

For Flow Cytometric:

Fixable Viability Dye eFluor™ 780, catalog # 65-0865-14

anti-SOX9 (Bioss, cat # bs4177R-A488) , anti-CD24-450 (Invitrogen cat # 48-0242-82) anti-BMPRI1A-A647 (Bioss cat # bs1509R-A647),

anti-P2Y1R-FITC (Biorbyt) and anti-pSmad PE (Cell Signaling 13820S)

eBioscience™ Foxp3 / Transcription Factor Fixation/Permeabilization Concentrate and Diluent, Catalog number 00-5521-00

For Immunofluorescence

Anti-INSULIN-Guinea Pig Polyclonal, Dako, Cat# A0564; RRID: AB_10013624

Anti-GLUCAGON-Rabbit Polyclonal, Dako, Cat# A0565; RRID: AB_10013726

Anti-KRT19-Rabbit Monoclonal, Abcam, Cat# ab52625; RRID: AB_2281020

Anti-Cytokeratin 7 antibody [EPR1619Y] – Rabbit monoclonal Abcam, Cat# ab68459, RRID: AB_1139824

Anti-SOMATOSTATIN-Rat Polyclonal Millipore Cat# MAB354 RRID: AB_2255365

Anti-ALPHA AMYLASE-Rabbit Polyclonal, Sigma Aldrich, Cat# A8273; RRID: AB_258380

Anti-SOX9-Rabbit Polyclonal Millipore, Cat# ab5535; RRID: AB_2239761

Anti-BrdU-Rat Polyclonal, Abcam Cat# ab6326; RRID: AB_305426

Nuclear Decloaker, 10X BioCare medical, cat # CB911M RRID: AB_2801515.

Protein Block, Serum-Free, AGILENT, Cat # X090930-2 RRID:AB_2801516.

ImmEdge Pen Vector Laboratories Cat# H-4000, RRID:AB_2336517

Normal Donkey Serum antibody, Jackson Immuno Research laboratories, Inc. Cat # 017-000-121

RRID: AB_2337258

Alexa Fluor® 594 AffiniPure Donkey Anti-Guinea Pig IgG (H+L) Jackson Immuno Research laboratories, Inc. Cat#706-585-148; RRID : AB_2340474

Alexa Fluor® 647 AffiniPure Donkey Anti-Guinea Pig IgG (H+L) Jackson Immuno Research laboratories, Inc. Cat#706-605-148; RRID: AB_2340476

Alexa Fluor Donkey polyclonal anti-rabbit IgG, Alexa Fluor 488, Thermofisher Scientific, Cat#A-21206; AB_2535792, Donkey Anti-Sheep IgG (H+L) Antibody, Alexa Fluor 488 Conjugated, Thermofisher Scientific, Cat# A-11015, RRID:AB_141362

Donkey Anti-Rabbit IgG (H+L) Polyclonal Antibody, Alexa Fluor 647 Conjugated, Thermofisher Scientific, Cat# A-31573, RRID: AB_2536183)

Donkey anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 594, Thermofisher Scientific, Cat# A-21203,

RRID: AB_2535789

Donkey anti-Goat IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 647 Thermofisher Scientific, Cat# A-21447, RRID: AB_2535864

Donkey anti-Rabbit IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 594, Thermofisher Scientific, Cat# A-21207, RRID: AB_141637

Validation

The antibodies used in this study have been validated by the companies before purchased. They are been used by the research community since 2008, with citations in over number of peer-reviewed publications. The lot number and the reproducibility for the application consistency was constantly checked.

Animals and other research organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research, and [Sex and Gender in Research](#)

Laboratory animals

We used CD1 and C57BL/6 mice as wild-type experimental mice and Ins2-Cre/mTmG as a transgenic variant for some experiments as described in the manuscript.

Wild animals

N/A

Reporting on sex

As discussed in the text, due to the widely reported challenge of consistently inducing diabetes in female mice through chemical means (Kilic et al, 2014; Deeds et al, 2011; Maxwell et al., 2022) we used male mice to test our hypothesis.

Field-collected samples

N/A

Ethics oversight

All animal experiments were conducted under the supervision and oversight of the Institutional Animal Care and Use Committee (IACUC) and Division of Veterinary Resources (DVR) at the University of Miami.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Plants

Seed stocks

N/A

Novel plant genotypes

N/A

Authentication

N/A

Flow Cytometry

Plots

Confirm that:

- ☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- ☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- ☒ All plots are contour plots with outliers or pseudocolor plots.
- ☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

Cells were harvested and filtered as described in the manuscript. Cells were stained in DPBS with the viability dye eFluor 780 (eBiosciences/ThermoFisher) for 15 minutes at 4°C in the dark, then they were washed three times in DPBS and fixed and permeabilized in a Foxp3 / Transcription Factor Fixation/Permeabilization Concentrate and Diluent kit. Subsequently, cells were blocked in blocking buffer for 1 hour. Cells were incubated with diluted primary antibodies for one hour, dilutions were according to the manufacturer's suggestion. After three consecutive washes cells were then resuspend in appropriate amounts of Flow Cytometry Staining Buffer (500µl) and analyze by flow cytometer in the LSRII or Cytotflex. Further analysis and generation of plots was completed using the Kaluza software v1.5a (BD). To compare two different time points paired with two tailed t test P<0.05 was run. To analyze the results, the Kruskal Wallis test was performed.

Instrument	Cytoflex S from Beckman Coulter, Inc.
Software	Kaluza Analysis V1.5a and V2.1.1, 2016)Beckman Coulter. https://www.beckman.com/flow-cytometry/software/kaluza
Cell population abundance	Single-cell suspensions were prepared and 1×10^5 cells and viability confirmed by stained with eFluor 780 viability dye. Viable cells constituting $70.2\% \pm 2.5\%$ of total events. Singlets were then gated using FSC-H vs FSC-A plots.
Gating strategy	To identify mNEPT cells with active BMP7 signaling, we employed two distinct gating strategies: Strategy 1: Identifying BMP-Active Cells We began by identifying viable cells using eFluor 780 to exclude the dead cell, in combination with forward scatter area (FSA). To exclude debris, we used side scatter width (SSC-W) versus side scatter height (SSC-H). The resulting population was further refined based on forward and side scatter characteristics. From this refined population, we generated a bivariate plot of P2RY1 (FITC-A) versus ALK3-A and selected the double-positive (P2RY1⁺/ALK3⁺) subpopulation. Within this group, we gated on pSMAD1/5/8 (ECD-A) positive events to identify cells with active BMP signaling. To identify the ALK3⁻/pSMAD⁻ population, we returned to the P2RY1 vs. ALK3 plot, gated on ALK3⁻ cells, and then selected the pSMAD1/5/8-negative events within this subset. Strategy 2: Identifying Ductal Cells with or without BMP Signaling Using the same initial gating strategy for viable and single cells, we plotted FSC-A versus CD24 (Pacific Blue-A) to identify the ductal population (CD24⁺). From this population, we gated on double-positive events for CD24 and pSMAD1/5/8 (PE-A). Within this fraction, we further identified triple-positive cells expressing ALK3 (APC-A), HNF1β, and SOX9 (FITC-A). The same approach was applied to the pSMAD⁻ population to identify ALK3⁺/HNF1β⁺/SOX9⁺ cells lacking active BMP signaling.

☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.