Novel biomarkers for glycaemic deterioration in type 2 diabetes: an IMI RHAPSODY study

Roderick C Slieker, PhD[1,2,*], Louise A Donnelly, PhD[3,*], Livia Lopez-Noriega, PhD[4], Hermine Muniangi-Muhitu[4], Elina Akalestou, PhD[4], Mahsa Sheikh[4], Eleni Georgiadou, PhD[4], Giuseppe N. Giordano, PhD[5], Mikael Åkerlund, PhD[5], Emma Ahlqvist, PhD[5], Ashfaq Ali, PhD[6], Marko Barovic[7], MSc, Gerard A Bouland, MSc[2], Frédéric Burdet, MSc[8], Mickaël Canouil, PhD[9], Iulian Dragan, MSc[8], Petra JM Elders, PhD[10], Celine Fernandez, PhD[5], Andreas Festa, MD[11,12], Hugo Fitipaldi, MSc[5], Phillippe Froguel[9,13], Valborg Gudmundsdottir, PhD[14,15], Vilmundur Gudnason, MD, PhD[14,15], Mathias J. Gerl, PhD[16], Amber A van der Heijden, PhD[10], Lori L Jennings, PhD[17], Michael K. Hansen, PhD[18], Min Kim, PhD[6,19], Isabelle Leclerc MD PhD[4],Christian Klose, PhD[16], Dmitry Kuznetsov, PhD[8], Dina Mansour Aly, MSc[5], Florence Mehl, PhD[8], Diana Marek, PhD[8], Prof Olle Melander, PhD[5], Anne Niknejad, MSc[8], Filip Ottosson, MSc[5], Imre Pavo, MD[11], Alexander Efanov, PhD[20], Kevin Duffin[20], Timothy J. Pullen, PhD[4,21], Prof Kai Simons, PhD[16], Prof. Michele Solimena, PhD[7,22], Tommi Suvitaival, PhD[6], Asger Wretlind, MSc[6], Prof Peter Rossing, MD[6,23], Prof Valeriya Lyssenko, PhD[24,25], Cristina Legido Quigley, PhD[6,19], Prof Leif Groop, PhD[5,26], Prof Bernard Thorens, PhD[27], Prof Paul W Franks, PhD[5,28], Mark Ibberson, PhD[7], Prof Joline WJ Beulens, PhD[1,29], Leen M 't Hart, PhD[1,2,30,*, †], Prof Ewan R Pearson, MD[3,*, ††], Prof Guy A Rutter, PhD[*,4,31, ††]

[1] Department of Epidemiology and Data Science, Amsterdam Public Health Institute, Amsterdam Cardiovascular Sciences, Amsterdam UMC, location VUMC, Amsterdam, the Netherlands

[2] Department of Cell and Chemical Biology, Leiden University Medical Center, Leiden, the Netherlands

[3] Population Health & Genomics, School of Medicine, University of Dundee, Dundee, UK.

[4] Section of Cell Biology and Functional Genomics, Division of Diabetes, Endocrinology and Metabolism, Department of Metabolism, Digestion and Reproduction, Imperial College London, London, U.K.

[5] Genetic and Molecular Epidemiology Unit, Lund University Diabetes Centre, Department of Clinical Sciences, CRC, Lund University, SUS, Malmö, Sweden

[6] Steno Diabetes Center Copenhagen, Gentofte, Denmark

[7] Paul Langerhans Institute Dresden (PLID) of the Helmholtz Center Munich at the University Hospital Carl Gustav Carus and Medical Faculty, Dresden, Germany.

[8] Vital-IT Group, SIB Swiss Institute of Bioinformatics, Lausanne, Switzerland

[9] INSERM U1283, CNRS UMR 8199, European Genomic Institute for Diabetes (EGID), Institut Pasteur de Lille, University of Lille, Lille University Hospital, Lille, F-59000, France

[10] Department of General Practice and Elderly Care Medicine, Amsterdam Public Health Research Institute, Amsterdam UMC–location VUmc, Amsterdam, the Netherlands

[11] Eli Lilly Regional Operations GmbH, Vienna, Austria.

[12] 1st Medical Department, LK Stockerau, Niederösterreich, Austria.

[13] Division of Systems Biology, Department of Diabetes, Endocrinology and Metabolism, Imperial College London, London, UK

[14] Faculty of Medicine, University of Iceland, Reykjavik, Iceland

[15] Icelandic Heart Association, Kopavogur, Iceland

[16] Lipotype GmbH Dresden Germany

[17] Novartis Institutes for Biomedical Research, Cambridge, MA 02139, USA

[18] Cardiovascular and Metabolic Disease Research, Janssen Research & Development, Spring House, Pennsylvania, U.S.A.

[19] Institute of Pharmaceutical Science, Faculty of Life Sciences and Medicines, King's College London, London, United Kingdom

[20] Lilly Research Laboratories, Eli Lilly and Company, Indianapolis, United States

[21] Department of Diabetes, Guy's Campus King's College London, U.K.

[22] Molecular Diabetology, University Hospital and Medical Faculty Carl Gustav Carus, TU Dresden, Dresden, Germany.

[23] Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark

[24] Department of Clinical Science, Center for Diabetes Research, University of Bergen, Bergen, Norway

[25] Genomics, Diabetes and Endocrinology Unit, Department of Clinical Sciences Malmö, Lund University Diabetes Centre, Skåne University Hospital, Malmö, Sweden

[26] Finnish Institute of Molecular Medicine, Helsinki University, Helsinki, Finland

[27] Center for Integrative Genomics, University of Lausanne, CH-1015, Lausanne, Switzerland

[28] Department of Nutrition, Harvard School of Public Health, Boston, MA, USA

[29] Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, the Netherlands

[30] Department of Biomedical Data Sciences, Section Molecular Epidemiology, Leiden University Medical Center, Leiden, the Netherlands

[31] Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore

* Contributed equally

Corresponding authors

[†] Leen 't Hart, PhD
Einthovenweg 20
2333ZC Leiden
The Netherlands
Telephone: +31 71 5269796
E-mail: lmthart@lumc.nl

[††] Prof Ewan Pearson, MD
Ninewells Hospital
Dundee
DD1 9SY
United Kingdom
Telephone: +44 1382 383387
E-mail: E.Z.Pearson@dundee.ac.uk

[†††] Prof Guy A Rutter, PhD
Section of Cell Biology and Functional Genomics,
Division of Diabetes, Endocrinology and Metabolism, Department of Metabolism, Digestion and Reproduction,
Imperial College London,
Hammersmith Hospital Campus
Du Cane Road, London, W12 0NN, U.K

ABSTRACT

We have deployed a multi-omics approach in large cohorts of patients with existing type 2 diabetes to identify biomarkers for disease progression across three molecular classes, metabolites, lipids and proteins. A Cox regression analysis for association with time to insulin requirement in 2,973 patients in the DCS, ANDIS and GoDARTS cohorts identified homocitrulline, isoleucine and 2-aminoadipic acid, as well as the bile acids glycocholic and taurocholic acids, as predictive of more rapid deterioration. Increased levels of eight triacylglycerol species, and lowered levels of the sphingomyelin SM 42:2;2 were also predictive of disease progression. Of ~1,300 proteins examined in two cohorts, levels of GDF-15/MIC1, IL-18RA, CRELD1, NogoR, FAS, and ENPP7 were associated with faster progression, whilst SMAC/DIABLO, COTL1, SPOCK1 and HEMK2 predicted lower progression rates. Strikingly, identified proteins and lipids were also associated with diabetes incidence and prevalence in external replication cohorts. Implicating roles in disease compensation, NogoR/RTN4R improved glucose tolerance in high fat-fed mice and tended to improved insulin signalling in liver cells whilst IL-18R antagonised inflammatory IL-18 signalling towards nuclear factor kappa-B *in vitro*. Conversely, high NogoR levels led to islet cell apoptosis. This comprehensive, multi-disciplinary approach thus identifies novel biomarkers with potential prognostic utility, provides evidence for new disease mechanisms, and identifies potential therapeutic avenues to slow diabetes progression.

INTRODUCTION

Type 2 diabetes is a progressive multifactorial disease which presently affects >400m worldwide, with numbers expected to increase to > 700 m by 2045.[1] Biomarkers for the disease, which provide a deeper understanding of the disease process, are therefore eagerly sought. Importantly, their identification may improve prediction and personalized approaches to disease treatment.[2]

Whilst many studies have examined associations between circulating biomarkers and incident disease[3,4], to date few studies have explored changes associated with glycaemic deterioration after the development of diabetes. Published studies (reviewed in [5]) have established that faster glycaemic deterioration is seen in those who are diagnosed younger, are more obese at diagnosis, have lower HDL, and higher HbA1c. A few studies have investigated genetic variants associated with more rapid progression with small and variable results[5], although a report on a Hong Kong Chinese population reported a replicated finding that a high polygenic risk score consisting of 123 T2D risk variants was associated with increased progression to insulin requirement.[6] To date, no studies that have adopted a multi-omic approach to biomarker discovery, or reported systematically how metabolites of different classes impact on progression. Such associations have the potential to be clinically useful in terms of prediction, as well as providing biological insights into the processes that drive glycaemic deterioration in T2D.

In a collaboration based around the EU Innovative Medicines Initiative-2 Risk Assessment and ProgreSsiOn of Diabetes (RHAPSODY) we have undertaken here to identify, in three large European cohorts, biomarkers of diabetes progression of three molecular classes: charged small molecules (metabolites), lipids and proteins. In this way, we identify novel species and, in the case of two of the identified proteins, provide evidence through functional studies in preclinical models for novel mechanisms of action in disease-relevant tissues.

METHODS

Discovery cohorts

Specific details on DCS[7], GoDARTS[8] and ANDIS[9] have been described elsewhere.[10] These cohorts were selected based in part on satisfactory quality control for biomarkers stability in stored samples.

Briefly, the Hoorn Diabetes Care System (DCS) cohort is a prospective cohort with currently over 14,000 individuals with routine care data. The Ethical Review Committee of the VU University Medical Center, Amsterdam approved the study. In 2008-2014, additional blood sampling was done in 5,500 participants, who provided written informed consent. These samples were used for this study. The turbidimetric inhibition immunoassay for haemolyzed whole EDTA blood (Cobas c501, Roche Diagnostics, Mannheim, Germany) was used to measure HbA1c. HDL (mmol/L) was measured enzymatically (Cobas c501, Roche Diagnostics). C-peptide was measured on a DiaSorin Liaison (DiaSorin, Saluggia, Italy).

The Genetics of Diabetes Audit and Research Tayside Study (GoDARTS) is a cohort of ~8,000 patients with T2D. The study was approved by the Tayside Medical Ethics Committee and all individuals provided informed consent. Laboratory measurements were measured in a non-fasted state. C-peptide was measured on a DiaSorin Liaison (DiaSorin, Saluggia, Italy).

In the All New Diabetics in Scania (ANDIS) cohort, people with incident diabetes within Scania County, Sweden were recruited from January 2008 until November 2016 and all participants gave written informed consent. Regional ethics review committee in Lund approved the study. An electro-chemiluminescence immunoassay was used to measure C-peptide on a Cobas e411 (Roche Diagnostics, Mannheim, Germany) or a radioimmunoassay

(Human C-peptide RIA; Linco, St Charles, MO, USA; or Peninsula Laboratories, Belmont, CA, USA). The Clinical Chemistry database was used to obtain HbA1c levels.

Validation cohorts

External validation was performed in four external cohorts, ACCELERATE, AGES-Reykjavik, MDC-CC and DESIR. ACCELERATE is a clinical trial aimed at investigating the effect of evacetrapib on major adverse cardiovascular outcomes and has been described elsewhere.[11] For the current study we only included the 6,054 individuals in the untreated arm. From this group, we selected 2,978 individuals with type 2 diabetes. In this group, 1003 individuals were excluded that did not have C-peptide levels or HbA1c levels, 72 were excluded because the age at diagnosis was < 35 years, 31 were excluded because they were on insulin at baseline and 22 were excluded because they had > 2 non-insulin glucose-lowering drugs and HbA1c levels > 8.5%. The final set consisted of 1,850 individuals of which 162 reached the primary endpoint.

AGES-Reykjavik is a prospective population-based study from Iceland.[12,13] In fasted blood samples protein levels were measured with the Somalogic platform. At baseline there were 4784 individuals free of diabetes and 654 with type 2 diabetes. Of 2,940 individuals free of diabetes at baseline and with 5-year follow-up information, 112 developed type 2 diabetes. [13] Identified proteins were tested against incident and prevalent type 2 diabetes using logistic regression, adjusted for age and sex.

Malmö Diet and Cancer Cardiovascular Cohort (MDC-CC) is a population-based cohort comprised of people living Malmö.[14] Lipids were measured using the Lipotype platform in 3,667 individuals of which 555 developed type 2 diabetes.[15] Proteins were measured using the Olink Proseek Multiplex proximity extension assay in 4915 individuals of

188 which 700 developed type 2 diabetes. Identified lipids and proteins were tested against incident
189 diabetes using Cox proportional hazard model adjusted for age, sex and BMI.
190 DESIR is a prospective population-based cohort comprised of middle-aged European
191 individuals. Metabolomics was measured by Metabolon (Durham, NC).[16] Logistic regression
192 adjusted for age, sex and BMI was used to tested for an association between metabolites and
193 prevalent (n = 43) and incident (n = 231) type 2 diabetes versus controls (n = 813).
194
195 *Molecular measurements*
196 In the three discovery cohorts, those were selected with an age at diagnosis > 35 years, GAD
197 negative, GWAS data and with a blood sample within three years after diagnosis. The
198 metabolomics, lipidomics and proteomics groups were of different sizes (see further details
199 below). Individuals were ranked based on the time between diagnosis and sampling date and
200 those with the smallest time between diagnosis and sampling were selected. For metabolomics,
201 we selected 1,267 in DCS, 900 in GoDARTS and 900 in ANDIS. For lipidomics, 900
202 individuals were selected in DCS, GoDARTS and ANDIS. For proteomics, we selected 600
203 individuals in DCS and GoDARTS.
204
205 *Small charged molecule analytes*
206 In 2,973 individuals 19 small charged molecule analytes (referred to as metabolomics) using
207 UHLPC-MS/MS (UHLPLC: 1290 Infinity system from Agilent Technologies, Santa Clara,
208 CA, USA; MS/MS: 6460 triple quadrupole system from Agilent Technologies) relevant to
209 diabetes were quantified, which was used as the discovery set.[17] In DCS, all samples passed
210 QC and were used in the analysis. In GoDARTS, three failed in QC and the remaining samples
211 were used for analysis. In ANDIS, 4 failed QC and of the 892 remaining samples, 811 were

212 free of the outcome at sampling. In addition, a validation set was generated comprised of 2,668
213 individuals (699 GoDARTS, 1,969 ANDIS).
214
215 *Lipid measurements*
216 Six hundred and fourteen (614) lipids were determined using a QExactive mass spectrometer
217 (Thermo Scientific) equipped with a TriVersa NanoMate ion source (Advion Biosciences) on
218 the Lipotype lipidomics platform (Lipotype, Dresden, Germany).[18] Samples (n = 2,608) were
219 measured in batches of 84 samples each. Unprocessed mass spectra files were used to identify
220 lipids with LipotypeXplorer.[19] Lipid identifications with a signal-to-noise ratio > 5, and a signal
221 intensity 5-fold higher than in corresponding blank samples were considered for further data
222 analysis. Eight reference samples were used to apply batch correction and amounts were further
223 adjusted for analytical drift (p-value slope ≤ 0.05 and $R^2 \geq 0.75$ and the relative drift > 5%).
224 Lipid nomenclature is used as described previously and SwissLipids database identifiers are
225 provided (**Table S1**).[20] After quality control 162 lipid species were used in this study. The
226 median coefficient of subspecies variation of the 162 lipids used as accessed by reference
227 samples was 9.49% across all three cohorts. In DCS, 900 individuals were included for
228 lipidomics measurements, all passed QC, and all were suitable for analysis. In GoDARTS, 898
229 individuals were included in the analysis, 1 failed QC and all 897 remaining samples were
230 included in the analysis. In ANDIS, 896 individuals were included in the analysis, 5 failed QC
231 and of the 891 remaining samples, 811 were free of the outcome at sampling.
232
233 *Protein measurements*
234 Proteins were measured on the SomaScan® Platform from Somalogic (n = 1195 proteins) on
235 the SomaLogic SOMAscan platform (Boulder, Colorado, USA) in 1,188 individuals. Top
236 associated proteins were validated in ANDIS (n = 1992) and ACCELERATE (n = 1850) using

ELISA with time to insulin requirement as the outcome. External validation was performed for the top proteins based on P-value and/or effect size for which an ELISA was available or could be developed.

Primary endpoint

The primary endpoint time to insulin requirement was defined as the period from diagnosis to a clinical endpoint of the earlier of (i) starting sustained (more than 6 months duration) insulin treatment or (ii) clinical requirement of insulin as indicated by two or more HbA1c measurements>8.5% more than three months apart when on two or more non-insulin diabetes therapies.[21] In DCS, 600 individuals were included for proteomics measurements, 11 failed QC and all were included for analysis. In GoDARTS, 600 individuals were included in the analysis, 1 failed QC and the 599 remaining samples were included in the analysis.

Federated database

All main analyses were performed on a federated database system. Opal, an open-source data warehouse (Open Source Software for BioBanks, OBiBa) was used to store cohort data on local nodes and remote analysis was performed in R using *DataSHIELD*[22] and *dsSwissKnife* R packages.[23] A central server was set up at the Swiss Institute of Bioinformatics to manage federated node access, user administrator and software deployment. Local nodes were set up at the respective cohorts. All data was harmonized according to the CDISC Study Data Tabulation Model (www.cdisc.org) prior to inclusion into the federated database.

Mendelian randomisation

Genetic instruments for proteins, lipids and metabolites predictive of progression in the three cohorts were obtained from published GWASs. Protein quantitative trait loci (QTLs) were

obtained from Gudmundsdottir et al.[13] Lipid QTLs were obtained from Tabassum et al.[24] Metabolite QTLs were obtained from Lotta et al.[25] Only QTLs with P-values < $5 \cdot 10^{-8}$ were included. For traits with one instrument Wald ratio was used and for multiple instruments inverse variance weighting. Instruments were excluded when in LD ($r^2 > 0.1$). Genetic instruments for type 2 diabetes were obtained from the latest GWAS on incident type 2 diabetes.[26] Horizontal pleiotropy was estimated based on MR-Egger intercept. Cochran Q-statistic was used to estimate heterogeneity of instruments.

Cells and cell culture

HEK-Blue IL-18 cells passages 1-16 (InvivoGen, USA) were cultured in 4.5 g/L glucose Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% foetal bovine serum (FBS), 2 mM L-Glutamine, 50 U/ml penicillin, 50 µg/ml streptomycin (Sigma, UK) and 100 µg/ml Normocin (InvivoGen, USA). HEK-Blue IL-18 cells were designed to detect bioactive IL-18 by monitoring the activation of the NF-κB and AP-1 pathways. They were generated by stable transfection of HEK293 derived cells with the genes encoding IL-18R and IL-18 receptor accessory protein (IL-18RAP). Additionally, the TNF-α and the IL-1β responses have been blocked to guarantee a specific respond to IL-18. Cells were seeded in T75 flasks and sustained at 37°C in a humidified incubator containing 5% CO2. Experiments were carried out with HEK293 cells as well as a negative control.

Cell transfection

HEK-Blue IL-18 and HEK293 cells were seeded in 12-well plates at 30,000 per well and transfected after 48 hours with both NF-κB and Renilla using Lipofectamine 2000 DNA Transfection Reagent (Invitrogen, USA). For each well, 1.5 µl of Lipofectamine reagent was diluted in 100µl Opti-MEM Gibco medium (Thermo Fisher, USA). The amount of 1 µl of NF-

287 κB DNA at a concentration of 472.6 ng/µl and 1 µl of Renilla at a concentration of 40 ng/µl
288 were also diluted in 100 µl of Opti-MEM medium for each well, the diluted DNA mix was then
289 added to diluted Lipofectamine reagent and the transfection mix was incubated for 20 minutes
290 at RT. Cells were washed with 1ml of PBS and 400 µl of Opti-MEM medium was added per
291 well. Cells were transfected with 200 µl transfection mix each well. Medium was then changed
292 to complete growth medium (DMEM supplemented with 10% FBS, 2 mM L-Glutamine, 50
293 U/ml penicillin, 50 µg/ml streptomycin and 100 µg/ml Normocin) after 4 hours.
294
295 *Cell treatment and stimulation*
296 Transfected cells were stimulated with different concentrations of IL-18, IL-18Rα or both to
297 measure changes in NF-κB activation. Concentrations were made by diluting different volumes
298 of 1 µM IL-18 and IL-18Rα stock (diluted from freeze-dried powder in in distilled water) in
299 serum-free medium to exclude the effect of serum factors and cells were stimulated with 300
300 µl of each condition per well of 12-well plate for 6 hours. Recombinant human IL-18 and rhIL-
301 18 Rα/Fc chimera were obtained from MBL and R&D Systems respectively. Normal rabbit
302 IgG from Abcam (ab171870) was used as a negative control.
303
304 *Dual luciferase reporter assay*
305 Cells transfected with NF-κB and stimulated with IL-18/IL-18Rα were lysed following six
306 hours of treatment by adding 200 µl of Promega Passive Lysis Buffer (PLB) to each well and
307 gentle shaking for 15 minutes at RT. Cell lysates were stored at -20°C and luciferase activity
308 was measured the next day using Promega Dual-Luciferase Reporter Assay System according
309 to manufacturer's instructions to study gene expression at the transcriptional level. Briefly,
310 Luciferase assay reagent II (LAR II) was dispensed into a luminometer tube for each condition.
311 Cell lysate was resuspended in tube containing LAR II and Firefly luciferase activity was

measured using Berthold Lumat LB 9507 Tube Luminometer. Next, Stop & Glo reagent was added to each tube and the Renilla luciferase activities were measured. Measurements were read with PuTTY software and all Firefly luciferase activities were normalized with Renilla luciferase activities to obtain the NF-κB signalling ratios using Microsoft Excel software.

HepG2 cells stimulation and Western (immuno-) blotting

HepG2 Cells (human hepatocytes) were cultivated in DMEM media with 1g/L of glucose, supplemented with 10% SVF. HepG2 cells line were maintained at 37 °C and 5% CO2. Cells were treated with 0, 1nM, 10 nM, or 100 nM of NogoR (see above) or CRELD1 (Mouse Fc-tagged, Sino Biological, Cat # 51149-M02H) recombinant proteins for 3 hours and then stimulated with insulin for 15 minutes prior protein extraction. Cells were lysed in radioimmunoprecipitation assay buffer (RIPA buffer: 20 mM Tris-HCl, 50 mM NaCl, 1 mM Na_2-EDTA, 1 mM EGTA, 1% NP-40, 1% sodium deoxycholate) with 1% Phosphatase Inhibitor Cocktails (P0044 and P5725, Sigma-Aldrich) and 1% Protease Inhibitor Cocktail (P8340, Sigma-Aldrich). Protein concentration was measured using a Pierce BCA Protein Assay Kit (Thermofisher, 23225) at 562 nm using a PHERAstar reader (BMG Labtech). After protein transfer to a polyvinylidene Difluoride (PVDF) membrane (Millipore), membranes were blocked with Tris-buffered saline 1X plus 0.1% (v/v/) Tween (TBST), containing 4% (w/v) BSA for 1h. Primary antibody incubation was performed overnight during rotation at 4 ºC in TBST-4% (w/v) BSA and horseradish peroxidase (HRP)-conjugated secondary antibody incubation was subsequently performed for 1 h in TBST-4% (w/v) milk powder. For development we used Clarity Western ECL Substrate (Bio-Rad).

In vivo metabolic tests

All *in vivo* procedures were approved by the UK Home Office, according to the Animals (Scientific Procedures) Act 1986 with local ethical committee (Imperial AWERB) approval under personal project license (PPL) number PA03F7F07 to I.L. For the oral glucose tolerance test (OGTT), C56BL6J mice (Charles River) maintained for four weeks on a high fat high sucrose (HFHS) diet (a 58 kcal% Fat and Sucrose diet (D12331, Research Diet, New Brunswick, NJ) were fasted for 16 h prior to experiment and received an oral glucose load (2 g/kg of body weight). For insulin tolerance tests, mice received an intraperitoneal injection of insulin (1 IU/kg) after 3 h fasting. Blood glucose levels were determined by tail venepuncture using a glucose meter (Accu-Chek; Roche, Burgess Hill, UK) at 0, 15, 30, 60 and 120 min. after the glucose load. For insulin measurements blood was collected in EDTA covered tubes at times 0, 15 and 30 min. after the glucose load (2 g/kg of body weight). Subsequently, blood was centrifuged at 4000 x g for 20 min. at 4 ºC and plasma was collected. Insulin was determined by ELISA (CrystalChem, 90080), according to the manufacturer's instructions. Daily intraperitoneal injections of NogoR were performed with indicated dose of NogoR (mouse, His and Fc tag; Sino Biologicals Cat # 50106-M03H).

Insulin secretion from mouse and human islets

Mouse pancreatic islets were isolated from male C57BL/6 mice (Envigo, Indianapolis, IN) by collagenase digestion. Use of animals was approved by Eli Lilly and Company's Institutional Animal Care and Use Committee. Human pancreatic islets from listed cadaver organ donors that were refused for pancreas or islet transplantation were obtained from Prodo Labs (Irvine, CA) and InSphero AG (Schlieren, Switzerland) and were used in accordance with internal review board ethical guidelines for use of human tissue. Islets were cultured in the complete PIM(S) Prodo Islet Media (Prodo Labs) and RPMI 1640 medium (Invitrogen) supplemented

with 11 mm glucose, 10% (v/v) heat-inactivated foetal bovine serum (Invitrogen), 100 IU/ml penicillin, and 100 µg/ml streptomycin (Invitrogen).

For insulin secretion in mouse islets, islets were incubated for 30 min. in Earle's balanced salt solution (EBSS) buffer supplemented with 3 mM glucose and 0.1% BSA. Then groups of three islets were selected and cultured with tested proteins at indicated glucose concentration in 300 µL of EBSS for 60 minutes at 37°C. At the end of incubation, the supernatant was collected and subjected to insulin analysis. To measure insulin secretion in human pancreatic islets, single islets placed in a GravityTRAP 96-well plate (InSphero) were washed and incubated for 30 minutes in 100 µL of EBSS supplemented with 0.1% BSA and 3 mM glucose. Then the buffer was replaced with 100 µL of EBSS containing indicated glucose and protein concentrations and further cultured for 60 minutes at 37°C. At the end of incubation, the supernatant was collected and submitted for insulin analysis. Insulin levels were determined with the Meso Scale Discovery (Gaithersburg, MD) electrochemiluminescence insulin assay.

For chronic incubation experiments, after overnight recovery human or mouse islets were cultured in 12 well plate (20-30 islets per well) in the RPMI-1640 culture media containing tested proteins for 72 h. At the end of incubation, islets were transferred into EBSS supplemented with 3mM glucose and 0.1% BSA. Then, 1h insulin secretion in response to elevated glucose in islets pre-treated with proteins was measured.

Quantification of β-cell proliferation

After overnight recovery, mouse or human islets were cultured for 72h in 12-well plates (200-300 islets per well) in RPMI-1640 medium containing 5mM glucose, 2% FBS, 10 µM EdU and tested proteins. At the end of incubation, islets were washed, dispersed into single cells with the Accutase solution (Sigma) and placed in 96 well plate coated the Cell-Tak Cell and

Tissue Adhesive (Corning). Cells were fixed, permeabilized and stained with Click-iT EdU HCS assay (ThermoFisher), PDX1 antibody (ab47308, Abcam) and Hoechst 33342 (ThermoFisher) nucleic acid dye. Cell images were captured and analysed using InSight Imaging System (ThermoFisher).

Islet cell apoptosis

After 3-4 day culture, mouse or human islets were plated (1 islet per well) in the GravityTRAP 96-well plate (InSphero) in 100 µl/well RPMI-1640 medium containing 11 mM glucose and 1% FBS. To induce cell death, islets were treated with glucose (25 mM) and palmitic acid (300 µM palmitate conjugated with BSA), cytokine mixture (120 ng/ml TNF-α, 60 ng/ml IL-1B and 240 ng/ml IFN-γ, R&D Systems) or tunicamycin (0.03 µg/ml, Tocris). After 72-hour incubation, islets were caspase activity was measured with the Caspase-Glo 3/7 Assay (Promega) according to manufacturer's protocol.

Statistical analyses

A Cox proportional hazard model was used to identify molecular risk factors for time to insulin requirement in R (v3.6.0) remotely on each cohort federated node using the *dssCoxph* function in *dsSwissKnife*. Three models with and without biomarkers were explored, where a primary model was adjusted for age, sex and BMI. A second model was further adjusted for HDL and C-peptide and the fully adjusted model additionally for diabetes duration and the use of glucose lowering drugs. Models were stratified for HbA1c (strata: < 53 mmol/mol, 53-64 mmol/mol and > 64 mmol/mol). Figures and meta-analysis was performed locally with R (v4.0.3). Estimates in each cohort were meta-analysed using the *metagen* function from the R package meta. P-values were adjusted for multiple testing using the Benjamini-Hochberg procedure and a $P_{FDR} < 0.05$ was considered significant. Figures were made using ggplot2 (v3.3.2). Analysis

412 of cellular and metabolic data were performed using GraphPad Prism versions 7.0-9.0 (San
413 Diego, CA, U.S.A.) deploying ANOVA with appropriate post hoc correction for multiple
414 testing.

RESULTS

Cohort characteristics and modelling of glycaemic deterioration

Individuals from three cohorts, DCS, GoDARTS and ANDIS were included. In a subset molecular characterization was performed of which characteristics are shown in **Table S2**. The characteristics across the cohorts were comparable (**Table S2**). Male subjects were more abundant in the cohorts (>55%), and the average age ranged from 61-67 years with a BMI of 30-32 kg/m^2. Glycated haemoglobin (HbA1c) levels were on average lowest in DCS (median[Q25-Q75]: 47.08[42-50] mmol/mol), followed by GoDARTS (55.54[46-61] mmol/mol) and ANDIS (60.06[46.0-68.0]). The time from diagnosis to sampling time ranged from 0 to 2.63 years. Three phenotypic models were explored in the included cohorts which showed concordance with BMI, use of glucose-lowering drugs being risk factors and age, HDL and C-peptide being protective (**Table S3**).

Metabolites are associated with increased diabetes risk and progression

Out of the 19 small metabolites examined, five were associated with disease progression with nominal significance in the base model (age, sex, BMI adjusted, P<0.05) in the meta-analysis of three cohorts. These were homocitrulline (Hcit), aminoadipic acid (AADA), isoleucine (Ile), glycocholic acid (GCA), taurocholic acid (TCA). Out of the five, the association of two remained significant after multiple testing adjustment, including aminoadipic acid (AADA, HR[CI]=1.11[1.04-1.19], P_{FDR} = 0.03) and homocitrulline (Hcit, 1.12[1.04-1.21], P_{FDR} = 0.04, **Fig. 1, Table S4).** Both metabolites showed associations in the same direction in the replication cohorts, but non-significant with attenuated effect sizes (AADA, HR[CI]=1.03[0.96-1.11]; Hcit 1.03[0.88-1.21]). In external validation cohorts, Hcit showed a trend as a risk factor for incident diabetes (HR[CI] = 1.05[0.74-1.48]) in MDC. Based on a logistic model in DESIR,

Hcit was a risk factor for prevalent diabetes (OR[CI] = 1.32[1.05-1.66]), but not incident diabetes (HR[CI] = 0.97[0.73-1.30]). AADA has previously been associated with a higher risk of incident type 2 diabetes in Wang et al. (OR[CI] = 1.60[1.19-2.16]).[27] Finally, the most consistent risk factor for time to insulin was isoleucine level, which was nominally significant in the discovery cohort (HR[CI] = 1.09[1.00-1.20]), a risk factor for incident diabetes in MDC (HR[CI] = 1.48[1.26-1.74]) and DESIR (OR[CI] = 23.88[3.13-182.31) as well as prevalent diabetes (OR[CI] = 10.94[3.94,30.32]). Finally, GCA and TCA were modest risk factors for time to insulin requirement, with hazard ratios of 1.09[1.01-1.17] and 1.06[1.01-1.12], respectively. In the replication set both TCA and GCA were in the same direction, but no longer significant with hazard ratios of 1.09[0.91-1.31] and 1.04[0.94,1.12].

Plasma triglyceride levels are markers of diabetes progression and incident diabetes

Among the 162 lipids investigated, the levels of nine reached significance in the base model. Among these eight lipids were a risk factor for early insulin requirement, and these were all triglycerides (**Fig. 2, Table S5**). These eight lipids were also a risk factor for incident diabetes in MDC (**Fig. 2**). A single lipid was *protective* for early insulin initiation (SM 42:2;2, HR[CI] = 0.85[0.77-0.94]). Interestingly, SM 42:2;2 was a *risk factor* for incident diabetes in MDC (HR[CI] = 1.16[1.06-1.27]). Further adjustment in the discovery cohort did attenuate the effect size but the direction remained the same and four and three remained significant in the partly and fully adjusted model, respectively **(Table S5)**.

Plasma proteins levels associate with diabetes progression and prevalent and incident diabetes

In the 1195 investigated plasma proteins, the levels of 98 were nominally associated with time to insulin in the base model. Additional adjustment attenuated the hazard ratios only minimally

in both the partly and fully adjusted model. MIC-1/GDF15 was the protein associated with the highest risk of progression (HR[CI] = 1.34[1.17-1.54]) and this association was replicated in ACCELERATE (HR[CI] = 1.22[1.04-1.42]). The protein associated with the second highest risk of progression was the Nogo receptor (NogoR, HR[CI] = 1.33[1.03-1.72], **Fig. 3, Table S6**). In ANDIS, NogoR also replicated (HR[CI] = 1.20[1.07-1.34], **Fig. 3**). NogoR was also a risk factor for incident (OR[CI] = 1.45[1.15-1.83]) and prevalent diabetes in AGES-Reykjavik (OR[CI] = 1.77[1.60-1.95]). In the top associated proteins, four were protective including SMAC, coactosin-like protein, testican-1 and HEMK2, of which HEMK2 was the most protective (HR[CI] = 0.78[0.68-0.89]). In the AGES-Reykjavik study, HEMK2 was also protective for prevalent diabetes (OR[CI] = 0.78[0.72,0.85]).

Evidence of causality of biomarkers on incident diabetes based on Mendelian Randomisation

To assess causality of the identified biomarkers we would ideally have tested against the genetics of time to insulin requirement in people with diabetes, but in the current study the outcome was underpowered (n =14000) and there is no publicly available data for time to insulin genetic variants. Instead, we investigated the causality of biomarkers on type 2 diabetes. We found no significant associations with incident diabetes for any of the top metabolites (**Table 1)**. However one of the nominally significant phosphatidyl ethanaolamines, PE 18:0;0-18:2;0 has support for a causal association with type 2 diabetes (**Table 1**, β[95%CI] = -0.07[-0.12 – -0.02], P = $3.89 \cdot 10^{-3}$). For the proteins, modest evidence of a causal relation was observer for three proteins, GDF15 (β[95%CI] = 0.03[0.01– 0.05], P= $2.68 \cdot 10^{-3}$), IL-18Ra (β[95%CI] = 0.02[0.003– 0.03], P= 0.014) and FAS (β[95%CI] = 0.05[0.005-0.09], P = 0.03). For the other protein biomarkers there was no evidence of a causal relation.

491 **Functional analyses of identified protein biomarkers**

492 *Glucose-stimulated insulin secretion*. We chose to study six protein biomarkers with greatest
493 effect size (GDF15, IL-18Ra, NogoR, CRELD1, FAS, ENPP7) and which accelerated
494 progression, *i.e.*, those which may plausibly exert a deleterious effect on insulin secretion or
495 action. None of these affected basal (3 mM glucose) or high (17 mM) glucose-stimulated
496 insulin secretion (GSIS) acutely (1h) or after longer incubations (48h) from either mouse (**Fig.**
497 **4a,b**) or human (**Fig. 4c**) islets. Glucagon-like peptide-1 (GLP-1) was used as a positive
498 control, and stimulated secretion at the higher glucose concentration, as expected.

499 To determine whether any of the examined protein biomarkers might affect pancreatic
500 beta cell mass, we next assessed their impact on beta cell apoptosis (**Fig. 5a-c**) and on
501 proliferation (d) in mouse islets. Of those examined, only IL-18Ra (3-fold), and NogoR (> 15-
502 fold) exerted an effect, increasing apoptosis in mouse islets. Dose response analyses (**Fig. 5b**)
503 revealed that the effects of NogoR were apparent only at high concentrations (> 1 nM) likely
504 to be above the normal physiological range. At 100 nM NogoR, but not IL-18Ra also enhanced
505 apoptosis in human islets (**Fig. 5c**). None of the tested compounds affected human islet
506 proliferation (**Fig. 5d**)

507

508 **NogoR improves glucose tolerance in vivo**

509 GDF15/MIC has been the subject of several earlier studies (see Discussion) and so was not
510 pursued further here. Since NogoR was the biomarker with the next-largest correlation with
511 disease progression, we next sought to determine whether this protein may influence glucose
512 homeostasis in vivo. Administered daily for two weeks to mice previously maintained on a
513 HFHS diet (four weeks[28]), we observed a clear improvement in glucose tolerance in vivo versus
514 vehicle-injected animals (**Fig. 6a-d**). These changes were observed with no change in insulin
515 secretion (**Fig. 6e**).

Since these results suggested that NogoR might affect insulin signalling in disease relevant tissues such as the liver, we measured the action of this protein on signalling events downstream to insulin receptor activation, tested in HepG2 cells. Although not achieving statistical significance, a trend was observed towards a potentiation of insulin-stimulated phosphorylation of the protein kinase AKT on Ser473 after culture at different concentrations of NogoR (1 nM, 10 nM and 100 nM) for 3h. (**SFig. 1a**). In contrast, no differences were observed in insulin-stimulated Akt phosphorylation in HepG2 cells when cultured with different concentrations of CRELD1 (**SFig. 1b**).

Effect of IL-18Ra on IL-18Ra signalling

After NogoR, IL-18Ra exerted the third strongest impact on diabetes progression (**Fig. 3**). We therefore tested the effects of IL-18Ra in a reporter cell line expressing the IL-18R and a luciferase construct under the control of the cytokine-regulated transcription factor, nuclear factor κB (NF-κB; Methods). IL-18Ra attenuated the actions of IL-18 over a range of concentrations at concentrations as low as 0.1 nM (**Fig. 7**).

DISCUSSION

We have undertaken a large multi-omic study, across three patient cohorts to discover novel lipid, metabolite and protein biomarkers for diabetes progression. Many of our findings are replicated in independent diabetes progression cohorts or validated for incident and/or prevalent diabetes. In particular, we identify 9 lipids, 3 small charged molecules and 11 protein biomarkers associated with accelerated glycaemic deterioration, and provide biological data in pre-clinical models demonstrating possible mechanisms of action for NogoR and IL-18Ra. Strikingly, measurements of proteins and lipid data reveal that the drivers of diabetes incidence and prevalence may be similar to those of progression.

Metabolites and glycaemic deterioration

Two metabolites, Aminoadipic Acid (AADA) and Homocitrulline (Hcit) were significantly associated with diabetes progression, after correcting for multiple testing; three metabolites, Isoleucine (Ile), and the bile acids GCA and TCA, were nominally associated with progression. Isoleucine is a branched chain amino acid (BCAA) and a well-established risk factor for insulin resistance and increased risk of type 2 diabetes.[29] Correspondingly, Ile was also associated with both prevalent and incident diabetes in the current study. GCA and TCA are both existing markers for pre-existing diabetes.[30]

In previous studies, AADA was shown to be a risk factor for incident diabetes[27] and Hcit with prevalent diabetes[31], consistent with our findings. AADA is an alpha amino acid formed as a downstream product of lysine oxidation by the action of myeloperoxidase (MPO).[32] Higher levels of plasma AADA have been associated with obesity, insulin resistance, and increased risk of diabetes.[27,33-35] Hcit levels have been associated with disrupted energy metabolism in rat brains and have been linked to chronic renal failure.[36,37] Of note, however, where there was a sufficient genetic instrument, none of the metabolites we assessed with mendelian randomisation were causally associated with diabetes risk, although it should be noted that BCAAs have been suggested to be causal in the aetiology of type 2 diabetes.[38]

Lipids and glycaemic deterioration

Nine lipids were associated with diabetes progression of which eight were associated with increased risk which all belonged to the Triacylglycerol (TAG) class. In the external validation data, all eight TAGs were associated with increased risk of type 2 diabetes. TAGs have previously been associated with incident T2D risk.[39] TAG species levels also strongly decline when obese individuals with diabetes undergo Roux-en-Y gastric bypass.[40] SM 42:2;2 was the

sole lipid associated with lower risk on progression towards to insulin, but was associated with increased risk on future diabetes in the external validation data. A possible explanation for this could be that metformin treatment influences the sphingomyelin levels including SM 42:2;2 as has been shown in two studies in metformin treated HFD animals and human hepatocytes respectively.[41,42] The top associated lipids were also investigated for causality but generally instruments were not available or no evidence was observed. The MR analysis of PE 18:0;0_18:2;0 supported a possible causal relation with incident diabetes. `

Proteins and glycaemic deterioration

The protein with the strongest association with time to insulin was GDF-15/MIC1. This protein has previously been implicated in diabetes incidence and the control of food intake[43,44], acting via receptors in the hind brain.[28] GDF-15 has also been reported to serve as a useful biomarker for impaired fasting glucose[45] and diabetic kidney disease[46] as well as a number of conditions including cardiovascular disease[47,48] GDF15 is strongly elevated following metformin exposure, due to release from the gut.[49] Of note, we did not show any attenuation of the GDF15 signal with progression when adjusting for whether the patients were treated with metformin at the time of blood sampling.

In addition, we identify several novel proteins biomarkers to be associated with glycaemic deterioration in diabetes, including NogoR (RTN4), IL-18Ra, CRELD1, ENPP7 and FAS. Interestingly, NogoR has previously been associated with prevalent diabetes and diabetes incidence.[13] In an effort to understand the potential impact of an increase in circulating NogoR on glucose metabolism, we demonstrated that injection of this biomarker improves glucose tolerance in high fat- high sucrose-fed mice, an effect likely to reflect improved insulin sensitivity; insulin secretion was not significantly affected (indeed tended to be *increased* after treatment). NogoR, (encoded by the *RTN4R* gene), is chiefly expressed in the central nervous

exist and mediates interactions between the myelin sheath and neurons, and has been implicated in Alzheimer's disease.[50] In this setting, NogoR interacts with Oligodendrocyte myelin glycoprotein (OMGP) and Nogo-A, present on myelin cells, to inhibit neuronal regeneration.[51] Thus NogoR serves as a receptor for Nogo-A but conceivably also the shorter homologues Nogo-B and Nogo-C.[52] Importantly in the context of systemic glucose homeostasis, Nogo-B interacts with the Nogo-B receptor (NGBR, encoded by *NUS1*) and knockout of *Nus1* in mice causes hepatic steatosis, possibly be interfering with insulin signalling.[53] Furthermore, variants in the human *NUS1* gene including rs4443534 are associated with altered type 2 diabetes risk.[54] Thus, by titrating Nogo-B (or other Nogo family members) circulating NogoR may act indirectly on the liver to affect glucose storage or glycogen breakdown. On the other hand, NogoR also enhanced cell death in pancreatic islets, at least at high concentrations of this biomarker. The mechanisms involved in this action are unclear given the absence of expression of the parental receptors of the interactors Nogo-A and OMGP in islets.[55] Furthermore, the physiological relevance of these changes is uncertain since concentrations at which NogoR exerted these effects in vitro ($\geq$ 3 nM) comfortably exceeded those which improved glucose metabolism in vivo (~ 0.7 nM). We suspect, therefore, that positive effects of NogoR to improve insulin sensitivity may predominant in most physiological settings, though we do not exclude a shift towards deleterious actions on the pancreas in disease settings.

CRELD1 (Cysteine-Rich with EGF-Like Domains 1, AVSD2, Cirrin) is a membrane-bound Ca^{2+}-binding member of the EGF family critically required for normal development of the heart[56] whose expression was recently shown to influence T-cell activity and immune homeostasis.[57] Although we were not able to test the effects of this agent on glucose homeostasis in vivo due to the unavailability of the human protein, we do provide evidence that CRELD1 may regulate insulin signalling *in vitro*.

A recent Mendelian randomisation study[13] provided nominally significant support for a causal association between IL-18Ra and T2D, which was repeated here. Consistent with a role in glucose homeostasis, IL-18 deletion in the mouse leads to obesity and insulin resistance.[58] Conversely, after weight loss following exercise/diet or bariatric surgery in man, a significant reduction in IL-18 concentrations was observed.[59,60] IL-18 secretion also increased in response to inflammasome activation and pyroptosis.[61] We confirmed earlier studies[62] which demonstrated that an IL-18Ra:Fc fusion fragment inhibits the pro-inflammatory action of IL-18. However, and in distinction to these earlier studies, measuring interferon-gamma (IFN-gamma) production from mononuclear cells, the actions of IL-18Ra did not depend upon the additional presence of IL-18Rbeta. We also noted that the concentrations of IL-18Ra as low as 0.1 nM efficiently inhibited the actions of a considerable excess (100 nM) of IL-18 potentially suggesting a non-competitive action (*i.e.*, binding to the IL-18R at a separate site on the cellular receptor to IL-18). Unexpectedly, IL-18Ra also influenced beta cell apoptosis in vitro. The mechanisms involved here are unclear, however, since IL-18Ra expression levels in the beta cell are low.[63] Finally, we note that IL-18Ra, CRELD1 and coactosin-like protein are all involved in immune regulation, and might thus influence the inflammatory changes known to be involved in T2D.[64]

ENPP7 (Ectonucleotide pyrophosphatase/phosphodiesterase-7) is strongly expressed in the small intestine where it is involved in sphingomyelin hydrolysis and the absorption of ceramide and phosphocholine.[65] These processes might therefore be influenced by changed circulating levels of ENPP7. FAS (FAS cell surface death receptor; TNF superfamily receptor-6) is involved in caspase 3 and 8 activation and cell death[66] and might thus influence cell survival in critical metabolic tissues. HSP 90B (heat shock protein 90 alpha family class B member 1), encoded by *HSP90AB1*, is a molecular chaperone and regulator of protein folding.[67]

Other identified proteins were associated with protection from glycaemic deterioration of diabetes. SMAC/IAPP-binding mitochondrial protein, also called DIABLO, is a mitochondrially-associated protein which migrates to the cytosol upon the activation of apoptosis, facilitating this process by restricting the activity of apoptotic inhibitors.[68] Lowered levels of this protein in the plasma thus seem likely to reflect diminished levels of cell death in disease-relevant (or other) tissues. Coactosin-like protein, encoded by the *COTL1* gene and also called Coactosin-like F-actin binding protein, and CLP, is enriched in haematopoietic cells (BioGPS), and regulates leukotriene synthesis. Low levels may therefore reflect more limited inflammation in individuals whose disease deteriorates quickly. Testican-1, also called SPOCK1, SPARC, and Osteonectin is enriched in the brain. The function of SPOCK1, a Ca^{2+} binding proteoglycan, is currently unknown, though roles in neuronal development[69] adipocyte differentiation[70] and as an extracellular matrix factor controlling epithelial to mesenchymal transition[71] have been suggested. Finally, HEMK2 (N6AMT1, PrmC) is a ubiquitously-expressed DNA methyl transferase.[72]

To investigate causal associations, we undertook MR analysis of six of our identified protein biomarkers that were associated with diabetes progression. We repeated previous findings that GDF15 is causally associated with diabetes progression[13] and found a possible causal association for IL-18Ra and FAS. These data suggest that these two proteins are likely to be causally associated with diabetes progression. The lack of a causal association of NogoR with diabetes risk does suggest that NogoR may similarly not be causally associated with diabetes progression, however, our functional studies do suggest a causal mechanism linking NogoR with abnormal glucose metabolism.

This study has several limitations. Between cohorts there was heterogeneity, and this was in part due to the nature of the cohorts, for example the ACCELERATE cohort is a clinical trial which is different in setup than the discovery studies. Nonetheless, the results of the study

were robust, for example a large number of biomarkers were also found to be associated with prevalent and incident diabetes. A third limitation is that for some protein biomarkers ELISA assays are currently unavailable to validate the signals obtained in the SomaLogic screen. Finally, in the Mendelian Randomization analysis we could only investigate the causal association with diabetes risk, and some of the genetic instruments were weak, so the absence of causality in MR analysis does not mean that our findings with diabetes progression were non-causal. Indeed, plausible evidence for causality was subsequently obtained in functional studies in preclinical models for NogoR and IL-18Ra.

CONCLUSION

Our findings provide new mechanistic insights into the mechanisms that underlie glycaemic deterioration once diabetes has developed. Importantly, we describe novel biomarkers of different chemical classes, several of which have not previously been associated with the incident or prevalent disease, suggestive of potentially distinct mechanisms driving the two stages of glycaemic deterioration - towards diabetes onset and after diabetes onset. We provide direct functional analyses implicating two of these (NogoR, IL-18Ra) as likely to contribute directly to disease progression. By better understanding the biological drivers of glycaemic deterioration in diabetes it may be possible to target therapies to these processes to prevent or slow diabetes progression, rather than simply to lower HbA1c, potentially transforming diabetes treatments.

ACKNOWLEDGEMENTS

We acknowledge the support of the Health Informatics Centre, University of Dundee for managing and supplying the anonymised data. The authors wish to thank the Core-IT group of SIB Swiss Institute of Bioinformatics and in particular Jorge Molina for expert technical help

691 in setting up and maintaining the federated database. GR thanks Claudio Elgueta Karstegl for
692 technical support. The authors thank participants of the included cohorts.
693
694 **FUNDING**
695 ERP holds a Wellcome Trust New Investigator Award (102820/Z/13/Z). GAR was supported
696 by a Wellcome Trust Investigator Award (212625/Z/18/Z), MRC Programme grants
697 (MR/R022259/1, MR/J0003042/1, MR/L020149/1), an Experimental Challenge Grant (DIVA,
698 MR/L02036X/1), a Diabetes UK Project grant (BDA16/0005485). This project has received
699 funding from the Innovative Medicines Initiative 2 Joint Undertaking, under grant agreement
700 no. 115881 (RHAPSODY). This Joint Undertaking receives support from the European
701 Union's Horizon 2020 research and innovation programme and EFPIA. This work is supported
702 by the Swiss State Secretariat for Education, Research and Innovation (SERI), under contract
703 no. 16.0097. VaG is supported by the Icelandic Research Fund (grant no. 184845-051). The
704 Hoorn DCS cohort was supported by grants from the Netherlands Organisation for Health
705 Research and Development (113102006, 459001015)
706
707 **CONFLICTS OF INTEREST**
708 KS is CEO of Lipotype GmbH. KS and CK are shareholders of Lipotype GmbH. MJG is
709 employee of Lipotype GmbH. GAR has received grant funding and consultancy fees from Sun
710 Pharmaceuticals and Les Laboratoires Servier. MKH is an employee of Janssen Research &
711 Development, LLC. AF and IP are employees of Eli Lilly Regional Operations GmbH. The
712 AGES-Reykjavik proteomics study was supported by the Novartis Institute for Biomedical
713 Research, and protein measurements for the AGES-Reykjavik cohort were performed at
714 SomaLogic. L.L.J. is an employee and stockholder of Novartis. PR (Peter Rossing) has
715 received honoraria to Steno Diabetes Center Copenhagen for consultancy and teaching from

Astellas, Astra Zeneca, Boehringer Ingelheim, Bayer, Novo Nordisk, Sanofi, Gilead and Vifor and research grants from Novo Nordisk and Astra Zeneca.

AUTHOR CONTRIBUTION

RCS, LAD, JWJB, LMTH, ERP, and GR designed the study. GR and RCS drafted the manuscript. RCS, LAD, HF, GAB, MA performed the analyses. GR oversaw biomarker shipments and assays, and coordinated all functional work in preclinical models. MS, MB, LMtH, AAWAvH, PJME, JWJB, EA, LG, LAD, ERP provided patient samples for analysis. LL, HM, EG, IL performed all studies on NogoR, EA, MS work on IL-18Ra. ID, DK, FB, DM, AN, MI set up a federated node system for data-analysis. RCS, LAD, HF, DMA, EA, AA, MJG, MK, FM, TS, AW, CLQ, MI were involved in the (clinical) data pre-processing and quality control. GNG, AF, MKH, DMA, IP, TJP, BT, VL, LG, PWF, GAR contributed to the data acquisition and project logistics. MJG, CK, KS generated the Lipotype data. CLQ, AA, PR, AW, TS, FM generated the metabolomics data and/or were involved in the quality control. FO, CF and OM acquired the MDC-CC data and performed the lipid and protein validation in the MDC-CC. MC and PF performed the metabolite validation in DESIR. VaG, ViG and LLJ performed the protein validation in the AGES-Reykjavik cohort. PJME and AAWAvH acquired the data from the Hoorn DCS cohort. All authors contributed to the data interpretation. All authors critically revised the manuscript and approved the final version. RCS and LAD are the guarantors of the work.

DATA DEPOSITION

Summary statistics of lipidomic, proteomic and metabolomic data will be available from an interactive Shiny dashboard available upon publication.

REFERENCES

1 International Diabetes Federation. Diabetes facts & figures. (2019).

2 Chung, W. K. *et al.* Precision Medicine in Diabetes: A Consensus Report From the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). *Diabetes Care* **43**, 1617-1635, doi:10.2337/dci20-0022 (2020).

3 Ahluwalia, T. S., Kilpeläinen, T. O., Singh, S. & Rossing, P. Editorial: Novel Biomarkers for Type 2 Diabetes. *Front Endocrinol (Lausanne)* **10**, 649, doi:10.3389/fendo.2019.00649 (2019).

4 Kolberg, J. A. *et al.* Development of a type 2 diabetes risk model from a panel of serum biomarkers from the Inter99 cohort. *Diabetes Care* **32**, 1207-1212, doi:10.2337/dc08-1935 (2009).

5 Thakarakkattil Narayanan Nair, A. *et al.* The impact of phenotype, ethnicity and genotype on progression of type 2 diabetes mellitus. *Endocrinol Diabetes Metab* **3**, e00108, doi:10.1002/edm2.108 (2020).

6 Jiang, G. *et al.* Obesity, clinical, and genetic predictors for glycemic progression in Chinese patients with type 2 diabetes: A cohort study using the Hong Kong Diabetes Register and Hong Kong Diabetes Biobank. *PLoS Med* **17**, e1003209, doi:10.1371/journal.pmed.1003209 (2020).

7 van der Heijden, A. A. *et al.* The Hoorn Diabetes Care System (DCS) cohort. A prospective cohort of persons with type 2 diabetes treated in primary care in the Netherlands. *BMJ open* **7**, e015599 (2017).

8 Hebert, H. L. *et al.* Cohort Profile: Genetics of Diabetes Audit and Research in Tayside Scotland (GoDARTS). *Int J Epidemiol* **47**, 380-381j, doi:10.1093/ije/dyx140 (2018).

9 *ANDIS - Alla Nya Diabetiker I Skåne*, <http://andis.ludc.med.lu.se/> (2020).

10 Slieker, R. C. *et al.* Replication and cross-validation of T2D subtypes based on clinical variables: an IMI-RHAPSODY study. *medRxiv*, doi:MEDRXIV/2020/248628 (2020).

11 Lincoff, A. M. *et al.* Evacetrapib and Cardiovascular Outcomes in High-Risk Vascular Disease. *N Engl J Med* **376**, 1933-1942, doi:10.1056/NEJMoa1609581 (2017).

12 Harris, T. B. *et al.* Age, Gene/Environment Susceptibility-Reykjavik Study: multidisciplinary applied phenomics. *Am J Epidemiol* **165**, 1076-1087, doi:10.1093/aje/kwk115 (2007).

13 Gudmundsdottir, V. *et al.* Circulating Protein Signatures and Causal Candidates for Type 2 Diabetes. *Diabetes* **69**, 1843-1853, doi:10.2337/db19-1070 (2020).

14 Ottosson, F. *et al.* A plasma lipid signature predicts incident coronary artery disease. *Int J Cardiol*, doi:10.1016/j.ijcard.2021.01.059 (2021).

15 Fernandez, C. *et al.* Plasma Lipidome and Prediction of Type 2 Diabetes in the Population-Based Malmö Diet and Cancer Cohort. *Diabetes Care* **43**, 366-373, doi:10.2337/dc19-1199 (2020).

16 Yengo, L. *et al.* Impact of statistical models on the prediction of type 2 diabetes using non-targeted metabolomics profiling. *Mol Metab* **5**, 918-925, doi:10.1016/j.molmet.2016.08.011 (2016).

17 Ahonen, L. *et al.* Targeted Clinical Metabolite Profiling Platform for the Stratification of Diabetic Patients. *Metabolites* **9**, doi:10.3390/metabo9090184 (2019).

18 Surma, M. A. *et al.* An automated shotgun lipidomics platform for high throughput, comprehensive, and quantitative analysis of blood plasma intact lipids. *Eur J Lipid Sci Technol* **117**, 1540-1549, doi:10.1002/ejlt.201500145 (2015).

19 Herzog, R. *et al.* LipidXplorer: a software for consensual cross-platform lipidomics. *PLoS One* **7**, e29851, doi:10.1371/journal.pone.0029851 (2012).

20 Aimo, L. *et al.* The SwissLipids knowledgebase for lipid biology. *Bioinformatics* **31**, 2860-2866, doi:10.1093/bioinformatics/btv285 (2015).
21 Zhou, K. *et al.* Clinical and genetic determinants of progression of type 2 diabetes: a DIRECT study. *Diabetes Care* **37**, 718-724, doi:10.2337/dc13-1995 (2014).
22 Wolfson, M. *et al.* DataSHIELD: resolving a conflict in contemporary bioscience--performing a pooled analysis of individual-level data without sharing the data. *Int J Epidemiol* **39**, 1372-1382, doi:10.1093/ije/dyq111 (2010).
23 Dragan, I., Sparsø, T., Kuznetsov, D., Slieker, R. & Ibberson, M. dsSwissKnife: An R package for federated data analysis. *bioRxiv*, 2020.2011.2017.386813, doi:10.1101/2020.11.17.386813 (2020).
24 Tabassum, R. *et al.* Genetic architecture of human plasma lipidome and its link to cardiovascular disease. *Nat Commun* **10**, 4329, doi:10.1038/s41467-019-11954-8 (2019).
25 Lotta, L. A. *et al.* A cross-platform approach identifies genetic regulators of human metabolism and health. *Nat Genet* **53**, 54-64, doi:10.1038/s41588-020-00751-5 (2021).
26 Mahajan, A. *et al.* Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *Nat Genet* **50**, 1505-1513, doi:10.1038/s41588-018-0241-6 (2018).
27 Wang, T. J. *et al.* 2-Aminoadipic acid is a biomarker for diabetes risk. *J Clin Invest* **123**, 4309-4317, doi:10.1172/jci64801 (2013).
28 Tsai, V. W. *et al.* GDF15 mediates adiposity resistance through actions on GFRAL neurons in the hindbrain AP/NTS. *Int J Obes (Lond)* **43**, 2370-2380, doi:10.1038/s41366-019-0365-5 (2019).
29 Lynch, C. J. & Adams, S. H. Branched-chain amino acids in metabolic signalling and insulin resistance. *Nat Rev Endocrinol* **10**, 723-736, doi:10.1038/nrendo.2014.171 (2014).
30 Andersen, E., Karlaganis, G. & Sjövall, J. Altered bile acid profiles in duodenal bile and urine in diabetic subjects. *European journal of clinical investigation* **18**, 166-172 (1988).
31 Suhre, K. *et al.* Metabolic footprint of diabetes: a multiplatform metabolomics study in an epidemiological setting. *PLoS One* **5**, e13953, doi:10.1371/journal.pone.0013953 (2010).
32 Lin, H. *et al.* Myeloperoxidase-mediated protein lysine oxidation generates 2-aminoadipic acid and lysine nitrile in vivo. *Free Radic Biol Med* **104**, 20-31, doi:10.1016/j.freeradbiomed.2017.01.006 (2017).
33 Wijekoon, E. P., Skinner, C., Brosnan, M. E. & Brosnan, J. T. Amino acid metabolism in the Zucker diabetic fatty rat: effects of insulin resistance and of type 2 diabetes. *Can J Physiol Pharmacol* **82**, 506-514, doi:10.1139/y04-067 (2004).
34 Gao, X. *et al.* Serum metabolic biomarkers distinguish metabolically healthy peripherally obese from unhealthy centrally obese individuals. *Nutr Metab (Lond)* **13**, 33, doi:10.1186/s12986-016-0095-9 (2016).
35 Cobb, J. *et al.* α-Hydroxybutyric Acid Is a Selective Metabolite Biomarker of Impaired Glucose Tolerance. *Diabetes Care* **39**, 988-995, doi:10.2337/dc15-2752 (2016).
36 Viegas, C. M. *et al.* Experimental evidence that ornithine and homocitrulline disrupt energy metabolism in brain of young rats. *Brain Res* **1291**, 102-112, doi:10.1016/j.brainres.2009.07.021 (2009).
37 Desmons, A. *et al.* Homocitrulline: a new marker for differentiating acute from chronic renal failure. *Clin Chem Lab Med* **54**, 73-79, doi:10.1515/cclm-2015-0398 (2016).
38 Lotta, L. A. *et al.* Genetic Predisposition to an Impaired Metabolism of the Branched-Chain Amino Acids and Risk of Type 2 Diabetes: A Mendelian Randomisation Analysis. *PLoS Med* **13**, e1002179, doi:10.1371/journal.pmed.1002179 (2016).
39 Rhee, E. P. *et al.* A genome-wide association study of the human metabolome in a community-based cohort. *Cell Metab* **18**, 130-143, doi:10.1016/j.cmet.2013.06.013 (2013).

40 Graessler, J. *et al.* Lipidomic profiling before and after Roux-en-Y gastric bypass in obese patients with diabetes. *Pharmacogenomics J* **14**, 201-207, doi:10.1038/tpj.2013.42 (2014).
41 Bornstein, S. *et al.* Metformin Affects Cortical Bone Mass and Marrow Adiposity in Diet-Induced Obesity in Male Mice. *Endocrinology* **158**, 3369-3385, doi:10.1210/en.2017-00299 (2017).
42 Wanninger, J. *et al.* Metformin reduces cellular lysophosphatidylcholine and thereby may lower apolipoprotein B secretion in primary human hepatocytes. *Biochim Biophys Acta* **1781**, 321-325, doi:10.1016/j.bbalip.2008.04.012 (2008).
43 Mullican, S. E. *et al.* GFRAL is the receptor for GDF15 and the ligand promotes weight loss in mice and nonhuman primates. *Nat Med* **23**, 1150-1157, doi:10.1038/nm.4392 (2017).
44 Patel, S. *et al.* GDF15 Provides an Endocrine Signal of Nutritional Stress in Mice and Humans. *Cell Metab* **29**, 707-718.e708, doi:10.1016/j.cmet.2018.12.016 (2019).
45 Hong, J. H. *et al.* GDF15 Is a Novel Biomarker for Impaired Fasting Glucose. *Diabetes Metab J* **38**, 472-479, doi:10.4093/dmj.2014.38.6.472 (2014).
46 Carlsson, A. C. *et al.* Growth differentiation factor 15 (GDF-15) is a potential biomarker of both diabetic kidney disease and future cardiovascular events in cohorts of individuals with type 2 diabetes: a proteomics approach. *Ups J Med Sci* **125**, 37-43, doi:10.1080/03009734.2019.1696430 (2020).
47 Khan, S. & Rasool, S. T. Current use of cardiac biomarkers in various heart conditions. *Endocr Metab Immune Disord Drug Targets*, doi:10.2174/1871530320999200831171748 (2020).
48 Pang, L., Ge, L., Yang, P., He, H. & Zhang, H. Degradation of organophosphate esters in sewage sludge: Effects of aerobic/anaerobic treatments and bacterial community compositions. *Data Brief* **17**, 1030-1035, doi:10.1016/j.dib.2018.02.039 (2018).
49 Coll, A. P. *et al.* GDF15 mediates the effects of metformin on body weight and energy balance. *Nature* **578**, 444-448, doi:10.1038/s41586-019-1911-y (2020).
50 Xu, Y. Q., Sun, Z. Q., Wang, Y. T., Xiao, F. & Chen, M. W. Function of Nogo-A/Nogo-A receptor in Alzheimer's disease. *CNS Neurosci Ther* **21**, 479-485, doi:10.1111/cns.12387 (2015).
51 Llorens, F., Gil, V. & del Río, J. A. Emerging functions of myelin-associated proteins during development, neuronal plasticity, and neurodegeneration. *Faseb j* **25**, 463-475, doi:10.1096/fj.10-162792 (2011).
52 Zhang, R., Tang, B. S. & Guo, J. F. Research advances on neurite outgrowth inhibitor B receptor. *J Cell Mol Med* **24**, 7697-7705, doi:10.1111/jcmm.15391 (2020).
53 Hu, W. *et al.* Nogo-B receptor deficiency increases liver X receptor alpha nuclear translocation and hepatic lipogenesis through an adenosine monophosphate-activated protein kinase alpha-dependent pathway. *Hepatology* **64**, 1559-1576, doi:10.1002/hep.28747 (2016).
54 Type 2 diabetes Knowledge Portal. <https://t2d.hugeamp.org/region.html?chr=6&end=118081803&phenotype=HEIGHT&start=117946665> (2021).
55 Wu, C., Jin, X., Tsueng, G., Afrasiabi, C. & Su, A. I. BioGPS: building your own mash-up of gene annotations and expression profiles. *Nucleic Acids Res* **44**, D313-316, doi:10.1093/nar/gkv1104 (2016).
56 Rupp, P. A. *et al.* Identification, genomic organization and mRNA expression of CRELD1, the founding member of a unique family of matricellular proteins. *Gene* **293**, 47-57, doi:10.1016/s0378-1119(02)00696-0 (2002).
57 Bonaguro, L. *et al.* CRELD1 modulates homeostasis of the immune system in mice and humans. *Nat Immunol* **21**, 1517-1527, doi:10.1038/s41590-020-00811-2 (2020).

58 Netea, M. G. *et al.* Deficiency of interleukin-18 in mice leads to hyperphagia, obesity and insulin resistance. *Nat Med* **12**, 650-656, doi:10.1038/nm1415 (2006).
59 Esposito, K. *et al.* Effect of weight loss and lifestyle changes on vascular inflammatory markers in obese women: a randomized trial. *Jama* **289**, 1799-1804, doi:10.1001/jama.289.14.1799 (2003).
60 Botella-Carretero, J. I. *et al.* The decrease in serum IL-18 levels after bariatric surgery in morbidly obese women is a time-dependent event. *Obes Surg* **17**, 1199-1208, doi:10.1007/s11695-007-9202-3 (2007).
61 Pickering, R. J. & Bryant, C. E. Preventing pores and inflammation. *Science* **369**, 1564-1565, doi:10.1126/science.abe0917 (2020).
62 Reznikov, L. L. *et al.* The combination of soluble IL-18Ralpha and IL-18Rbeta chains inhibits IL-18-induced IFN-gamma. *J Interferon Cytokine Res* **22**, 593-601, doi:10.1089/10799900252982070 (2002).
63 Blodgett, D. M. *et al.* Novel Observations From Next-Generation RNA Sequencing of Highly Purified Human Adult and Fetal Islet Cell Subsets. *Diabetes* **64**, 3172-3181, doi:10.2337/db15-0039 (2015).
64 Jafaripour, S., Sedighi, S., Jokar, M. H., Aghaei, M. & Moradzadeh, M. Inflammation, diet, and type 2 diabetes: a mini-review. *J Immunoassay Immunochem* **41**, 768-777, doi:10.1080/15321819.2020.1750423 (2020).
65 Duan, R. D. *et al.* Purification, localization, and expression of human intestinal alkaline sphingomyelinase. *J Lipid Res* **44**, 1241-1250, doi:10.1194/jlr.M300037-JLR200 (2003).
66 Scott, F. L. *et al.* The Fas-FADD death domain complex structure unravels signalling by receptor clustering. *Nature* **457**, 1019-1022, doi:10.1038/nature07606 (2009).
67 Chadli, A. *et al.* GCUNC-45 is a novel regulator for the progesterone receptor/hsp90 chaperoning pathway. *Mol Cell Biol* **26**, 1722-1730, doi:10.1128/mcb.26.5.1722-1730.2006 (2006).
68 Du, C., Fang, M., Li, Y., Li, L. & Wang, X. Smac, a mitochondrial protein that promotes cytochrome c-dependent caspase activation by eliminating IAP inhibition. *Cell* **102**, 33-42, doi:10.1016/s0092-8674(00)00008-8 (2000).
69 Charbonnier, F., Périn, J. P., Roussel, G., Nussbaum, J. L. & Alliel, P. M. [Cloning of testican/SPOCK in man and mouse. Neuromuscular expression perspectives in pathology]. *C R Seances Soc Biol Fil* **191**, 127-133 (1997).
70 Alshargabi, R. *et al.* SPOCK1 induces adipose tissue maturation: New insights into the function of SPOCK1 in metabolism. *Biochem Biophys Res Commun* **533**, 1076-1082, doi:10.1016/j.bbrc.2020.09.129 (2020).
71 Sun, L. R., Li, S. Y., Guo, Q. S., Zhou, W. & Zhang, H. M. SPOCK1 Involvement in Epithelial-to-Mesenchymal Transition: A New Target in Cancer Therapy? *Cancer Manag Res* **12**, 3561-3569, doi:10.2147/cmar.S249754 (2020).
72 Figaro, S., Scrima, N., Buckingham, R. H. & Heurgué-Hamard, V. HemK2 protein, encoded on human chromosome 21, methylates translation termination factor eRF1. *FEBS Lett* **582**, 2352-2356, doi:10.1016/j.febslet.2008.05.045 (2008).

TABLES

Table 1. Mendelian randomization analysis on metabolites, lipids and proteins against incident type 2 diabetes

Metabolites

variable	*method*	*nsnp*	*Beta*	*Lower*	*Upper*	*P-value*	*Egger intercept*	*Q*	*Q(df)*	*Q(Pvalue)*
AADA	IVW	2	0.02	-0.09	0.13	0.77		0.02	1	0.88
Citrulline	IVW	6	-0.01	-0.29	0.28	0.97	-0.04	83.02	5	<0.001
Isoleucine	IVW	5	-0.06	-0.94	0.83	0.90	-0.11	127.69	4	<0.001
Leucine	IVW	6	-0.08	-0.74	0.58	0.81	-0.07	127.77	5	<0.001

Lipids

variable	*method*	*nsnp*	*Beta*	*Lower*	*Upper*	*P-value*	*Egger intercept*	*Q*	*Q(df)*	*Q(Pvalue)*
PE 18:0;0_18:2;0	IVW	3	-0.07	-0.12	-0.02	$3.89 \cdot 10^{-3}$	0.02	5.53	2	0.06
SM 42:2;2	IVW	2	0.00	-0.07	0.07	0.99		0.91	1	0.34
TAG 50:1;0	IVW	2	0.01	-0.04	0.05	0.75		0.00	1	0.97

Proteins

variable	*method*	*nsnp*	*Beta*	*Lower*	*Upper*	*pval*	*Egger intercept*	*Q*	*Q(df)*	*Q(Pvalue)*
GDF15	IVW	8	0.03	0.01	0.05	$2.68 \cdot 10^{-3}$	0.01	8.83	7	0.27
IL-18Ra	IVW	14	0.02	0.003	0.03	0.01	0.00	6.75	13	0.91
FAS	IVW	4	0.05	0.005	0.09	0.03	0.00	6.34	3	0.10
RTN4R	IVW	10	-0.01	-0.03	0.01	0.28	-0.01	4.16	9	0.90
ENPP7	IVW	4	0.01	-0.01	0.02	0.34	0.00	0.85	3	0.84
CRELD1	IVW	7	0.01	-0.01	0.02	0.43	0.01	7.33	6	0.29

IVW, inverse variance weighting; nsnp, number of instruments

938 **FIGURE LEGENDS**

939 **Figure 1 Metabolites associated with diabetes development and progression. a)** Hazards
940 of a time to insulin model in the three discovery cohorts plus two replication sets in two of the
941 three cohorts and their respective meta-analyses (Model 1). The figure shows the five
942 nominally significant metabolites, with Hcit and AADA being also significant after multiple
943 testing. **b)** Hazards of incident diabetes in MDC based on a Cox proportional hazards model
944 adjusted for age, sex and BMI. **c)** Odds ratios of incident and prevalent diabetes in DESIR
945 based on a logistic regression model adjusted for age, sex and BMI.

946

947 **Figure 2. Lipids associated with diabetes development and progression. a)** Hazards of a
948 time to insulin model in the three discovery cohorts and the meta-analysed hazards (Model
949 1). The figure shows the nine significant lipids after multiple testing. **b)** Hazard models of
950 incident diabetes in MDC based on a Cox proportional hazards model.

951

952 **Figure 3. Proteins in plasma or serum associated with time to insulin requirement. a)**
953 Top proteins associated with time to insulin requirement. Shown is the top 10 based on P-
954 value plus Nogo receptor, which showed the largest risk of the top hundred proteins. X-axis,
955 hazard ratio on a log2 scale and studies on the y-axis. **b)** Association between protein levels
956 and incident diabetes. X-axis, odds ratio on a log2 scale. **c)** Association between protein
957 levels and prevalent diabetes. X-axis, odds ratio on a log2 scale.

958

959 **Figure 4. Impact of identified biomarkers on insulin secretion from mouse (a,b) and**
960 **human** (c) islets. Incubations were performed for 30 min. at the indicated concentrations of
961 glucose, and secreted insulin measured using an electrochemiluminescence assay. ***, ** p =

962 0.001, 0.01 compared to vehicle; ##p<0.01 vs 3 mM glucose. Comparisons by one-way
963 ANOVA in each case. Other details are given in the Methods Section.
964
965 **Figure. 5** Impact of identified biomarkers on apoptosis (a-c) or proliferation (d) in mouse or
966 human islets as indicated. Test compounds were added at 100 nM unless otherwise indicated.
967 *,***p<0.05, 0.001 by one-way ANOVA for the effects of the indicated compounded *versus*
968 vehicle. See Methods for other details.
969
970 **Figure 6. NogoR enhances glucose clearance and insulin sensitivity in mice**. Two separate
971 cohorts of wild-type male C57BL/6J mice were maintained on a high fat diet for six weeks,
972 then injected for 14 consecutive days with saline or 100 ng (2.1 pmol/animal) recombinant
973 NogoR. **a,b)** Body weights of cohort one and circulating glucose levels during an oral glucose
974 tolerance test (OGTT;2g/kg) pre-NogoR treatment *n* =5. **c,d)** Body weights and blood glucose
975 levels after an oral glucose load (2g/kg) of cohort one after NogoR treatment. *n* = 5. **e)** Plasma
976 insulin levels after an oral glucose load (2g/kg) in cohort 1. *n* = 5 per group. **f, g)** Post-treatment
977 body weights of cohort 2 and circulating glucose levels after receiving an intraperitoneal
978 injection of 1 IU/kg of insulin. Data are represented as the mean ± SEM. *p-value < 0.05
979 Student T-test.
980
981 **Figure 7. NF-κB activation in HEK293 cells overexpressing IL-18R and co-stimulated**
982 **with Il-18 and IL-18Rα.** IL-18R HEK cells stimulated for 6 h with 100, 50, 20, 10, 5, 2 and
983 1 nM of IL-18 alone, and also with 100, 10, 1 and 0.1nM of IL-18Rα. IL-18 and IL-18Rα
984 concentrations were tested in triplicate and NF-κB activation was measured using dual
985 luciferase reporter assay. Curve fitting was done using non-linear regression with GraphPad

Prism 9.0.0 for log(agonist) vs. response (a) and inhibitor vs. response (b).

SUPPLEMENTARY FIGURE LEGENDS

Supp Figure 1. Effects of NogoR and of CRELD1 on insulin signalling in vitro. HepG2 cells were cultured for 3 hours with 0, 1, 10 or 100 nM of NogoR (a) or CRELD1 (b) and 100 nM of insulin for 15 minutes prior protein extraction. a) Protein AKT and pSer473 AKT levels in Hepg2 treated with different concentrations of NogoR (n = 3). b) Protein AKT and pSer473 AKT levels in HepG2 treated with different concentrations of NogoR (n = 3). Data are represented as the mean ± SEM and were analysed by one-way ANOVA.

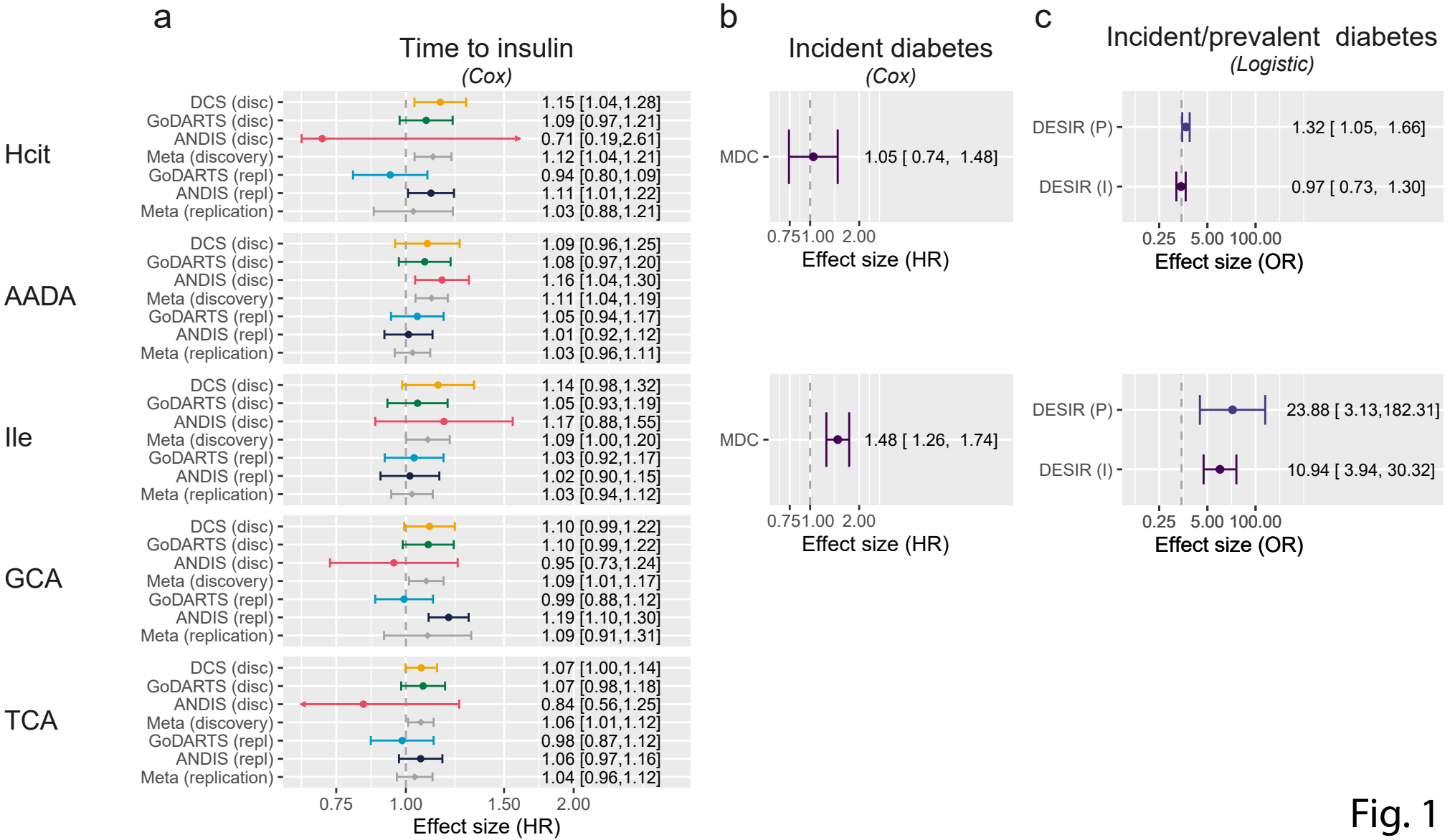

a
Time to insulin
(Cox)
Hcit
DCS (disc) 1.15 [1.04,1.28]
GoDARTS (disc) 1.09 [0.97,1.21]
ANDIS (disc) 0.71 [0.19,2.61]
Meta (discovery) 1.12 [1.04,1.21]
GoDARTS (repl) 0.94 [0.80,1.09]
ANDIS (repl) 1.11 [1.01,1.22]
Meta (replication) 1.03 [0.88,1.21]
AADA
DCS (disc) 1.09 [0.96,1.25]
GoDARTS (disc) 1.08 [0.97,1.20]
ANDIS (disc) 1.16 [1.04,1.30]
Meta (discovery) 1.11 [1.04,1.19]
GoDARTS (repl) 1.05 [0.94,1.17]
ANDIS (repl) 1.01 [0.92,1.12]
Meta (replication) 1.03 [0.96,1.11]
Ile
DCS (disc) 1.14 [0.98,1.32]
GoDARTS (disc) 1.05 [0.93,1.19]
ANDIS (disc) 1.17 [0.88,1.55]
Meta (discovery) 1.09 [1.00,1.20]
GoDARTS (repl) 1.03 [0.92,1.17]
ANDIS (repl) 1.02 [0.90,1.15]
Meta (replication) 1.03 [0.94,1.12]
GCA
DCS (disc) 1.10 [0.99,1.22]
GoDARTS (disc) 1.10 [0.99,1.22]
ANDIS (disc) 0.95 [0.73,1.24]
Meta (discovery) 1.09 [1.01,1.17]
GoDARTS (repl) 0.99 [0.88,1.12]
ANDIS (repl) 1.19 [1.10,1.30]
Meta (replication) 1.09 [0.91,1.31]
TCA
DCS (disc) 1.07 [1.00,1.14]
GoDARTS (disc) 1.07 [0.98,1.18]
ANDIS (disc) 0.84 [0.56,1.25]
Meta (discovery) 1.06 [1.01,1.12]
GoDARTS (repl) 0.98 [0.87,1.12]
ANDIS (repl) 1.06 [0.97,1.16]
Meta (replication) 1.04 [0.96,1.12]
0.75 1.00 1.50 2.00
Effect size (HR)
b
Incident diabetes
(Cox)
MDC 1.05 [0.74, 1.48]
0.75 1.00 2.00
Effect size (HR)
MDC 1.48 [1.26, 1.74]
0.75 1.00 2.00
Effect size (HR)
c
Incident/prevalent diabetes
(Logistic)
DESIR (P) 1.32 [1.05, 1.66]
DESIR (I) 0.97 [0.73, 1.30]
0.25 5.00 100.00
Effect size (OR)
DESIR (P) 23.88 [3.13,182.31]
DESIR (I) 10.94 [3.94, 30.32]
0.25 5.00 100.00
Effect size (OR)

Fig. 1

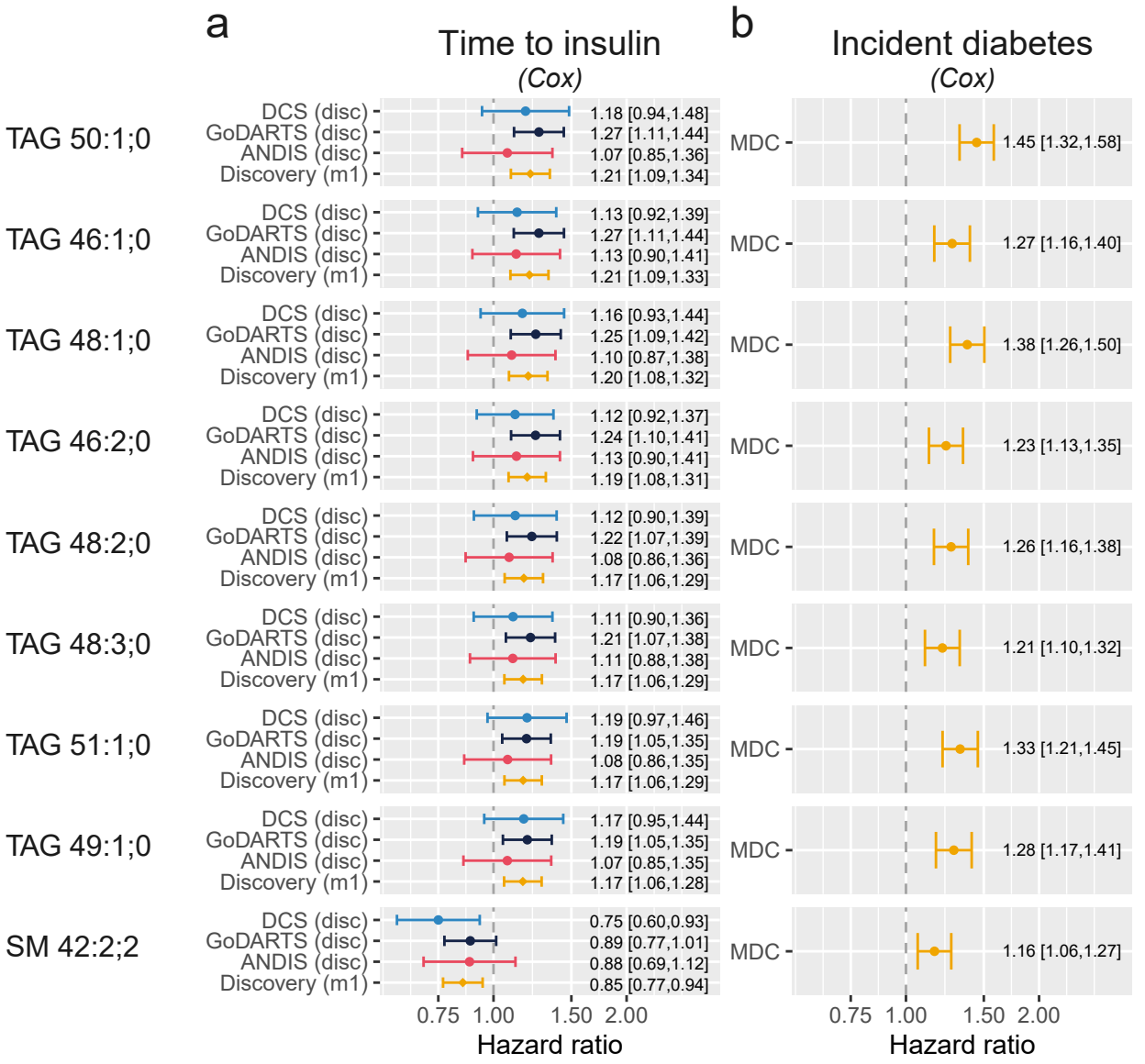

a
Time to insulin
(Cox)
b
Incident diabetes
(Cox)
TAG 50:1;0
DCS (disc) 1.18 [0.94,1.48]
GoDARTS (disc) 1.27 [1.11,1.44]
ANDIS (disc) 1.07 [0.85,1.36]
Discovery (m1) 1.21 [1.09,1.34]
MDC 1.45 [1.32,1.58]
TAG 46:1;0
DCS (disc) 1.13 [0.92,1.39]
GoDARTS (disc) 1.27 [1.11,1.44]
ANDIS (disc) 1.13 [0.90,1.41]
Discovery (m1) 1.21 [1.09,1.33]
MDC 1.27 [1.16,1.40]
TAG 48:1;0
DCS (disc) 1.16 [0.93,1.44]
GoDARTS (disc) 1.25 [1.09,1.42]
ANDIS (disc) 1.10 [0.87,1.38]
Discovery (m1) 1.20 [1.08,1.32]
MDC 1.38 [1.26,1.50]
TAG 46:2;0
DCS (disc) 1.12 [0.92,1.37]
GoDARTS (disc) 1.24 [1.10,1.41]
ANDIS (disc) 1.13 [0.90,1.41]
Discovery (m1) 1.19 [1.08,1.31]
MDC 1.23 [1.13,1.35]
TAG 48:2;0
DCS (disc) 1.12 [0.90,1.39]
GoDARTS (disc) 1.22 [1.07,1.39]
ANDIS (disc) 1.08 [0.86,1.36]
Discovery (m1) 1.17 [1.06,1.29]
MDC 1.26 [1.16,1.38]
TAG 48:3;0
DCS (disc) 1.11 [0.90,1.36]
GoDARTS (disc) 1.21 [1.07,1.38]
ANDIS (disc) 1.11 [0.88,1.38]
Discovery (m1) 1.17 [1.06,1.29]
MDC 1.21 [1.10,1.32]
TAG 51:1;0
DCS (disc) 1.19 [0.97,1.46]
GoDARTS (disc) 1.19 [1.05,1.35]
ANDIS (disc) 1.08 [0.86,1.35]
Discovery (m1) 1.17 [1.06,1.29]
MDC 1.33 [1.21,1.45]
TAG 49:1;0
DCS (disc) 1.17 [0.95,1.44]
GoDARTS (disc) 1.19 [1.05,1.35]
ANDIS (disc) 1.07 [0.85,1.35]
Discovery (m1) 1.17 [1.06,1.28]
MDC 1.28 [1.17,1.41]
SM 42:2;2
DCS (disc) 0.75 [0.60,0.93]
GoDARTS (disc) 0.89 [0.77,1.01]
ANDIS (disc) 0.88 [0.69,1.12]
Discovery (m1) 0.85 [0.77,0.94]
MDC 1.16 [1.06,1.27]
0.75 1.00 1.50 2.00
Hazard ratio
0.75 1.00 1.50 2.00
Hazard ratio

Fig. 2

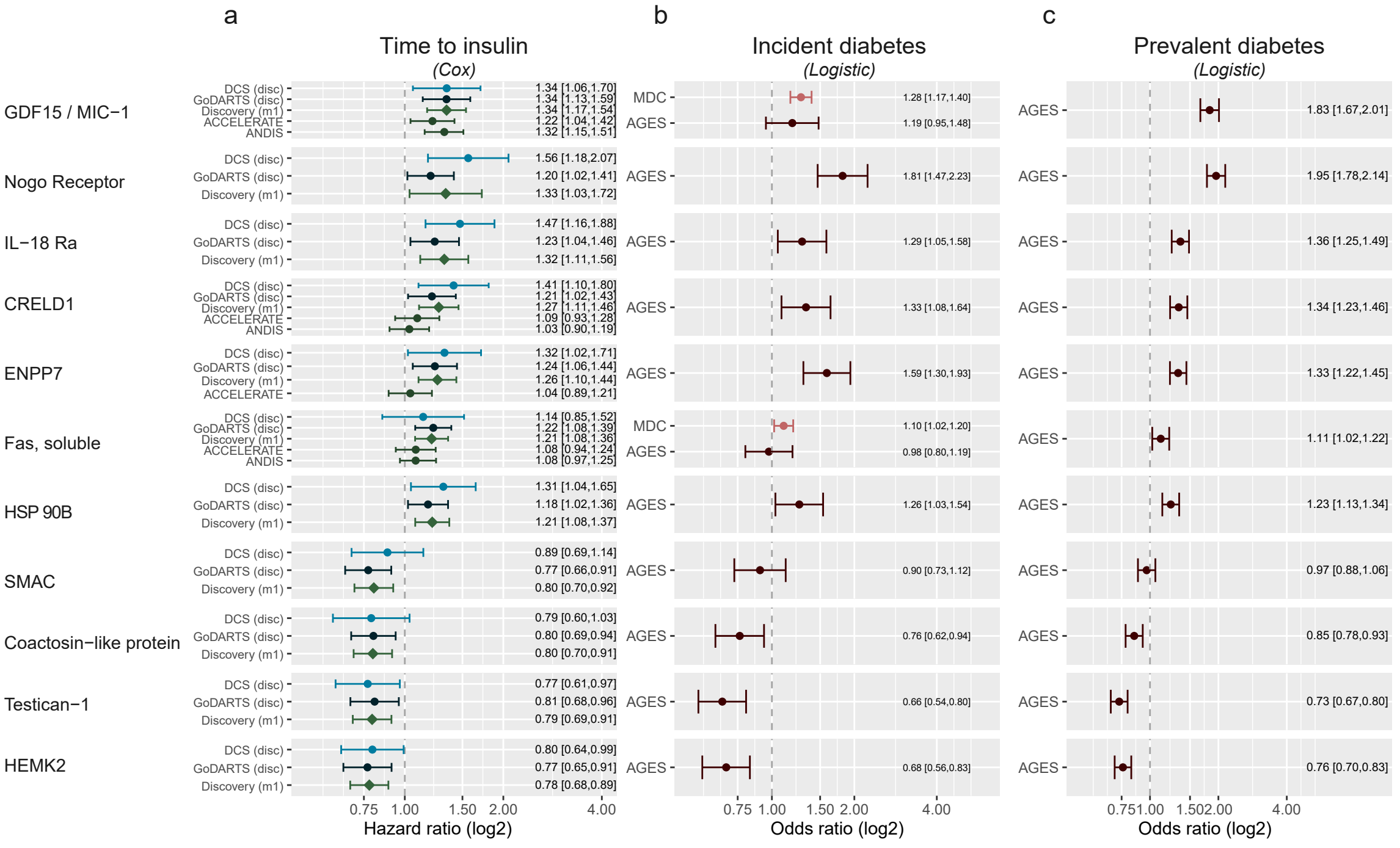

a
Time to insulin
(Cox)
b
Incident diabetes
(Logistic)
c
Prevalent diabetes
(Logistic)
GDF15 / MIC−1
DCS (disc) 1.34 [1.06,1.70]
GoDARTS (disc) 1.34 [1.13,1.59]
Discovery (m1) 1.34 [1.17,1.54]
ACCELERATE 1.22 [1.04,1.42]
ANDIS 1.32 [1.15,1.51]
MDC 1.28 [1.17,1.40]
AGES 1.19 [0.95,1.48]
AGES 1.83 [1.67,2.01]
Nogo Receptor
DCS (disc) 1.56 [1.18,2.07]
GoDARTS (disc) 1.20 [1.02,1.41]
Discovery (m1) 1.33 [1.03,1.72]
AGES 1.81 [1.47,2.23]
AGES 1.95 [1.78,2.14]
IL−18 Ra
DCS (disc) 1.47 [1.16,1.88]
GoDARTS (disc) 1.23 [1.04,1.46]
Discovery (m1) 1.32 [1.11,1.56]
AGES 1.29 [1.05,1.58]
AGES 1.36 [1.25,1.49]
CRELD1
DCS (disc) 1.41 [1.10,1.80]
GoDARTS (disc) 1.21 [1.02,1.43]
Discovery (m1) 1.27 [1.11,1.46]
ACCELERATE 1.09 [0.93,1.28]
ANDIS 1.03 [0.90,1.19]
AGES 1.33 [1.08,1.64]
AGES 1.34 [1.23,1.46]
ENPP7
DCS (disc) 1.32 [1.02,1.71]
GoDARTS (disc) 1.24 [1.06,1.44]
Discovery (m1) 1.26 [1.10,1.44]
ACCELERATE 1.04 [0.89,1.21]
AGES 1.59 [1.30,1.93]
AGES 1.33 [1.22,1.45]
Fas, soluble
DCS (disc) 1.14 [0.85,1.52]
GoDARTS (disc) 1.22 [1.08,1.39]
Discovery (m1) 1.21 [1.08,1.36]
ACCELERATE 1.08 [0.94,1.24]
ANDIS 1.08 [0.97,1.25]
MDC 1.10 [1.02,1.20]
AGES 0.98 [0.80,1.19]
AGES 1.11 [1.02,1.22]
HSP 90B
DCS (disc) 1.31 [1.04,1.65]
GoDARTS (disc) 1.18 [1.02,1.36]
Discovery (m1) 1.21 [1.08,1.37]
AGES 1.26 [1.03,1.54]
AGES 1.23 [1.13,1.34]
SMAC
DCS (disc) 0.89 [0.69,1.14]
GoDARTS (disc) 0.77 [0.66,0.91]
Discovery (m1) 0.80 [0.70,0.92]
AGES 0.90 [0.73,1.12]
AGES 0.97 [0.88,1.06]
Coactosin−like protein
DCS (disc) 0.79 [0.60,1.03]
GoDARTS (disc) 0.80 [0.69,0.94]
Discovery (m1) 0.80 [0.70,0.91]
AGES 0.76 [0.62,0.94]
AGES 0.85 [0.78,0.93]
Testican−1
DCS (disc) 0.77 [0.61,0.97]
GoDARTS (disc) 0.81 [0.68,0.96]
Discovery (m1) 0.79 [0.69,0.91]
AGES 0.66 [0.54,0.80]
AGES 0.73 [0.67,0.80]
HEMK2
DCS (disc) 0.80 [0.64,0.99]
GoDARTS (disc) 0.77 [0.65,0.91]
Discovery (m1) 0.78 [0.68,0.89]
AGES 0.68 [0.56,0.83]
AGES 0.76 [0.70,0.83]
0.75 1.00 1.50 2.00 4.00
Hazard ratio (log2)
0.75 1.00 1.50 2.00 4.00
Odds ratio (log2)
0.75 1.00 1.50 2.00 4.00
Odds ratio (log2)

Fig. 3

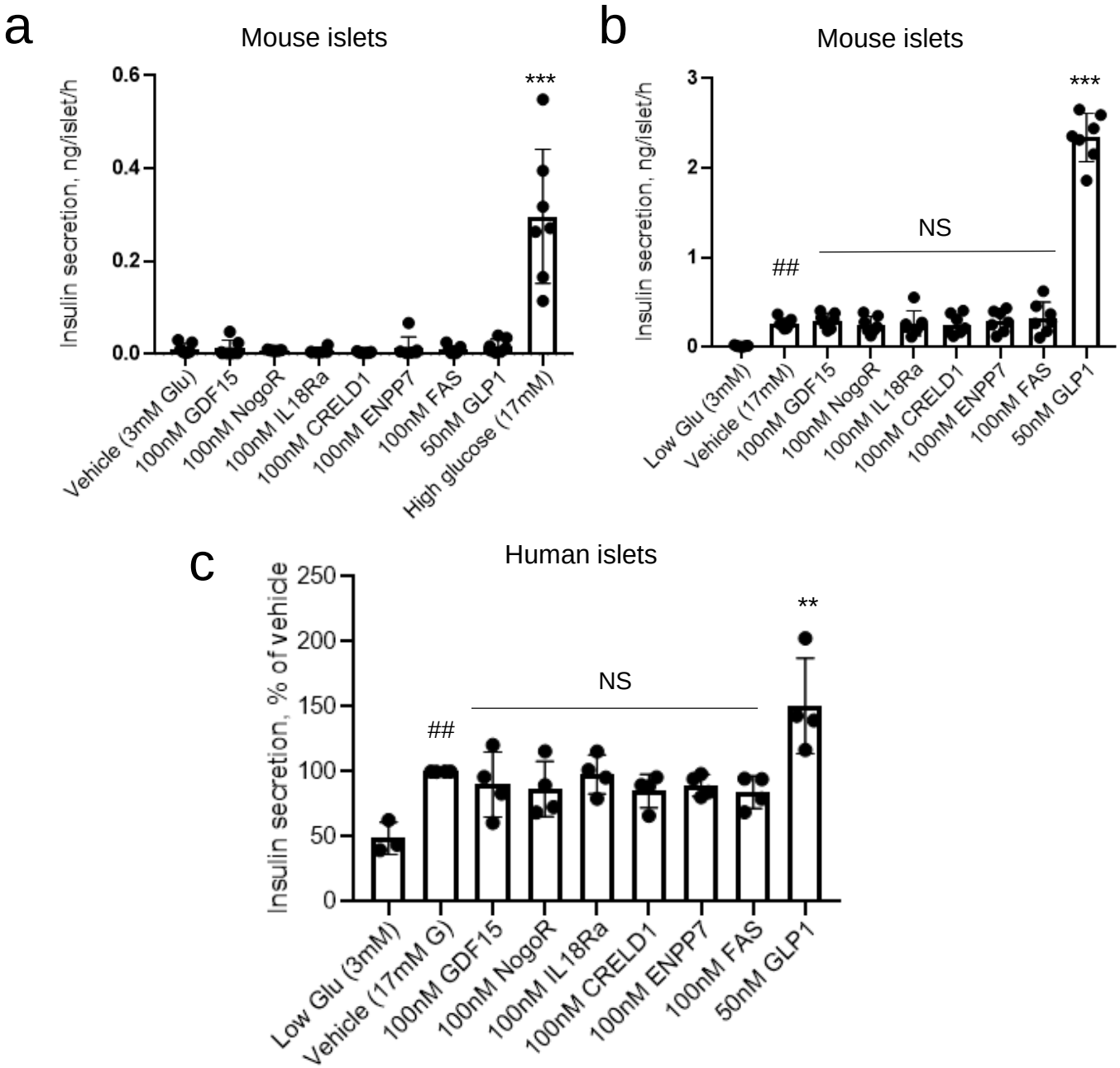

a
Mouse islets
Insulin secretion, ng/islet/h
0.6
0.4
0.2
0.0
***
Vehicle (3mM Glu)
100nM GDF15
100nM NogoR
100nM IL18Ra
100nM CRELD1
100nM ENPP7
100nM FAS
50nM GLP1
High glucose (17mM)
b
Mouse islets
Insulin secretion, ng/islet/h
3
2
1
0
***
##
NS
Low Glu (3mM)
Vehicle (17mM)
100nM GDF15
100nM NogoR
100nM IL18Ra
100nM CRELD1
100nM ENPP7
100nM FAS
50nM GLP1
c
Human islets
Insulin secretion, % of vehicle
250
200
150
100
50
0
**
##
NS
Low Glu (3mM)
Vehicle (17mM G)
100nM GDF15
100nM NogoR
100nM IL18Ra
100nM CRELD1
100nM ENPP7
100nM FAS
50nM GLP1

Fig. 4

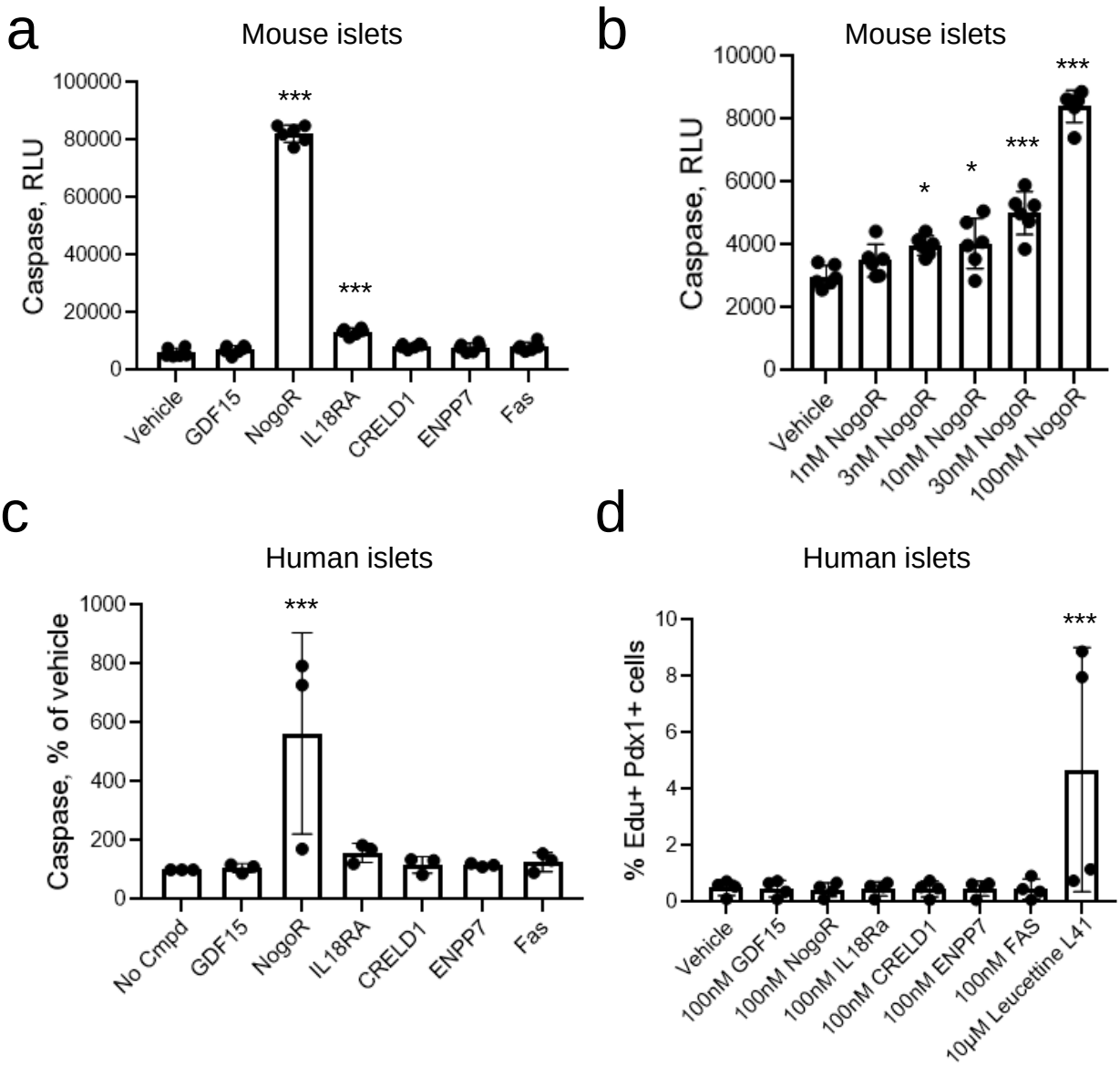
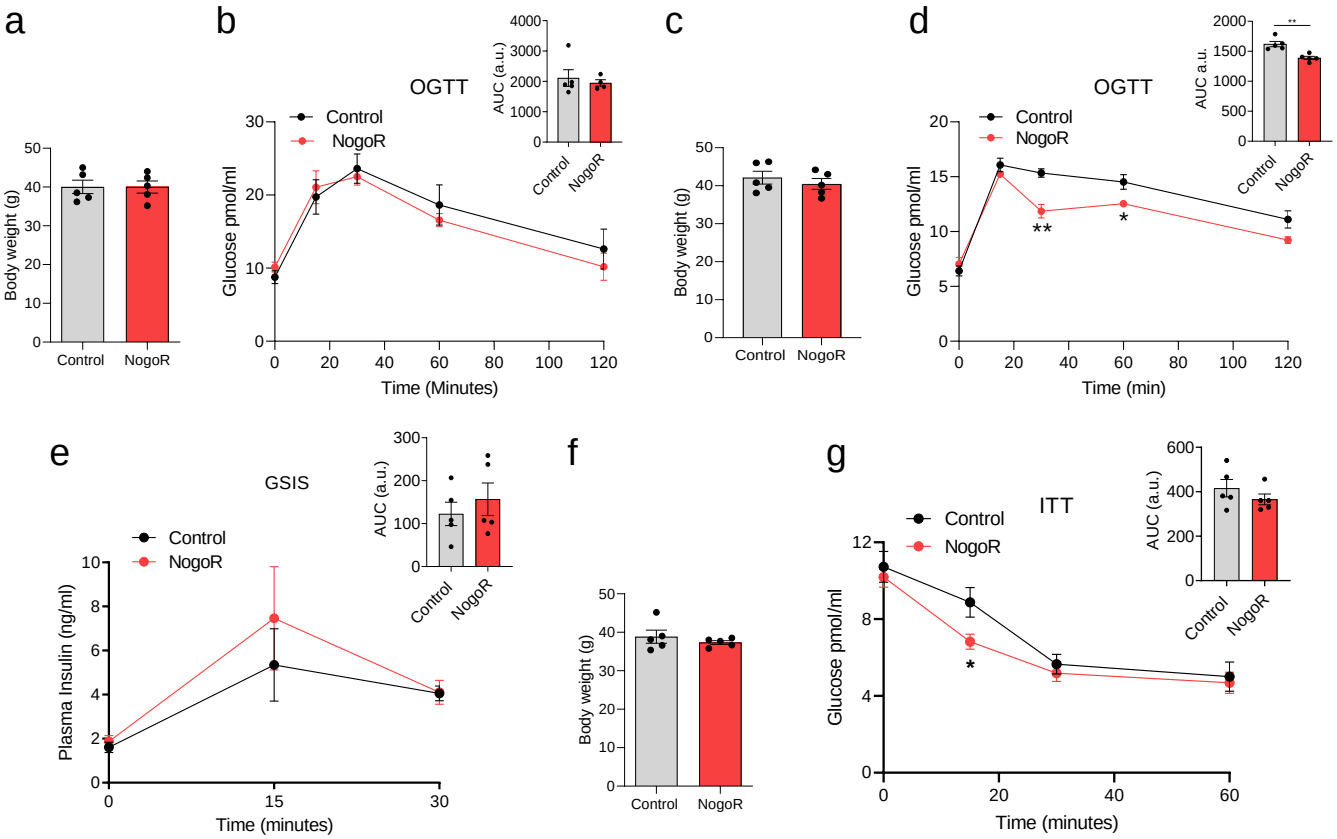

a
Mouse islets
Caspase, RLU
100000
80000
60000
40000
20000
0
***
***
Vehicle
GDF15
NogoR
IL18RA
CRELD1
ENPP7
Fas
b
Mouse islets
Caspase, RLU
10000
8000
6000
4000
2000
0
*
*
***
***
Vehicle
1nM NogoR
3nM NogoR
10nM NogoR
30nM NogoR
100nM NogoR
c
Human islets
Caspase, % of vehicle
1000
800
600
400
200
0
***
No Cmpd
GDF15
NogoR
IL18RA
CRELD1
ENPP7
Fas
d
Human islets
% Edu+ Pdx1+ cells
10
8
6
4
2
0
***
Vehicle
100nM GDF15
100nM NogoR
100nM IL18Ra
100nM CRELD1
100nM ENPP7
100nM FAS
10μM Leucettine L41

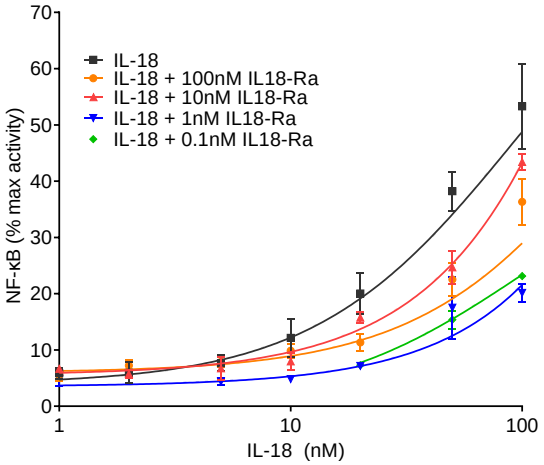
Fig. 5


a
Body weight (g)
Control
NogoR
b
OGTT
Control
NogoR
Glucose pmol/ml
Time (Minutes)
AUC (a.u.)
c
Body weight (g)
Control
NogoR
d
OGTT
Control
NogoR
Glucose pmol/ml
Time (min)
AUC a.u.
**
*

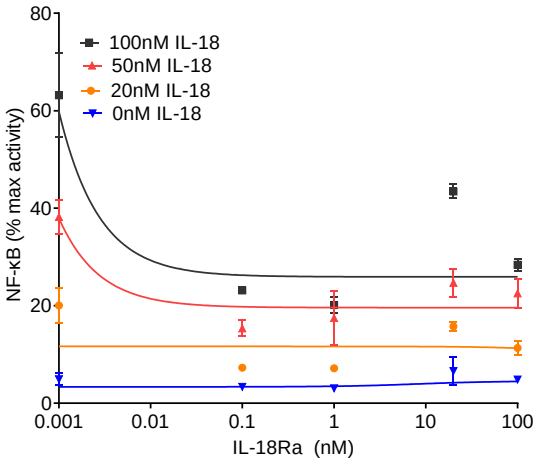
e
GSIS
Control
NogoR
Plasma Insulin (ng/ml)
Time (minutes)
AUC (a.u.)
f
Body weight (g)
Control
NogoR
g
ITT
Control
NogoR
Glucose pmol/ml
Time (minutes)
AUC (a.u.)
*

Fig. 6

a



b



	0nM IL18-Ra	100nM IL18-Ra	10nM IL18-Ra	1nM IL18-Ra	0.1nM IL18-Ra
EC50,nM	93.78	300.9	925.2	794445	102.4
logEC50	1.972	2.478	2.966	5.900	2.010

	100nM IL-18	50nM IL-18	20nM IL-18	0nM Il-18
IC50, nM	1.368	1.402	1.723	0.8448
logIC50	0.1361	0.1467	0.2363	-0.0732

Fig. 7