

RESEARCH ARTICLE

# Integrative network analysis highlights biological processes underlying GLP-1 stimulated insulin secretion: A DIRECT study

Valborg Gudmundsdottir<sup>1</sup>, Helle Krogh Pedersen<sup>1</sup>, Karla Viviani Allebrandt<sup>2</sup>, Caroline Brorsson<sup>1</sup>, Nienke van Leeuwen<sup>3</sup>, Karina Banasik<sup>4,5</sup>, Anubha Mahajan<sup>6</sup>, Christopher J. Groves<sup>7</sup>, Martijn van de Bunt<sup>6,7</sup>, Adem Y. Dawed<sup>8</sup>, Andreas Fritsche<sup>9</sup>, Harald Staiger<sup>10</sup>, Annemarie M. C. Simonis-Bik<sup>11</sup>, Joris Deelen<sup>12,13</sup>, Mark H. H. Kramer<sup>11</sup>, Axel Dietrich<sup>2</sup>, Thomas Hübschle<sup>14</sup>, Gonneke Willemsen<sup>15</sup>, Hans-Ulrich Häring<sup>9</sup>, Eco J. C. de Geus<sup>15,16</sup>, Dorret I. Boomsma<sup>15</sup>, Elisabeth M. W. Eekhoff<sup>11</sup>, Jorge Ferrer<sup>17,18,19</sup>, Mark I. McCarthy<sup>5,6,7</sup>, Ewan R. Pearson<sup>8</sup>, Ramneek Gupta<sup>1</sup>, Søren Brunak<sup>1,4</sup>, Leen M. 't Hart<sup>3,12\*</sup>



**1** Department of Bio and Health Informatics, Technical University of Denmark, Kongens Lyngby, Denmark, **2** Department of Translational Bioinformatics, R&D Operations, Sanofi-Aventis Deutschland GmbH, Industriepark Höchst, Frankfurt, Germany, **3** Department of Molecular Cell Biology, Leiden University Medical Center, Leiden, The Netherlands, **4** Novo Nordisk Foundation Center for Protein Research, Disease Systems Biology, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark, **5** Wellcome Trust Center for Human Genetics, University of Oxford, Oxford, United Kingdom, **6** Oxford NIHR Biomedical Research Center, Oxford, United Kingdom, **7** Oxford Center for Diabetes Endocrinology and Metabolism, University of Oxford, Oxford, United Kingdom, **8** Division of Molecular & Clinical Medicine, School of Medicine, University of Dundee, Dundee, United Kingdom, **9** Department of Internal Medicine, Division of Endocrinology, Diabetology, Angiology, Nephrology and Clinical Chemistry, Eberhard Karls University Tübingen, Member of the German Centre for Diabetes Research (DZD), Tübingen, Germany, **10** Institute of Pharmaceutical Sciences, Department of Pharmacy and Biochemistry, Eberhard Karls University, Tübingen, Germany, **11** Department of Internal Medicine, Diabetes Center and Endocrinology, VU University Medical Center, Amsterdam, The Netherlands, **12** Section Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands, **13** Max Planck Institute for Biology of Ageing, Cologne, Germany, **14** Department GI Endocrinology, R&D Diabetes Division, Sanofi-Aventis Deutschland GmbH, Industriepark Höchst, Frankfurt, Germany, **15** Department of Biological Psychology, Vrije Universiteit and the EMGO Institute for Health and Care Research, VU University Medical Center, Amsterdam, The Netherlands, **16** Netherlands Consortium for Healthy Aging, Leiden, The Netherlands, **17** Section of Epigenomics and Disease, Department of Medicine, Imperial College London, London, United Kingdom, **18** Genomic Programming of Beta Cells Laboratory, Institut d'Investigacions Biomediques August Pi I Sunyer (IDIBAPS), Barcelona, Spain, **19** Centro de Investigación Biomédica en Red de Diabetes y Enfermedades Metabólicas Asociadas (CIBERDEM), Madrid, Spain

**OPEN ACCESS**

**Citation:** Gudmundsdottir V, Pedersen HK, Allebrandt KV, Brorsson C, van Leeuwen N, Banasik K, et al. (2018) Integrative network analysis highlights biological processes underlying GLP-1 stimulated insulin secretion: A DIRECT study. PLoS ONE 13(1): e0189886. <https://doi.org/10.1371/journal.pone.0189886>

**Editor:** Tao Cai, NIDCR/NIH, UNITED STATES

**Received:** March 17, 2017

**Accepted:** December 4, 2017

**Published:** January 2, 2018

**Copyright:** © 2018 Gudmundsdottir et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Individual genotypes and phenotypes cannot be released due to restrictions in the informed consent signed by the participants, which was collected in the pre-GWAS era and did not include potential transfer of the data to a repository outside of the control of the primary investigators. However, requests for data may be submitted to Michiel Verburgh, secretary for the NTR study ([ntr@psy.vu.nl](mailto:ntr@psy.vu.nl)). Raw mouse expression data has been deposited in the ArrayExpress database (accession number E-MTAB-6015). In addition, summary statistics (P-

\* [lmhart@lumc.nl](mailto:lmhart@lumc.nl)

## Abstract

Glucagon-like peptide 1 (GLP-1) stimulated insulin secretion has a considerable heritable component as estimated from twin studies, yet few genetic variants influencing this phenotype have been identified. We performed the first genome-wide association study (GWAS) of GLP-1 stimulated insulin secretion in non-diabetic individuals from the Netherlands Twin register (n = 126). This GWAS was enhanced using a tissue-specific protein-protein interaction network approach. We identified a beta-cell protein-protein interaction module that was significantly enriched for low gene scores based on the GWAS P-values and found support at the network level in an independent cohort from Tübingen, Germany (n = 100). Additionally, a polygenic risk score based on SNPs prioritized from the network was associated (P < 0.05) with glucose-stimulated insulin secretion phenotypes in up to 5,318 individuals in

values, betas, and covariates for all SNPs that were tested) for the NTR cohort are available for download from the figshare repository ([10.6084/m9.figshare.5683975](https://doi.org/10.6084/m9.figshare.5683975)).

**Funding:** The work leading to this publication has received support from the Innovative Medicines Initiative Joint Undertaking under grant agreement n°115317 (DIRECT), resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution. Sanofi (an EFPIA member) provided support in the form of salaries for authors [KA, TH, and AD] and funding of the experimentation conducted in mice. Beyond the contribution of these authors, Sanofi did not have any additional role in the study design, data collection and analysis that pertain to this manuscript, as well as on the decision to publish, or preparation of the manuscript. The specific roles of these authors are articulated in the 'author contributions' section. The study received funding from the Netherlands Organization for Health Research and Development (Priority Medicines Elderly Program 113102006); by BBMRI-NL, a Research Infrastructure financed by the Dutch government (NWO 184.021.007). The Netherlands Twin Register acknowledges funding from the Netherlands Organization for Scientific Research (NWO: Netherlands Twin Registry Repository: researching the interplay between genome and environment). The department of internal medicine section Endocrinology and Diabetic Center VUmc acknowledges funding from ZonMw.

**Competing interests:** IMI DIRECT has received funding from EFPIA members, including Sanofi. KA, TH, and AD are employees of Sanofi. Sanofi contributed to this manuscript with experimental data on mice adipose tissue gene expression alterations in response to a GLP-1 agonist treatment. With this experiment Sanofi strived to understand peripheral effects of GLP-1 agonists. This does not alter our adherence to PLOS ONE policies on sharing data and materials.

MAGIC cohorts. The network contains both known and novel genes in the context of insulin secretion and is enriched for members of the focal adhesion, extracellular-matrix receptor interaction, actin cytoskeleton regulation, Rap1 and PI3K-Akt signaling pathways. Adipose tissue is, like the beta-cell, one of the target tissues of GLP-1 and we thus hypothesized that similar networks might be functional in both tissues. In order to verify peripheral effects of GLP-1 stimulation, we compared the transcriptome profiling of ob/ob mice treated with liraglutide, a clinically used GLP-1 receptor agonist, versus baseline controls. Some of the upstream regulators of differentially expressed genes in the white adipose tissue of ob/ob mice were also detected in the human beta-cell network of genes associated with GLP-1 stimulated insulin secretion. The findings provide biological insight into the mechanisms through which the effects of GLP-1 may be modulated and highlight a potential role of the beta-cell expressed genes *RYR2*, *GDI2*, *KIAA0232*, *COL4A1* and *COL4A2* in GLP-1 stimulated insulin secretion.

## Introduction

Glucagon-like peptide-1 (GLP-1) receptor agonists and DPP4-inhibitors are increasingly used therapeutic agents for type 2 diabetes, as they stimulate insulin secretion from the pancreatic beta-cells by potentiating glucose-dependent insulin secretion. In addition to the effects on the pancreas these drugs also operate via effects on other tissues. For instance, liraglutide, a clinically used GLP-1 receptor agonist, was shown to have beneficial effects on cardiovascular outcome and body weight loss [1]. However, the response to these drugs varies considerably between individuals. A large part of this variability is expected to be explained by underlying genetic differences as GLP-1 stimulated insulin secretion has an estimated heritability of 0.53 (95% CI, 0.33–0.70) [2]. Identification of these genetic determinants may aid patient stratification with regard to treatment response and shed light on the differential properties of the complex signaling networks controlling GLP-1 stimulated insulin secretion, which to date are not well understood. Previous studies have used targeted genotyping approaches to identify variants associated with GLP-1 stimulated insulin secretion, which mostly focused on GWAS loci for type 2 diabetes or related traits. Among the loci nominally associated with GLP-1 stimulated insulin secretion are variants in the *TCF7L2* [3], *GLP1R* [4], *WFS1* [5] and *CTRB1/2* loci [6] (all  $P < 0.05$ ), which highlights the potential of further genetic studies of GLP-1 stimulated insulin secretion.

Genome-wide association studies (GWAS) have successfully been used to identify genetic variants underlying complex phenotypes but for disease case-control status require large sample sizes to reveal variants with modest or small effect sizes. However, the use of more proximal phenotypes may reduce sample size requirements and furthermore, such analyses can be enhanced using integrative network approaches [7], by integrating genetic information with complementary data types such as tissue-specific gene expression and protein-protein interaction (PPI) data [8,9]. The aim of the current study was to provide insight into the biological mechanisms underlying GLP-1 stimulated insulin secretion using an untargeted integrative genomics approach. We performed a GWAS on 126 nondiabetic individuals from the Netherlands Twin Register (NTR) who underwent a hyperglycemic clamp [10], and the association analysis was augmented with a beta-cell specific PPI network analysis. We identified a set of genes that contained variants associated with GLP-1 stimulated insulin secretion, which at the

same time have the potential to physically interact in the beta-cell and are enriched for pathways important for insulin secretion. We carried out validation studies to assess the importance of the prioritized GLP-1 response subnetwork through associations with: i) GLP-1 stimulated insulin secretion in an independent collection of 100 unrelated individuals from Tübingen, Germany [10], ii) glucose stimulated insulin secretion phenotypes in up to 5,318 individuals from MAGIC [11] and iii) gene expression alterations in white adipose tissue as a response to liraglutide (a GLP-1 receptor agonist) treatment in ob/ob mice (a mouse model of obesity).

## Results

### GWAS and PPI network analysis

Clinical characteristics for the NTR cohort are shown in Table 1. No single nucleotide polymorphism (SNP) association reached genome-wide significance in the NTR cohort association analysis adjusted for age, gender, BMI, glucose tolerance status and insulin sensitivity (S1 and S2 Figs), while six independent signals were identified with  $P < 1.0 \times 10^{-5}$  (S1 Table). Using the integrative analysis workflow described in Fig 1, we next sought to identify significant signals at a cellular network-level. A more detailed analysis flowchart is shown in S3 Fig. As tissue-specific PPI networks have previously been shown to perform better for gene prioritization than global networks [12], we mapped gene significance values for GLP-1 stimulated insulin secretion onto a PPI network containing 8,457 genes that are expressed in pancreatic beta-cells [13] (see Methods). We then identified modules in the network that were enriched for genes with the strongest significance values using the jActiveModules algorithm [14] (see Methods for details). The top ranked network module contained 179 genes and had a z-score of 10.11, which was significantly higher than the z-scores of modules obtained by permuted gene scores (S4 Fig).

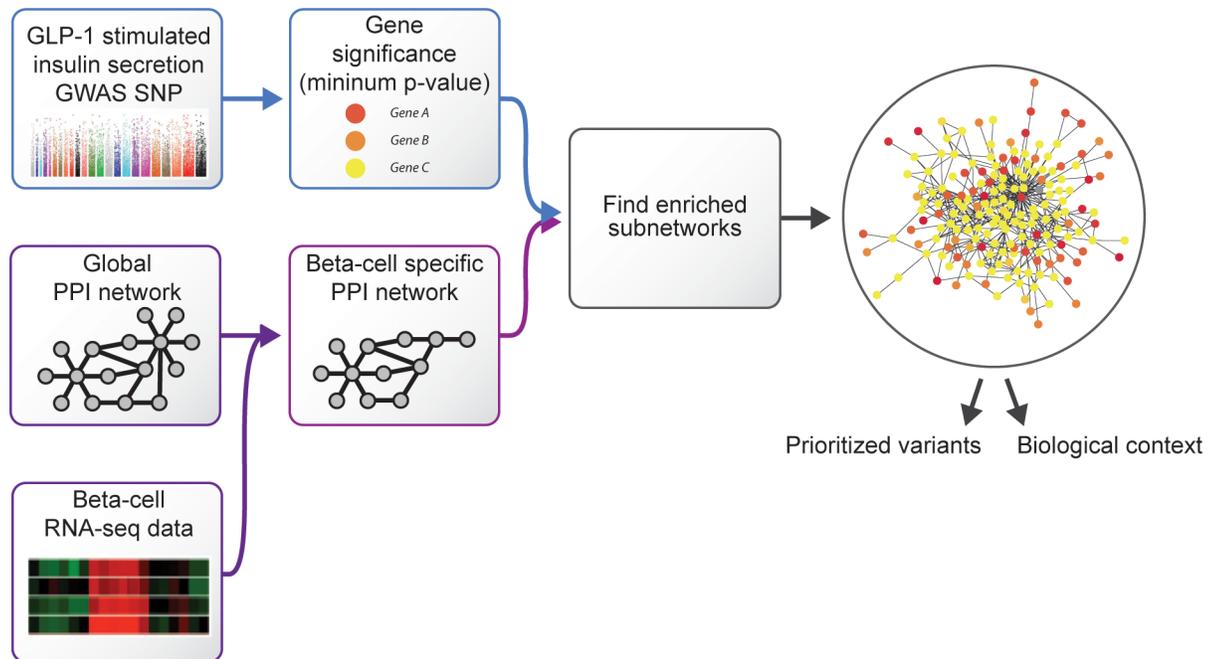
To focus on the most important part of this module, we reran the jActiveModules algorithm and created a consensus network from genes occurring in more than one of the top 15 second order modules (see Methods). The resulting consensus network contained 53 genes, whose significance was driven by 51 SNPs (Fig 2A, S2 Table). It contained genes already known to be involved in GLP-1 stimulated insulin secretion (*WFS1* [5], *RYR2* [15], *RAP1A* [16]), glucose stimulated insulin secretion (*VAV2* [17]), mediating the effects of GLP-1 on beta-cell mass (*FOXO1* [18]) and genes implicated in type 2 diabetes through GWAS of Han Chinese [19] and Mexican [20] populations (*PTPRD*), a gene expression-based genome-wide association study [21] (*CD44*) and a linkage study in an African American population [22], (*MAGI2* and *CTNNA2*).

**Table 1. Clinical characteristics of the study groups.**

	NTR cohort	German cohort
n (NGT/IGT)	120/6	68/32
Age (years)	31.5 ± 6.3	39.7 ± 12.8
Gender (M/F, n)	60/66	44/56
BMI (kg/m <sup>2</sup> )	24.1 ± 3.5	25.8 ± 5.5
Fasting glucose (mmol/l)	4.6 ± 0.4	5.2 ± 0.7
2-hr glucose (mmol/l)	5.4 ± 1.2	6.6 ± 2.1
Fasting insulin (pmol/l)	35 (27–52)	47 (32–67)

Data are means ± SD; median (interquartile range) or number (n).

<https://doi.org/10.1371/journal.pone.0189886.t001>



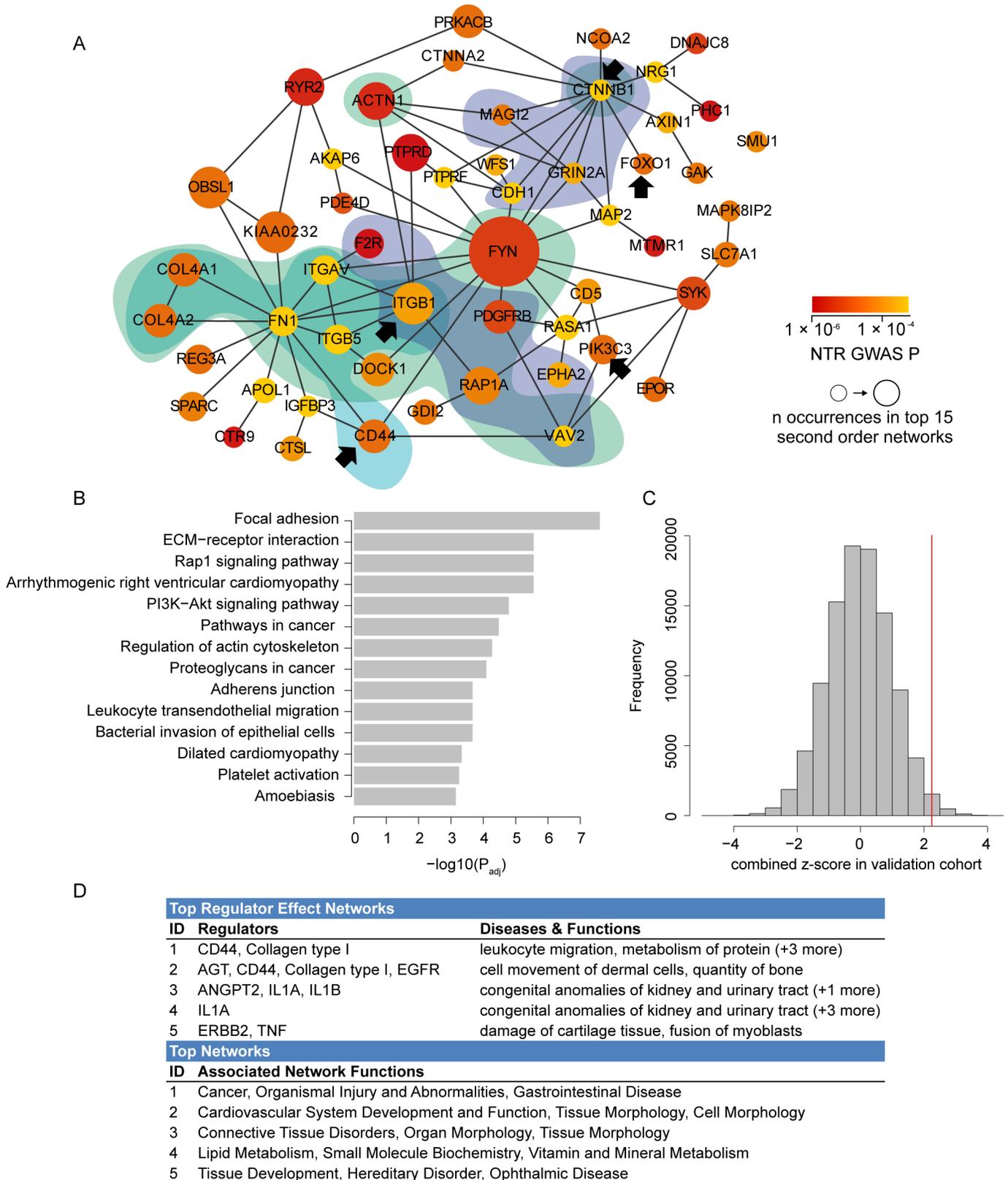
**Fig 1. Integrative network analysis workflow overview.** GLP-1 stimulated insulin secretion GWAS SNP P-values were converted to gene significance scores, which were then mapped onto a beta-cell specific PPI network created by pruning the global network using beta-cell gene expression data. The jActiveModules algorithm was used to identify network modules that were enriched for association signal. The top scoring network modules were used to prioritize genetic variants and explore the biological context of the genetic associations.

<https://doi.org/10.1371/journal.pone.0189886.g001>

### Functional investigation of the GLP-1 response consensus network

By definition, all of the genes in the consensus network are expressed in beta-cells, but four (*CTNNA2*, *RYR2*, *GRIN2A*, *NRG1*) have additionally been described as particularly enriched in beta-cells compared to non-beta-cell islet components [13]. The consensus network was significantly enriched for gene ontology (GO) terms related to plasma membrane and signaling transduction (S5 Fig) and enriched KEGG pathways included focal adhesion, extracellular matrix-receptor interaction, the Rap1 and PI3K-Akt signaling pathways and regulation of actin cytoskeleton (Fig 2B).

We investigated if the SNPs driving the significance of the network acted as expression quantitative trait loci (eQTL) for their corresponding genes in the network. We found seven potential eQTL pairs (unadjusted  $P < 0.01$ ) in human islets ( $n = 118$ ), involving the genes *CTR9*, *RYR2*, *PRKACB*, *DOCK1*, *APOL1*, *ITGB5* and *MAP2* (S3 Table). In addition, we found two potential eQTL pairs for *F2R* and *CD5* in pancreas, and three eQTL pairs in blood samples for *KIAA0232*, *SPARC* and *RAP1A* (S3 Table). We also investigated the overlap of network loci with pancreatic islet regulatory elements and found 15 loci to overlap islet promoter or enhancer regions [23] (S3 Table). Of those, nine loci overlapped clusters of active enhancers, but such clusters are enriched for variants associated with type 2 diabetes and fasting glycemia [23], This was a higher fraction than was observed for the top SNPs from a GWAS performed by MAGIC investigators of corrected insulin response, but lower than for genome-wide significant SNPs for fasting glucose and type 2 diabetes (S6 Fig). Five of the consensus network SNPs (mapping to *APOL1*, *ITGB5*, *MAP2*, *CD5* and *KIAA0232*) reside in loci overlapping active enhancers or clusters of active enhancers in islets and were a part of a potential eQTL pair.



**Fig 2. Results from network analysis of GLP-1 stimulated insulin secretion GWAS.** A) The beta-cell specific GLP-1 response consensus network, annotated with the top enriched KEGG pathways: Focal adhesion (green), ECM-receptor interaction (blue) and Rap1 signaling (purple). Arrows indicate genes that were

identified as upstream regulators of differentially expressed genes in the transcriptome analyses of the liraglutide treated mice versus baseline controls. B) The KEGG pathways enriched (BH adjusted  $P$ -value  $< 1 \times 10^{-3}$ ) within the GLP-1 response consensus network, compared to the whole beta-cell PPI network. C) The red line denotes the combined z-score in the Tübingen validation cohort for 28 consensus network SNPs with discovery GWAS  $P < 5 \times 10^{-4}$  compared to 100,000 z-scores obtained from randomly selected sets of SNPs from the beta-cell network (histogram), empirical  $P$ -value = 0.012. D) Top panel: Top regulators for networks of differentially expressed genes in the liraglutide treated mice transcriptome experiment. Bottom panel: Prioritized network modules from human and mouse experiments map to connective tissue and focal adhesion related pathways.

<https://doi.org/10.1371/journal.pone.0189886.g002>

## Validation of the GLP-1 response consensus network

In order to validate the genetic associations driving the network enrichment we calculated a combined z-score for the consensus network SNPs in an independent dataset of 100 unrelated individuals from Tübingen, Germany who underwent a hyperglycemic clamp procedure similar to the NTR cohort (clinical characteristics are shown in [Table 1](#)). We restricted all validation attempts to the 31 SNPs (28 of which passed quality control in the Tübingen data) with a discovery GWAS  $P < 5.0 \times 10^{-4}$ . The combined z-score based on these 28 SNPs was significantly higher ( $P = 0.01$ ) than those obtained from 100,000 randomly selected sets of SNPs from the beta-cell network ([Fig 2C](#)). However, this set of SNPs was not enriched for directional consistency in the validation dataset (14/28 SNPs directionally consistent,  $P = 0.57$ ).

Finally, we investigated if the network analysis prioritization of SNPs has an additional value over the individual SNPs with the lowest  $P$ -values from the GWAS analysis. We therefore calculated a similar combined z-score from the top 31 ranked independent GWAS SNPs in the NTR discovery cohort but, in contrast to the network-based SNPs, this z-score was not higher in the validation dataset than expected by chance ( $P = 0.95$ ) ([S7 Fig](#)).

At the single SNP level, two (rs7669558 and rs72509) of the 31 consensus network SNPs with a discovery GWAS  $P < 5.0 \times 10^{-4}$  had a  $P$ -value  $< 0.05$  in the validation dataset ([S2 Table](#)) but none were significant after Bonferroni correction for multiple testing. We performed a meta-analysis of the discovery and validation dataset ([S2 Table](#)), where none of the consensus network SNPs reached a genome-wide significance but four had a meta-analysis  $P < 5.0 \times 10^{-4}$  (*KIAA0232* rs7669558:  $P = 5.9 \times 10^{-5}$ , *COL4A1/COL4A2* rs72509:  $P = 7.0 \times 10^{-5}$ , *RYR2* rs6429033:  $P = 9.4 \times 10^{-5}$  and *GDI2* rs871748:  $P = 1.5 \times 10^{-4}$ ).

Of the 31 consensus network SNPs, 18 ([S3 Table](#)) had available information on both insulin secretion and action indices published by MAGIC (see [Methods](#) and [S4 Table](#) for detailed overview of phenotypes). Rs871748 (*GDI2*) was nominally associated with four measures of glucose stimulated insulin secretion in the MAGIC data, with a consistent direction of effect compared to the GLP-1 stimulated insulin secretion in the NTR cohort ([S5 Table](#)). Furthermore, a weighted polygenic risk score (PRS) made from the 18 SNPs common between the two datasets showed a nominally significant association ( $P < 0.05$ ) with the oral glucose tolerance test derived variables area under the curve (AUC) for insulin, AUC for insulin/AUC for glucose, insulin at 30 minutes and insulin sensitivity index ([S6 Table](#)). In contrast, no significant associations were observed when the same number of SNPs was selected based on their discovery GWAS  $P$ -value alone ([S7 Table](#)).

Finally, an independent gene set enrichment analysis was conducted from the transcriptome profiling of white adipose tissue from mice treated with liraglutide (an analogue of GLP-1) versus baseline controls (see [S8 Fig](#) and [Methods](#) for experimental details). Adipose tissue is, like the beta-cell, one of the target tissues of GLP-1 [24] and we thus hypothesized that similar networks might be functional in both tissues. We identified upstream regulators (see [Methods](#)) for the differentially expressed genes ([Fig 2D](#), [S8 Table](#)). Interestingly, five upstream regulators predicted to regulate genes in the transcriptome dataset (*CD44*, *FOXO1*, *ITGB1*, *CTNNB1*, and the PI3K complex—which *PIK3CA* is a member of) were also present in the human beta-cell GLP-1 response consensus network (as highlighted in [Fig 2A](#)) and the PI3K signaling pathway

was furthermore among its top enriched pathways (Fig 2B). Additionally, a member of the collagen Type IV family of genes (*COL4A3*, related to *COL4A1* and *COL4A2*) ranked as a top upstream regulator. *COL4A3* interacts with both *COL4A1* and *COL4A2* and other genes (*ITGAV*, *FN1*, *ITGB1* and *ITGB5*) that appear in the human consensus network, and which connect the collagen pathway to the GLP-1 receptor (S9). *CD44* is predicted to be an upstream regulator of collagen genes and *CTNNB1*, further illustrating how these genes are connected (S9 Fig). Finally, the genes prioritized for GLP-1 response in the human and mouse study were found to be highly connected in tissue-specific functional networks from GIANT [25] for both pancreatic islet and adipose tissue (S10 Fig).

## Discussion

In the present study we describe the first reported GWAS of GLP-1 stimulated insulin secretion. As genetic variants underlying complex phenotypes and diseases are expected to collectively perturb functional modules within the cellular machinery [26], we used the uniquely phenotyped NTR cohort to explore the underlying mechanisms of GLP-1 stimulated insulin secretion using a network analysis approach. While a previous study on the NTR cohort using MetaboChip genotyping identified three strong signals for GLP-1 stimulated insulin secretion [6], illustrating the potential for the discovery in this limited sample size, our genome-wide analysis did not reveal any additional loci associated at a genome-wide significance. Instead, we identified a module within a beta-cell specific PPI network that was significantly enriched for gene scores derived from the GWAS. We found support for the network level association in an independent dataset from Germany and a PRS constructed from SNPs selected from the network showed an association with glucose-stimulated insulin secretion phenotypes in the MAGIC consortium. The same support was not observed for a matched number of SNPs selected by the discovery GWAS *P*-value alone, suggesting that the network prioritization approach to some extent enhanced the GWAS findings. While GLP-1 agonist expression response could not be investigated in pancreatic islets, GLP-1 agonism is known to stimulate brown adipose tissue thermogenesis and browning through hypothalamic AMPK [27]. Thereby, we hypothesized that components of the GLP-1 response beta-cell subnetwork might be functional in both tissues. Our findings provide biological insight into the common mechanisms through which the effects of GLP-1 may be modulated in these tissues.

A few genes from the consensus network were highlighted by additional support from the meta-analysis of the discovery and validation dataset. The four SNPs with the lowest meta-analysis *P*-values had been assigned to the genes *GDI2*, *RYR2*, *KIAA0232*, *COL4A1* and *COL4A2*. Of those, the rs871748 variant (*GDI2*) was in addition found to be nominally associated with four insulin secretion phenotypes in the MAGIC data. The gene product of *GDI2* is a GDP dissociation inhibitor, which is involved in vesicular trafficking between cellular organelles by regulating GDP-GTP exchange reactions of Rab proteins. In the GLP-1 response consensus network, *GDI2* interacts with *RAP1A*, which encodes the Rap1 protein. Rap1 has been shown to be essential for cAMP mediated potentiation of glucose stimulated insulin secretion, such as the one stimulated by GLP-1, through its activation by Epac2 [16]. The importance of the Rap1 signaling pathway was further highlighted by it being among the top enriched pathways in the consensus network. Interestingly, the Ryr2 channel encoded by *RYR2* is also activated by Epac2 and plays a role in GLP-1 stimulated insulin secretion by intracellular Ca<sup>2+</sup> mobilization [15].

The uncharacterized protein *KIAA0232* is another direct interaction partner of *RYR2* in the consensus network and is additionally supported as the causal gene for the associated rs7669558 variant by its strong eQTL association in blood. The rs7669558 SNP is located in a

cluster of active islet enhancers, making this protein an intriguing target for further characterization.

The two collagen IV genes, *COL4A1* and *COL4A2*, were supported by the meta-analysis *P*-value for the rs72509 variant and their role in GLP-1 response was additionally supported by evidence from the transcriptome experiment in mice treated with the GLP-1 receptor agonist liraglutide. These two genes are part of the focal adhesion and extracellular matrix-receptor interaction pathways, the two most strongly enriched pathways in the consensus network. The integrin mediated ligation of pancreatic beta-cells to collagen IV is known to promote the secretion of insulin [28] and focal adhesion is important for glucose stimulated insulin secretion [29], which is here highlighted in the context of GLP-1 stimulation. Both of these pathways, in addition to the regulation of actin cytoskeleton pathway, have been implicated in type 2 diabetes, as they were found to be overrepresented among genes differentially methylated in pancreatic islets of type 2 diabetes patients compared to non-diabetic controls [30]. Moreover, the *COL4A1* gene (along with two other consensus network members, *VAV2* and *ITGB5*) was found to be among the significantly hypomethylated genes. In addition, *CD44*, a widely expressed cell surface glycoprotein known to induce integrin-mediated adhesion, was the most significant up-regulator of the expression dataset for liraglutide response [31].

As liraglutide is a GLP-1 receptor agonist, liraglutide treatment-induced gene expression alterations can highlight networks and pathways related to GLP-1 response. In our transcriptomic experiment for liraglutide response in adipose tissue, we observed an overlap with the pathways prioritized by the GLP-1 stimulated insulin secretion consensus network in the pancreatic beta-cell. This suggests that some additional effects of GLP-1 in peripheral tissue could be mediated by genes that are not islet-specific, such as collagen type IV. Studies performed in isolated adipocytes have demonstrated that GLP-1 has the ability to induce both lipogenic and lipolytic mechanisms in white adipose tissue through activation of ERK, PKC and AKT signaling pathways [32]. These are also active in the islets, as transgene expression of the GLP-1R in the islets of *Glp1r*<sup>-/-</sup> mice restored GLP-1R dependent stimulation of cAMP and Akt phosphorylation and was sufficient for restoration of GLP-1 stimulated insulin secretion in perfused islets [33]. Thereby, the GLP-1 receptors in the islet have an essential physiological role in the regulation of beta-cell function and glucose homeostasis through the Akt pathway, which in parallel affects adipogenesis.

Remodeling in heart and vasculature is linked to alterations in extracellular matrix and integrin expression [34]. This relates to a recently reported clinically relevant downstream effect of GLP-1 stimulation, cardioprotection via the PI3K/Akt/Bad pathway [35], leading to stabilization of atherosclerosis in increase of plaque collagen content in arteriosclerotic mice [36]. Both human and animal model networks presented in the current study point to relevant signals for the PI 3-kinase signaling pathway, which may be important for both insulin secretion and diabetes comorbidities originating in other tissues, such as cardiovascular disease.

One of the main challenges in GWAS is to identify the causal genes mediating the effects of associated variants. Here, each variant was assigned to nearby genes but in addition, the most likely causal gene for each variant can be considered to be the one that also physically interacts with other candidate genes at the protein level in the context of the beta-cell, based on the network module prioritization. We found a number of the consensus network SNPs to either have a direct potential eQTL association with the corresponding gene either in human islets, pancreas or blood, or reside in loci overlapping clustered active islet enhancers that are known to be enriched for type 2 diabetes and fasting glucose associated loci [23], suggesting that many of the SNPs driving the consensus network significance have the potential to confer regulatory effects on gene expression.

A limitation of our study is clearly the limited sample size of the GWAS cohorts impacting statistical power. Nevertheless, the modified hyperglycemic clamp procedure has benefits in terms of producing a very detailed measure, which is more proximal to the phenotype compared to those based on for instance OGTT and thus likely to reveal larger effect sizes, which is a general observation for pharmacogenomic traits [37]. We employed the integrative network analysis to reduce the number of false positives, by prioritizing signals in vicinity of genes that have the potential to physically interact in the pancreatic beta-cell, and furthermore focused our attention on the genetic variants from the network with the best meta-analysis P-values from the two cohorts. The network and transcriptomic analyses provide biological hypotheses in the form of prioritized genes and pathways for future functional studies that will be required to confirm their role in GLP-1 stimulated insulin secretion and as such this study should be seen as an exploratory study.

In conclusion, we have identified a beta-cell PPI network module enriched for nominal associations with GLP-1 stimulated insulin secretion. This network module highlights genes and pathways already known to be of importance for insulin secretion, and indicates new potential target genes that operate in the same network context. The genetic variants prioritized through the network approach were collectively associated with insulin secretion capability in the general population and many overlap with islet-active regulatory regions, suggesting a possible influence on the gene expression of network members. Consistent with this hypothesis, alterations in gene expression in response to liraglutide treatment in mice showed that main network regulators are connected to genes nominally associated with the GLP-1 stimulated insulin response. Furthermore, the results demonstrate how data integration can highlight biological mechanisms underlying a phenotype where GWAS results on their own may be insufficient.

## Materials and methods

### Hyperglycemic clamp cohorts

GLP-1 stimulated insulin secretion was measured with a modified hyperglycemic clamp in 126 twins and sibs from the Netherlands twin register (NTR), aged 20–51 years [2]. This cohort consists of a mixed sample of twins and non-twin sibs recruited from 54 families (family size 2–9). In total, the NTR twin sample included 33 monozygotic twin pairs ( $n = 66$ ), 14 same sex dizygotic twin pairs ( $n = 28$ ) and 32 single twins and same sex sibs of the twins. The validation cohort consisted of 100 unrelated subjects, aged 18–68 years, from Tübingen, Germany [10] (68 with NGT/32 with IGT) who were examined with an identical hyperglycemic clamp [38]. The human studies were conducted according to the principles expressed in the Declaration of Helsinki. The medical ethics committee at VU University Amsterdam, the Netherlands, approved the NTR study protocol. The study protocol was approved by the ethical committee of the University of Tübingen, Germany. All participants gave written informed consent before the study was started.

### Hyperglycemic clamp procedure

Hyperglycemic clamp studies were performed in 2005–2007 and 1998–1999 for the NTR and Tübingen cohort respectively. All participants underwent a modified hyperglycemic clamp at 10 mmol/l glucose for three hours with additional GLP-1 and arginine stimulation during the last hour. After a priming infusion of glucose to acutely raise blood glucose levels, blood glucose levels were measured with a glucose analyzer and kept constant at 10 mmol/l during the whole clamp. Insulin levels were measured with immunoassays as previously described. Insulin sensitivity index was calculated as described previously. GLP-1-stimulated insulin release

was measured as the mean incremental area under the curve (160–180 min) following GLP-1 stimulation. Exact details of the modified hyperglycemic clamp procedure can be found in Simonis-Bik et al. [2].

### Genotyping and association analysis

Genotyping of the two cohorts (NTR and Tübingen) and subsequent data analysis was performed between September 2013 and September 2016. Genotyping using the HumanCoreExome chip was performed according to the manufacturer's protocol (Illumina Inc. San Diego, CA, USA). For quality control we used the following settings: a cut-off for the genotyping call rate of 99%, Gentrain and clusters score  $< 0.6$  and  $0.4$  respectively, and the  $P$ -value cut-off for Hardy-Weinberg equilibrium was set at  $10^{-4}$ . In total 513,444 SNPs passed quality control. Imputation up to the March 2012 1000 Genomes reference panel was done using SHAPEIT (v2.r644) and IMPUTE (v2.3.0). SNPs with a low frequency ( $MAF < 5\%$ ) and/or imputation quality ( $R_T^2 < 0.4$ ) were excluded, leaving 6.6 M SNPs for the analysis. The test statistics were not adjusted for inflation (population stratification) because of the low genomic inflation factor ( $\lambda = 1.02$ ). In order to account for the family relationships in the twin cohort we used QTassoc [39], a software tool based on SNPtest that is capable of handling familial data (using the kinship coefficients matrix) and genotype uncertainty. Data from both cohorts were analyzed using linear regression under an additive model and were adjusted for age, gender, glucose tolerance status, insulin sensitivity index, and familiarity (NTR only) as potential confounders. Fixed-effect meta-analysis (quantitative trait) of the two studies was performed using GWAMA [40] with double genomic control, i.e. for the results from the individual studies (-gc) and from the meta-analysis (-gco).

### Gene significance scoring

$P$ -values from the discovery GWAS analysis on GLP-1 stimulated insulin secretion were used to assign gene significance. Each gene was assigned the lowest  $P$ -value mapping to its boundaries, defined as 110kb upstream and 40kb downstream from transcription start/stop sites. These boundaries represent the 99<sup>th</sup> percentile of cis-eQTLs from their associated genes [41]. All SNPs not mapping to any predefined gene-window were excluded from the analysis.

### Construction of a beta-cell specific PPI network

A beta-cell specific PPI network was created by pruning the InWeb database [42] of high confidence physical PPIs (154,168 interactions between 12,778 proteins) using published beta-cell specific RNAseq data [13]. More specifically, genes with 25th percentile RPKM  $< 1$  were considered less likely to be expressed in the beta-cell and thus removed from the pruned beta-cell specific network. Special care was taken not to remove beta-cell or pancreatic transcription factors (S9 Table) or other lowly expressed beta-cell specific genes, such as the ones defined as beta-cell enriched in the study by Nica et al. [13].

### Identifying hotspots in the PPI network

The Cytoscape plugin jActiveModules [14] was used to identify modules in the network that were enriched for high scoring genes (that is with low SNP  $P$ -values). The jActiveModules algorithm is described in detail in the original publication [14] but in brief, each node representing a gene  $i$  in the network is assigned a z-score  $z_i = \Phi^{-1}(1 - p_i)$ , where  $\Phi^{-1}$  is the inverse normal cumulative distribution function and  $p_i$  is the gene significance score. An aggregated z-score is calculated for each module  $A$  with  $k$  nodes as the normalized sum of z-scores of all

genes in the module:

$$z_A = \sum_{i \in A} Z_i / \sqrt{k}$$

A module z-score  $> 3$  is generally considered significant, according to the jActiveModules authors. jActiveModules searches for top scoring modules within the full network, starting from each of the (in our case top 100 highest scoring) nodes in the beta-cell network and adding nodes to the network module using a greedy search algorithm.

The z-scores generated by jActiveModules are a measure of the enrichment significance of the modules compared to 100,000 permutations of randomly selected genes from the whole beta-cell PPI network. The z-scores were additionally compared to 10 sets of permuted gene significance values. As the top scoring module was too large ( $n = 179$  genes) for manual inspection, the jActiveModules algorithm was rerun on the top scoring module and the consequent submodules within it used to build a consensus network, where a node was included if it appeared more than once in any of the top 15 second order modules. The enrichment of association signals in the consensus network was validated in the Tübingen cohort by comparing the combined z-score for the network SNPs to those obtained from 100,000 randomly sampled sets of SNPs that had previously been assigned to each of the genes in the beta-cell network during the gene significance scoring step. For comparison we attempted the same validation for a matched number of the top independent signals from the discovery GWAS.

### Gene set enrichment of beta-cell specific PPI network module

Gene set enrichment analysis was performed using ConsensusPathDB [43], testing Gene Ontology (GO) and KEGG pathway gene sets and using all 8,457 genes present in the beta-cell PPI as background. Benjamini & Hochberg [44] adjusted  $P$ -values  $< 0.05$  were considered statistically significant.

### eQTL associations

We extracted eQTL associations for each of the SNP-gene pairs in the consensus network from the GTEx portal [45] and from human pancreatic islets [46]. We searched for eQTL associations in human islets ( $n_{\text{samples}} = 118$ ), pancreatic tissue ( $n_{\text{samples}} = 58$ ) and blood ( $n_{\text{samples}} = 168$ ). Due to the predefined SNP-gene pairs tested and the small number of samples, we considered SNPs with  $P < 0.01$  as potential eQTLs. In addition we looked up the network SNPs amongst significant eQTLs in whole blood from 5,311 individuals [47].

### Islet regulatory element overlap

Positions for six types of regulatory elements in pancreatic islets (promoters, inactive enhancers, active enhancers, clustered active enhancers, CTCF bound sites and other) were obtained from a recent study [23]. All SNPs in the consensus network and SNPs in high LD ( $r^2 > 0.8$ ) were tested for overlap with any of the regulatory regions.

### Polygenic risk scores

Weighted PRS were tested for association with oral glucose tolerance test phenotypes (S4 Table) from MAGIC [11]. The association testing was performed with the “gtx” R package. The analysis was limited to SNPs common to both datasets. The effect raising allele of the best SNP for each gene (based on its discovery GWAS  $P$ -value) was chosen for the PRS, and

weighted by the effect size in the discovery GWAS. For comparison, another PRS was created from a matched number of the top discovery GWAS SNPs (LD-pruned at  $r^2 < 0.8$ ).

### Liraglutide treatment in ob/ob mice

Twenty female B6.Cg-Lep ob/ob mice were obtained from Charles River (Sulzfeld, Germany) in an age of 6–8 weeks. Upon receipt, animals were housed 5 per cage in an air-conditioned, pathogen-free barrier facility maintained at  $22 \pm 2^\circ\text{C}$  with an air humidity of 45–65% and a 12-h dark–light cycle (lights on at 06:00 am). The mice had ad libitum access to a standard rodent diet from ssniff® (R/M-H, V1534-0, 10 mm pellets, Soest, Germany) and tap water during the entire experiment. The mice were surgically implanted in an age of 10–12 weeks with Alzet osmotic mini pumps (model 1002, 0,25µl/h, Cupertino, CA, USA) under isoflurane (2.5% using a flow of 0.6 L/min) inhalation anesthesia with peri-operative, subcutaneous, analgesic Buprenorphine treatment (0.05 mg/kg Temgesic®). Due to wound healing problems after the mini pump surgery three mice were sacrificed during the time course of the experiment and therefore are not represented in the study data. The mice were split into two groups to perform continuous subcutaneous infusions of Dulbecco's phosphate buffered saline (vehicle;  $n = 8$ ) or GLP-1 (GLP-1;  $n = 9$ ) agonist Liraglutide (Victoza®, batch# CS6C214, Novo Nordisk A/S, Bagsvaerd, Denmark) to investigate the effects of the respective treatment on peri-gonadal white adipose tissue gene expression. Vehicle or Liraglutide at a dose of 600 µg/kg/d were continuously infused (at a dose of 600 µg/kg/d) over a 14 days period. On the morning after the osmotic mini pump reservoirs liquid content should have been consumed mice were dissected to remove the peri-gonadal fat pad for white adipose tissue gene expression analysis. To dissect the white adipose tissue, mice were anaesthetized under isoflurane (3.0% using a flow of 0.6 L/min) inhalation anesthesia and finally sacrificed by cervical dislocation. At least 220 mg of white adipose tissue was quickly removed, shock-frozen in liquid nitrogen and thereafter stored at  $-80^\circ\text{C}$  for mRNA extraction. The animal study conformed to the German law for the protection of animal guidelines and the guide for the care and use of laboratory animals published by the US National Institutes of Health (NIH Publications No 85–23, revised 2011) as well as to Sanofi-Aventis Ethical Committee guidelines and were approved by a local authority ethics review board (RP Darmstadt).

### Transcriptome profiling of liraglutide treated mice and baseline controls

To perform RNA-Seq analysis, samples were single-end sequenced at a depth of 75 – 80M reads per sample with a read-length of 51 bp using an Illumina Hiseq2500. Raw sequencing files were quality controlled with FastQC [48]. Alignment and trimming of reads was performed using the OSA [49] algorithm against the mouse reference genome b38.1 with RefSeq as the gene model as implemented in OmicSoft® ArraySuite® software, version 8. RNA transcripts were quantified using RSEM methods [50] as implemented in Arraystudio counting count the read fragments mapping to each individual gene and quantify expression by the corresponding FPKM. In summary, expression was measured as FPKM for 25,054 unique genes. Principal component analysis was then performed to check for possible batch effects and outliers complemented by calculating the RNA-Seq 5'→3' trend for each sample. One sample for the vehicle control as well as for the Liraglutide group were identified as outliers and removed from subsequent analysis, resulting in 7 and 8 samples remaining for each group respectively. Abundance values (counts) were normalized and compared between liraglutide treated mice versus baseline controls using DESeq2 [51]. All *P*-values were adjusted for multiple testing by the Benjamini-Hochberg method [44].

## Gene set enrichment analysis of the transcriptome profiling of liraglutide treated mice vs. baseline controls

Gene set enrichment analysis was performed using QIAGEN's Ingenuity® Pathway Analysis, a web-based bioinformatics tool [Qiagen, Redwood City, CA, USA]. A given set of input genes was associated with molecular networks based on their connectivity in the Ingenuity Pathways Knowledge Base. Fisher's exact test was used to determine the probability that each biological function assigned to that data set was attributable to chance alone. The goal of the IPA Upstream Regulator analytic is to identify the cascade of upstream transcriptional regulators (any molecule that can affect the expression of other molecule) that can explain the observed gene expression changes in a given dataset. For each potential transcriptional regulator two statistical measures, an overlap *P*-value and an activation z-score are computed. The overlap *P*-value calls likely upstream regulators based on significant overlap between dataset genes and known targets regulated by a transcriptional regulator. The activation z-score is used to infer likely activation states of upstream regulators based on comparison with a model that assigns random regulation directions (S8 Table). Analyses included direct and indirect relationships that have been experimentally observed in mice, rat or human studies. All differentially expressed genes ( $n = 342$  with a  $FDR < 0.05$ ) were used for the analysis.

## Tissue specific network analysis of genes identified in both human and animal experiments

To verify if the genes identified in the independent human and mouse experiments had interactions in specific interaction networks for the islet and adipose tissue, we used the Genome-scale Integrated Analysis of gene Networks in Tissues—GIANT [25]. GIANT leverages a tissue-specific gold standard to automatically up-weight datasets relevant to a tissue from a large data compendium of diverse tissues and cell-types. The resulting functional networks accurately capture tissue-specific functional interactions. We only included interactions with a relationship confidence  $> 0.28$  (within the range of 0 to 1).

## Supporting information

**S1 Fig. Q-Q plot of the GWAS on GLP-1 stimulated insulin secretion as measured with the hyperglycemic clamp in the NTR cohort.**

(TIF)

**S2 Fig. Manhattan plot of the GWAS on GLP-1 stimulated insulin secretion as measured with the hyperglycemic clamp in the NTR cohort.**

(TIF)

**S3 Fig. Flowchart illustrating the analysis workflow. A** The discovery analysis consisted of an integrative network analysis where GLP-1 stimulated insulin secretion GWAS *P*-values were combined with a beta-cell specific PPI network to identify enriched network modules or 'network hotspots'. The top scoring network module was distilled into a smaller consensus network, by combining top selected nodes from a second network module search. **B** Functional annotation of the consensus network genes and SNPs consisted of pathway overrepresentation analysis, eQTL lookups in pancreatic islets and blood and overlaps with islet regulatory elements (promoters and enhancer clusters). **C** We attempted validation of the network by calculating a combined z-score for the top-scoring network SNPs in an independent cohort and investigating polygenic risk scores from the same SNPs in OGTT data from MAGIC investigators. These results for the network SNPs were compared to those obtained by a matched

number of top GWAS SNPs. Finally the results from a study of the effects of liraglutide (a GLP-1 agonist) on mouse adipose tissue were compared to the findings from the network and pathway analysis of human genomic data.

(TIF)

**S4 Fig. Network module z-scores derived from real (red) and randomized (black) gene significance scores.** Network module z-scores based on randomized gene significance scores are shown as the mean of 10 randomizations with 95% confidence intervals ( $SEM^* 1.96$ ).

(TIF)

**S5 Fig. The GO terms enriched (BH adjusted  $P$ -value  $< 1 \times 10^{-3}$ ) within the GLP-1 response network, compared to the whole beta-cell PPI network.** Level 2 GO terms are shown for each of the categories; molecular function (red), cellular component (yellow) and biological process (blue).

(TIF)

**S6 Fig. The percentage of GLP-1 stimulated insulin secretion consensus network loci (green) overlapping islet regulatory elements, shown in comparison to genome-wide significant fasting glucose associated loci (blue), genome-wide significant T2D associated loci (red) and loci with  $P < 1 \times 10^{-4}$  in a GWAS of corrected insulin response adjusted for insulin sensitivity index (orange).**

(TIF)

**S7 Fig. The combined z-score in the Tübingen validation cohort for the top 31 independent GWAS SNPs (red line) compared to 100,000 z-scores obtained from randomly selected sets of SNPs from the beta-cell network (histogram), empirical  $P$ -value = 0.95.**

(TIF)

**S8 Fig. Mouse experiment overview.** B6.Cg-Lep ob/ob mice were treated with 600  $\mu\text{g}/\text{kg}/\text{d}$  liraglutide ( $n = 9$ ) or vehicle ( $n = 8$ ).

(TIF)

**S9 Fig. Visualization of subnetworks prioritized from the mouse adipose tissue transcriptome experiment.** A) Collagen genes interact with genes that also appear in the human consensus network, and which connect the collagen pathway to the GLP-1 receptor. Gene nodes are colored by up- (red) and down- (green) regulation in the liraglutide treated animals versus untreated controls. B) *CD44* is an upstream regulator of collagen genes and *CTNNB1*. Interactions are based on the Ingenuity Pathway Analyses library.

(TIF)

**S10 Fig. Tissue-specific functional interaction networks from GIANT.** Relevant genes (largest nodes) overlapping in the human and mouse studies were used to query tissue-specific interactions. The networks indicate that these genes have functional interactions in both pancreatic islet (A) and adipose tissue (B). Minimum relationship confidence is highlighted by the colors from green (0) to dark red (1).

(TIF)

**S1 Table. The six independent signals from the discovery GWAS on GLP-1 stimulated insulin secretion with  $P < 1 \times 10^{-5}$ , well as the validation and meta-analysis statistics.**

(XLSX)

**S2 Table. An overview of the genes in the GLP-1 response consensus network.** The SNP with the minimum discovery GWAS  $P$ -value mapping to each gene is shown together with the

discovery GWAS, validation and meta-analysis statistics.  
(XLSX)

**S3 Table. An overview of the eQTL and islet regulatory element overlap lookup for the GLP-1 response consensus network SNPs.**

(XLSX)

**S4 Table. OGTT-derived phenotypes for which summary statistics from MAGIC were used to investigate SNP and polygenic risk score associations.** Table adapted from Prokopenko et al. [11]

(XLSX)

**S5 Table. Nominally significant ( $P < 0.05$ , highlighted in bold) associations between SNPs from the GLP-1 response consensus network and quantitative metabolic traits from MAGIC.** The last column shows the effect (beta) of the effect allele used in MAGIC on GLP-1 stimulated insulin secretion in the NTR cohort.

(XLSX)

**S6 Table. The association between weighted polygenic risk scores for SNPs in the GLP-1 response consensus network with discovery GWAS  $P < 5 \times 10^{-4}$  and OGTT-derived phenotypes from MAGIC.**  $P$ -values  $< 0.05$  are highlighted in bold. The phenotypes are described in more detail in [S4 Table](#).

(XLSX)

**S7 Table. The association between weighted polygenic risk scores for LD-pruned top GWAS SNPs (matched number of SNPs compared to [S6 Table](#)) and OGTT-derived phenotypes from MAGIC.**

(XLSX)

**S8 Table. Upstream regulators identified in the transcriptome experiment of differentially expressed genes in liraglutide treated ob/ob mice versus baseline controls.** The overlap  $P$ -value calls likely upstream regulators based on significant overlap between dataset genes and known targets regulated by a transcriptional regulator. The  $z$ -score algorithm is designed to produce either a prediction of activation or inhibition (or no prediction). The analysis examines the known targets of each upstream regulator in the dataset, compares the targets' actual direction of change to expectations derived from the literature, then issues a prediction for each upstream regulator. The direction of change is the gene expression in the experimental samples relative to a control. If the direction of change is consistent with the literature across most targets, IPA predicts that the upstream regulator is more active in the experimental sample than in the control. Mostly inconsistent with the literature (anti-correlated with the literature), IPA predicts that the upstream regulator is less active in the experimental sample than in the control. If there is a random pattern relative to the literature, IPA does not make an activation or inhibition prediction for the upstream regulator. However, in these case, there may still be a significant overlap (Fisher's Exact  $P$ -value), just no clear pattern to predict a direction of activation. Genes highlighted in yellow were also present in the human beta-cell GLP-1 response consensus network.

(XLSX)

**S9 Table. Islet and pancreatic transcription factors that were specifically included in the beta-cell PPI network.**

(XLSX)

## Acknowledgments

We are very grateful to all participants who took part in these studies. Data on glycaemic traits have been contributed by MAGIC investigators and have been downloaded from [www.magicinvestigators.org](http://www.magicinvestigators.org).

## Author Contributions

**Conceptualization:** Valborg Gudmundsdottir, Helle Krogh Pedersen, Caroline Brorsson, Christopher J. Groves, Jorge Ferrer, Mark I. McCarthy, Ewan R. Pearson, Ramneek Gupta, Søren Brunak, Leen M. 't Hart.

**Data curation:** Nienke van Leeuwen, Karina Banasik, Anubha Mahajan, Martijn van de Bunt, Annemarie M. C. Simonis-Bik, Elisabeth M. W. Eekhoff.

**Formal analysis:** Valborg Gudmundsdottir, Helle Krogh Pedersen, Karla Viviani Allebrandt, Nienke van Leeuwen, Karina Banasik, Joris Deelen, Axel Dietrich, Leen M. 't Hart.

**Funding acquisition:** Jorge Ferrer, Mark I. McCarthy, Ewan R. Pearson, Søren Brunak, Leen M. 't Hart.

**Investigation:** Karina Banasik, Anubha Mahajan, Christopher J. Groves, Adem Y. Dawed, Andreas Fritsche, Harald Staiger, Annemarie M. C. Simonis-Bik, Thomas Hübschle, Gonneke Willemsen, Eco J. C. de Geus, Dorret I. Boomsma, Elisabeth M. W. Eekhoff.

**Methodology:** Valborg Gudmundsdottir, Helle Krogh Pedersen, Andreas Fritsche, Gonneke Willemsen, Hans-Ulrich Häring, Eco J. C. de Geus, Dorret I. Boomsma, Elisabeth M. W. Eekhoff.

**Project administration:** Valborg Gudmundsdottir, Caroline Brorsson, Ewan R. Pearson, Leen M. 't Hart.

**Resources:** Nienke van Leeuwen, Harald Staiger, Mark H. H. Kramer, Gonneke Willemsen, Hans-Ulrich Häring, Eco J. C. de Geus, Dorret I. Boomsma, Elisabeth M. W. Eekhoff, Jorge Ferrer, Leen M. 't Hart.

**Software:** Joris Deelen.

**Supervision:** Mark I. McCarthy, Ewan R. Pearson, Ramneek Gupta, Søren Brunak, Leen M. 't Hart.

**Validation:** Karla Viviani Allebrandt, Adem Y. Dawed, Andreas Fritsche, Harald Staiger, Hans-Ulrich Häring.

**Visualization:** Valborg Gudmundsdottir, Helle Krogh Pedersen, Karla Viviani Allebrandt, Leen M. 't Hart.

**Writing – original draft:** Valborg Gudmundsdottir, Helle Krogh Pedersen, Karla Viviani Allebrandt, Caroline Brorsson, Søren Brunak, Leen M. 't Hart.

**Writing – review & editing:** Valborg Gudmundsdottir, Helle Krogh Pedersen, Karla Viviani Allebrandt, Caroline Brorsson, Nienke van Leeuwen, Karina Banasik, Anubha Mahajan, Christopher J. Groves, Martijn van de Bunt, Adem Y. Dawed, Andreas Fritsche, Harald Staiger, Annemarie M. C. Simonis-Bik, Joris Deelen, Mark H. H. Kramer, Axel Dietrich, Thomas Hübschle, Gonneke Willemsen, Hans-Ulrich Häring, Eco J. C. de Geus, Dorret I. Boomsma, Elisabeth M. W. Eekhoff, Jorge Ferrer, Mark I. McCarthy, Ewan R. Pearson, Ramneek Gupta, Søren Brunak, Leen M. 't Hart.

## References

1. Mehta A, Marso SP, Neeland IJ. Liraglutide for weight management: a critical review of the evidence. *Obes Sci Pract*. 2016; <https://doi.org/10.1002/osp4.84> PMID: 28392927
2. Simonis-Bik AMC, Eekhoff EMW, De Moor MHM, Kramer MHH, Boomsma DI, Heine RJ, et al. Genetic influences on the insulin response of the beta cell to different secretagogues. *Diabetologia*. 2009; 52: 2570–2577. <https://doi.org/10.1007/s00125-009-1532-3> PMID: 19802603
3. Schäfer SA, Tschrirter O, Machicao F, Thamer C, Stefan N, Gallwitz B, et al. Impaired glucagon-like peptide-1-induced insulin secretion in carriers of transcription factor 7-like 2 (TCF7L2) gene polymorphisms. *Diabetologia*. 2007; 50: 2443–2450. <https://doi.org/10.1007/s00125-007-0753-6> PMID: 17661009
4. Sathananthan A, Man CD, Micheletto F, Zinsmeister AR, Camilleri M, Giesler PD, et al. Common genetic variation in GLP1R and insulin secretion in response to exogenous GLP-1 in nondiabetic subjects: a pilot study. *Diabetes Care*. 2010; 33: 2074–6. <https://doi.org/10.2337/dc10-0200> PMID: 20805279
5. Schäfer SA, Müssig K, Staiger H, Machicao F, Stefan N, Gallwitz B, et al. A common genetic variant in WFS1 determines impaired glucagon-like peptide-1-induced insulin secretion. *Diabetologia*. 2009; 52: 1075–82. <https://doi.org/10.1007/s00125-009-1344-5> PMID: 19330314
6. 't Hart LM, Fritsche A, Nijpels G, Van Leeuwen N, Donnelly LA, Dekker JM, et al. The CTRB1/2 locus affects diabetes susceptibility and treatment via the incretin pathway. *Diabetes*. 2013; 62: 3275–3281. <https://doi.org/10.2337/db13-0227> PMID: 23674605
7. Califano A, Butte AJ, Friend S, Ideker T, Schadt E. Leveraging models of cell regulation and GWAS data in integrative network-based association studies. *Nat Genet*. 2012; 44: 841–847. <https://doi.org/10.1038/ng.2355> PMID: 22836096
8. Baranzini SE, Galwey NW, Wang J, Khankhanian P, Lindberg R, Pelletier D, et al. Pathway and network-based analysis of genome-wide association studies in multiple sclerosis. *Hum Mol Genet*. 2009; 18: 2078–2090. <https://doi.org/10.1093/hmg/ddp120> PMID: 19286671
9. Taneera J, Lang S, Sharma A, Fadista J, Zhou Y, Ahlqvist E, et al. A systems genetics approach identifies genes and pathways for type 2 diabetes in human islets. *Cell Metab*. 2012; 16: 122–134. <https://doi.org/10.1016/j.cmet.2012.06.006> PMID: 22768844
10. 't Hart LM, Simonis-Bik AM, Nijpels G, Van Haeften TW, Schäfer SA, Houwing-Duistermaat JJ, et al. Combined risk allele score of eight type 2 diabetes genes is associated with reduced first-phase glucose-stimulated insulin secretion during hyperglycemic clamps. *Diabetes*. 2010; 59: 287–292. <https://doi.org/10.2337/db09-0736> PMID: 19808892
11. Prokopenko I, Poon W, Mägi R, Prasad B R, Salehi SA, Almgren P, et al. A Central Role for GRB10 in Regulation of Islet Function in Man. *PLoS Genet*. 2014; 10: e1004235. <https://doi.org/10.1371/journal.pgen.1004235> PMID: 24699409
12. Magger O, Waldman YY, Ruppin E, Sharan R. Enhancing the prioritization of disease-causing genes through tissue specific protein interaction networks. *PLoS Comput Biol*. 2012; 8: e1002690. <https://doi.org/10.1371/journal.pcbi.1002690> PMID: 23028288
13. Nica AC, Ongen H, Irminger JC, Bosco D, Berney T, Antonarakis SE, et al. Cell-type, allelic, and genetic signatures in the human pancreatic beta cell transcriptome. *Genome Res*. 2013; 23: 1554–1562. <https://doi.org/10.1101/gr.150706.112> PMID: 23716500
14. Ideker T, Ozier O, Schwikowski B, Siegel AF. Discovering regulatory and signalling circuits in molecular interaction networks. *Bioinformatics*. 2002; 18 Suppl 1: S233–S240. [https://doi.org/10.1093/bioinformatics/18.suppl\\_1.S233](https://doi.org/10.1093/bioinformatics/18.suppl_1.S233)
15. Leech CA, Dzhura I, Chepurny OG, Kang G, Schwede F, Genieser HG, et al. Molecular physiology of glucagon-like peptide-1 insulin secretagogue action in pancreatic  $\beta$  cells. *Prog Biophys Mol Biol*. 2011; 107: 236–247. <https://doi.org/10.1016/j.pbiomolbio.2011.07.005> PMID: 21782840
16. Shibasaki T, Takahashi H, Miki T, Sunaga Y, Matsumura K, Yamanaka M, et al. Essential role of Epac2/Rap1 signaling in regulation of insulin granule dynamics by cAMP. *Proc Natl Acad Sci U S A*. 2007; 104: 19333–19338. <https://doi.org/10.1073/pnas.0707054104> PMID: 18040047
17. Veluthakal R, Tunduguru R, Arora DK, Sidarala V, Syeda K, Vlaar CP, et al. VAV2, a guanine nucleotide exchange factor for Rac1, regulates glucose-stimulated insulin secretion in pancreatic beta cells. *Diabetologia*. 2015; 58: 2573–2581. <https://doi.org/10.1007/s00125-015-3707-4> PMID: 26224100
18. Buteau J, Spatz ML, Accili D. Transcription factor FoxO1 mediates glucagon-like peptide-1 effects on pancreatic  $\beta$ -cell mass. *Diabetes*. 2006; 55: 1190–1196. <https://doi.org/10.2337/db05-0825> PMID: 16644672

19. Tsai F-J, Yang C-F, Chen C-C, Chuang L-M, Lu C-H, Chang C-T, et al. A genome-wide association study identifies susceptibility variants for type 2 diabetes in Han Chinese. *PLoS Genet.* 2010; 6: e1000847. <https://doi.org/10.1371/journal.pgen.1000847> PMID: 20174558
20. Below JE, Gamazon ER, Morrison J V, Konkashbaev A, Pluzhnikov A, McKeigue PM, et al. Genome-wide association and meta-analysis in populations from Starr County, Texas, and Mexico City identify type 2 diabetes susceptibility loci and enrichment for expression quantitative trait loci in top signals. *Diabetologia.* 2011; 54: 2047–55. <https://doi.org/10.1007/s00125-011-2188-3> PMID: 21647700
21. Kodama K, Horikoshi M, Toda K, Yamada S, Hara K, Irie J, et al. Expression-based genome-wide association study links the receptor CD44 in adipose tissue with type 2 diabetes. *Proc Natl Acad Sci.* 2012; 109: 7049–7054. <https://doi.org/10.1073/pnas.1114513109> PMID: 22499789
22. Hasstedt SJ, Highland HM, Elbein SC, Hanis CL, Das SK. Five linkage regions each harbor multiple type 2 diabetes genes in the African American subset of the GENNID Study. *J Hum Genet.* 2013; 58: 378–83. <https://doi.org/10.1038/jhg.2013.21> PMID: 23552671
23. Pasquali L, Gaulton KJ, Rodríguez-Seguí SA, Mularoni L, Miguel-Escalada I, Akerman I, et al. Pancreatic islet enhancer clusters enriched in type 2 diabetes risk-associated variants. *Nat Genet.* 2014; 46: 136–43. <https://doi.org/10.1038/ng.2870> PMID: 24413736
24. Hayes MR, Kanoski SE, Alhadeff AL, Grill HJ. Comparative Effects of the Long-Acting GLP-1 Receptor Ligands, Liraglutide and Exendin-4, on Food Intake and Body Weight Suppression in Rats. *Obesity.* 2011; 19: 1342–1349. <https://doi.org/10.1038/oby.2011.50> PMID: 21415845
25. Greene CS, Krishnan A, Wong AK, Ricciotti E, Zelaya RA, Himmelstein DS, et al. Understanding multi-cellular function and disease with human tissue-specific networks. *Nat Genet.* 2015; 47: 569–76. <https://doi.org/10.1038/ng.3259> PMID: 25915600
26. Schadt EE. Molecular networks as sensors and drivers of common human diseases. *Nature.* 2009; 461: 218–223. <https://doi.org/10.1038/nature08454> PMID: 19741703
27. Beiroa D, Imbernon M, Gallego R, Senra A, Herranz D, Villarroya F, et al. GLP-1 agonism stimulates brown adipose tissue thermogenesis and browning through hypothalamic AMPK. *Diabetes.* 2014; 63: 3346–58. <https://doi.org/10.2337/db14-0302> PMID: 24917578
28. Kaido T, Yebra M, Cirulli V, Montgomery AM. Regulation of human  $\beta$ -cell adhesion, motility, and insulin secretion by collagen IV and its receptor  $\alpha 1\beta 1$ . *J Biol Chem.* 2004; 279: 53762–53769. <https://doi.org/10.1074/jbc.M411202200> PMID: 15485856
29. Rondas D, Tomas A, Halban PA. Focal adhesion remodeling is crucial for glucose-stimulated insulin secretion and involves activation of focal adhesion kinase and paxillin. *Diabetes.* 2011; 60: 1146–57. <https://doi.org/10.2337/db10-0946> PMID: 21357465
30. Dayeh T, Volkov P, Saló S, Hall E, Nilsson E, Olsson AH, et al. Genome-wide DNA methylation analysis of human pancreatic islets from type 2 diabetic and non-diabetic donors identifies candidate genes that influence insulin secretion. *PLoS Genet.* 2014; 10: e1004160. <https://doi.org/10.1371/journal.pgen.1004160> PMID: 24603685
31. Wang HS, Hung Y, Su CH, Peng ST, Guo YJ, Lai MC, et al. CD44 Cross-linking induces integrin-mediated adhesion and transendothelial migration in breast cancer cell line by up-regulation of LFA-1 (alpha L beta2) and VLA-4 (alpha4beta1). *Exp Cell Res.* 2005; 304: 116–126. <https://doi.org/10.1016/j.yexcr.2004.10.015> PMID: 15707579
32. Challa TD, Beaton N, Arnold M, Rudofsky G, Langhans W, Wolfrum C. Regulation of adipocyte formation by GLP-1/GLP-1R signaling. *J Biol Chem.* 2012; 287: 6421–30. <https://doi.org/10.1074/jbc.M111.310342> PMID: 22207759
33. Lamont BJ, Li Y, Kwan E, Brown TJ, Gaisano H, Drucker DJ, et al. Pancreatic GLP-1 receptor activation is sufficient for incretin control of glucose metabolism in mice. *J Clin Invest.* 2012; 122: 388–402. <https://doi.org/10.1172/JCI42497> PMID: 22182839
34. Lal H, Guleria RS, Foster DM, Lu G, Watson LE, Sanghi S, et al. Integrins: novel therapeutic targets for cardiovascular diseases. *Cardiovasc Hematol Agents Med Chem.* 2007; 5: 109–32. <https://doi.org/10.2174/187152507780363223> PMID: 17430135
35. Zhang H, Xiong Z, Wang J, Zhang S, Lei L, Yang L, et al. Glucagon-like peptide-1 protects cardiomyocytes from advanced oxidation protein product-induced apoptosis via the PI3K/Akt/Bad signaling pathway. *Mol Med Rep.* 2016; 13: 1593–1601. <https://doi.org/10.3892/mmr.2015.4724> PMID: 26717963
36. Burgmaier M, Liberman A, Möllmann J, Kahles F, Reith S, Lebherz C, et al. Glucagon-like peptide-1 (GLP-1) and its split products GLP-1(9–37) and GLP-1(28–37) stabilize atherosclerotic lesions in *apoe*<sup>-/-</sup> mice. *Atherosclerosis.* 2013; 231: 427–35. <https://doi.org/10.1016/j.atherosclerosis.2013.08.033> PMID: 24267262
37. Maranville JC, Cox NJ. Pharmacogenomic variants have larger effect sizes than genetic variants associated with other dichotomous complex traits. *Pharmacogenomics J. NIH Public Access.* 2016; 16: 388–392. <https://doi.org/10.1038/tpj.2015.47> PMID: 26149738

38. Fritsche A, Stefan N, Hardt E, Schützenauer S, Häring H, Stumvoll M. A novel hyperglycaemic clamp for characterization of islet function in humans: Assessment of three different secretagogues, maximal insulin response and reproducibility. *Eur J Clin Invest*. 2000; 30: 411–418. <https://doi.org/10.1046/j.1365-2362.2000.00649.x> PMID: 10809901
39. Uh H-W, Deelen J, Beekman M, Helmer Q, Rivadeneira F, Hottenga J-J, et al. How to deal with the early GWAS data when imputing and combining different arrays is necessary. *Eur J Hum Genet*. 2012; 20: 572–6. <https://doi.org/10.1038/ejhg.2011.231> PMID: 22189269
40. Mägi R, Morris AP. GWAMA: software for genome-wide association meta-analysis. *BMC Bioinformatics*. 2010; 11: 288. <https://doi.org/10.1186/1471-2105-11-288> PMID: 20509871
41. Veyrieras JB, Kudaravalli S, Kim SY, Dermitzakis ET, Gilad Y, Stephens M, et al. High-resolution mapping of expression-QTLs yields insight into human gene regulation. *PLoS Genet*. 2008; 4: e1000214. <https://doi.org/10.1371/journal.pgen.1000214> PMID: 18846210
42. Li T, Wernersson R, Hansen RB, Horn H, Mercer JM, Slodkowitz G, et al. A scored human protein-protein interaction network to catalyze genomic interpretation. *Nat Methods*. 2017; 14: 61–64. <https://doi.org/10.1038/nmeth.4083> PMID: 27892958
43. Kamburov A, Stelzl U, Lehrach H, Herwig R. The ConsensusPathDB interaction database: 2013 Update. *Nucleic Acids Res*. 2013; 41: D793–D800. <https://doi.org/10.1093/nar/gks1055> PMID: 23143270
44. Benjamini Y, Hochberg Y. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B*. 1995; 57: 289–300. <https://doi.org/10.2307/2346101>
45. The GTEx Consortium. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science (80-)*. 2015; 348: 648–60. <https://doi.org/10.1126/science.1262110> PMID: 25954001
46. van de Bunt M, Manning Fox JE, Dai X, Barrett A, Grey C, Li L, et al. Transcript expression data from human islets links regulatory signals from genome-wide association studies for type 2 diabetes and glycemic traits to their downstream effectors. *PLoS Genet*. 2015; 11: e1005694. <https://doi.org/10.1371/journal.pgen.1005694> PMID: 26624892
47. Westra H-J, Peters MJ, Esko T, Yaghootkar H, Schurmann C, Kettunen J, et al. Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat Genet*. 2013; 45: 1238–1243. <https://doi.org/10.1038/ng.2756> PMID: 24013639
48. Andrews S. FastQC: A quality control tool for high throughput sequence data. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
49. Hu J, Ge H, Newman M, Liu K. OSA: a fast and accurate alignment tool for RNA-Seq. *Bioinformatics*. 2012; 28: 1933–4. <https://doi.org/10.1093/bioinformatics/bts294> PMID: 22592379
50. Li B, Ruotti V, Stewart RM, Thomson JA, Dewey CN. RNA-Seq gene expression estimation with read mapping uncertainty. *Bioinformatics*. 2010; 26: 493–500. <https://doi.org/10.1093/bioinformatics/btp692> PMID: 20022975
51. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014; 15: 550. <https://doi.org/10.1186/s13059-014-0550-8> PMID: 25516281