Supported by wellcome trust

Oxford Biomedical Research Centre

HORIZON 2020

NHS National Institute for Health Research

MRC Medical Research Council

imi Innovative Medicines Initiative

NIH National Institute of Diabetes and Digestive and Kidney Diseases

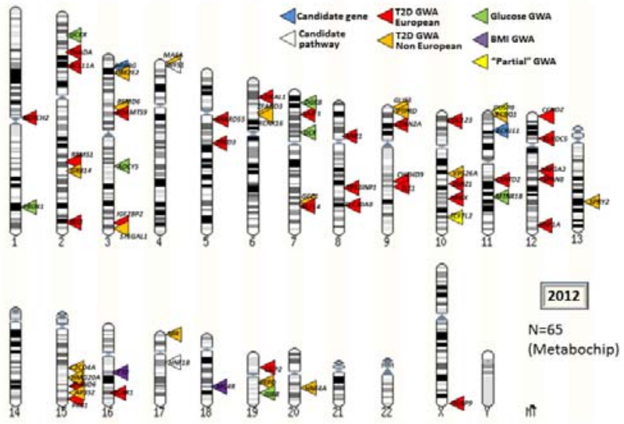
DIABETES UK CARE. CONNECT. CAMPAIGN.

Sara Althari, Amy Barratt, Nicola Beer, Amanda Bennett, Carla Burrows, Massimo Costa, Chris Groves, Benoit Hastoy, Agata Juszczak, Katia Mattis, Mark McCarthy, Natasha Ng, Katharine Owen, Anne Raimondo, Søren Thomsen, Matthias Turner, Mahesh Umpathysivam, Martijn van de Bunt

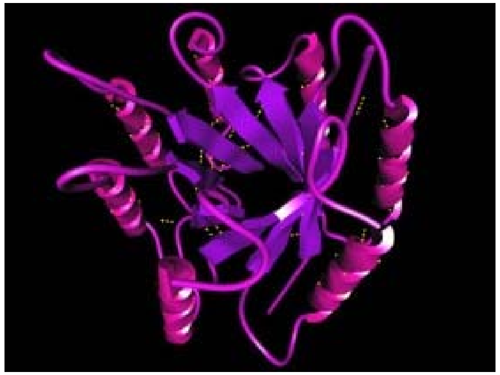
Alumni: Laura Jablonka, Neelam Hassanali, Bahram Jafar-Mohamadi, Laura McCulloch, Saima Mughal, Kara Oskak, Aparna Pal, Matthew Rees, Jana Rundle, Sara Suliman, Gaya Thanabalasingham, Mary Travers, Nicholas Tribble

How do we unlock the biology from GWAS?

The challenge……



transcript uncertainty

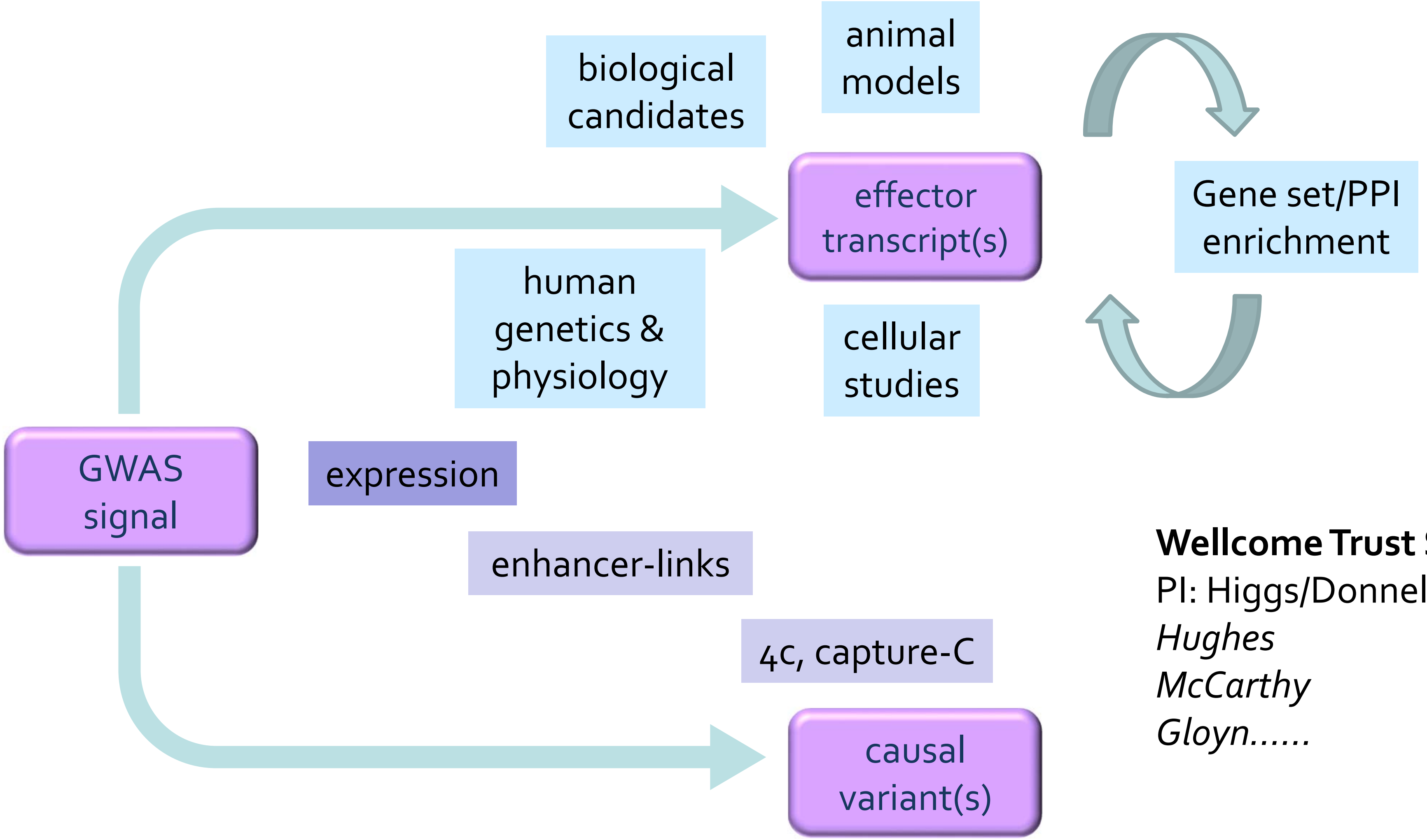


Effector transcripts
(proteins)

>100 GWAS regions for T2D risk
Most in non-coding sequence and presumed to be regulatory
Majority work through beta-cell dysfunction

Morris et al Nature Genetics 2012; Mahajan et al Nature Genetics 2014; Dimas et al Diabetes 2015

Unlocking the biology: From GWAS to function…



GWAS signal

expression

enhancer-links

4c, capture-C

causal variant(s)

human genetics & physiology

biological candidates

animal models

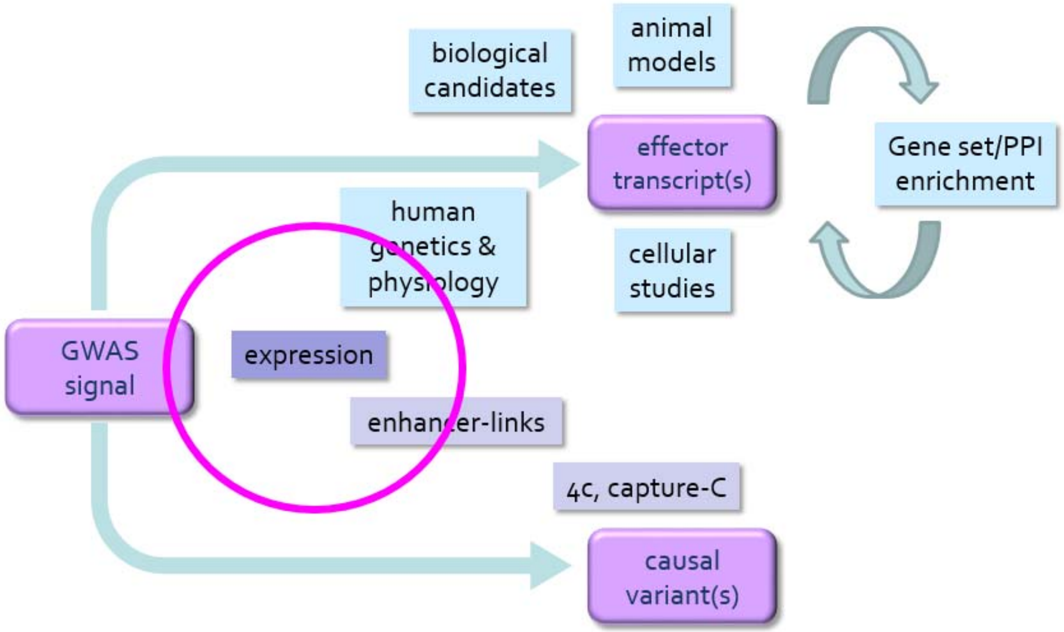
effector transcript(s)

cellular studies

Gene set/PPI enrichment

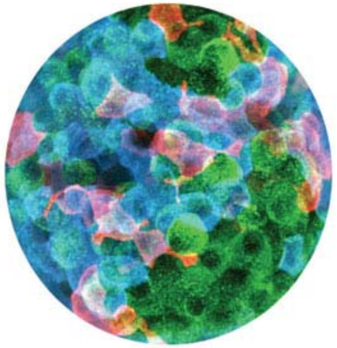
Wellcome Trust SA
PI: Higgs/Donnelly
Hughes
McCarthy
Gloyn……

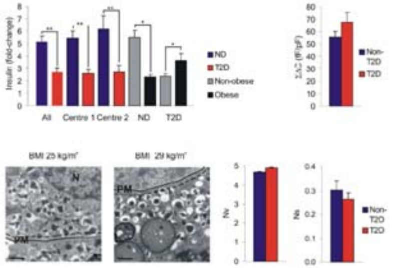
Can we identify "effector" transcripts through effects on gene expression?



Oxford/Edmonton Human Islet Biobank

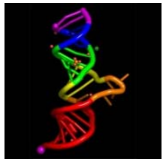
n >200





Islet function & morphology

Genotype (omni2.5 + exome) ,
DNA methylation studies (450K, WGBS) ,
Chromatin state (ATAC-seq),
Transcription Factor Binding (ChIP-seq)
DNA-interactions (Capture-C)

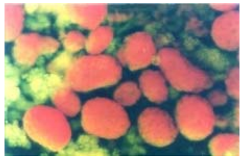


Transcript profiling (RNA-seq)
miRNA, mRNA, LncRNA
Single Cell Transcriptomics (SCTx)

Mark McCarthy, Patrick MacDonald, Patrik Rorsman

Exon eQTL analysis

Martijn van de Bunt



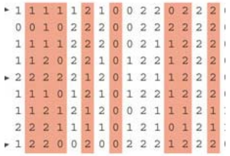
127 human islet preparations

RNA-seq
on Illumina
HiSeq2000

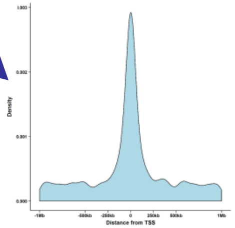
Read mapping with Tophat v2

Exon counts with RNA-SeQC package

Genotyping
on Illumina OMNI2.5+
exome array



Imputation into 1000 Genomes
phase 1 ALL panel with
IMPUTE2



eQTL mapping
(118 samples)

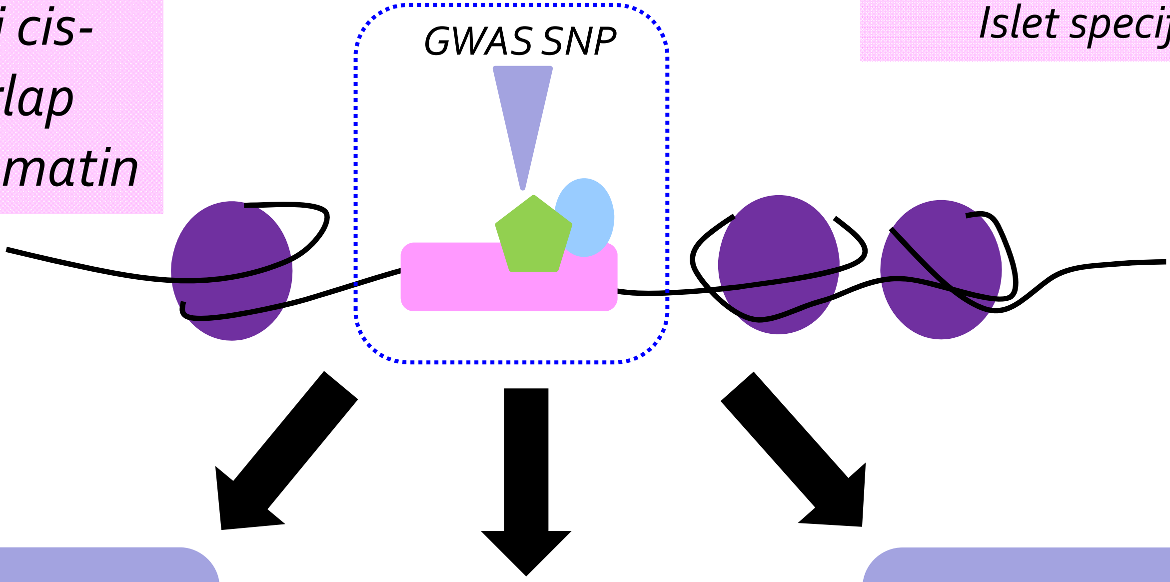
- European descent
- MatrixEQTL with 15 PEER covariates + gender
- MAF > 0.05
- Consistent expression

van de Bunt et al PLoS Genetics 2015

Exon cis eQTL analysis in human islets delivers effector transcripts at T2D & glycaemic trait GWAS loci

14 T2D/glycaemic trait GWAS loci cis-exon eQTL overlap active islet chromatin

~20% islet eQTLS
Islet specific (GTEx)



Confirmed in disease relevant tissue
e.g. CAMK1D

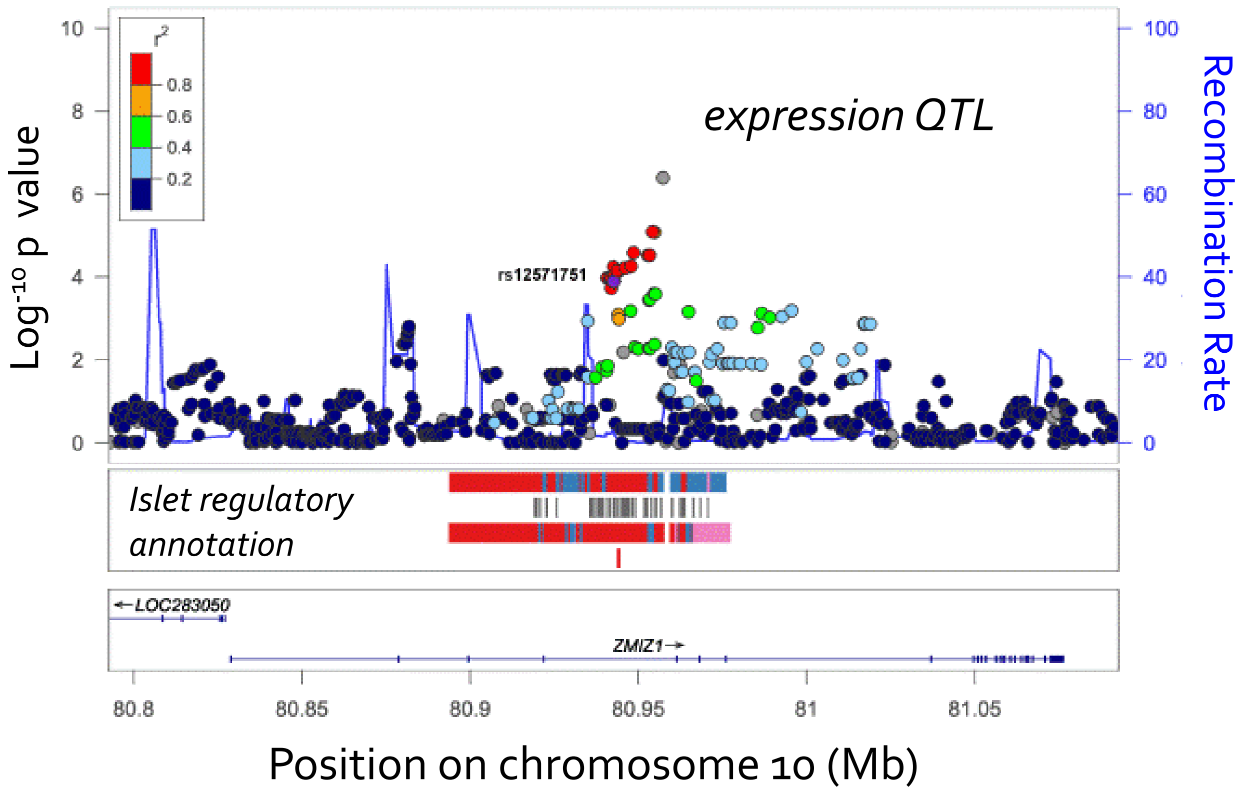
Experimental link to biological candidate
e.g ADCY5

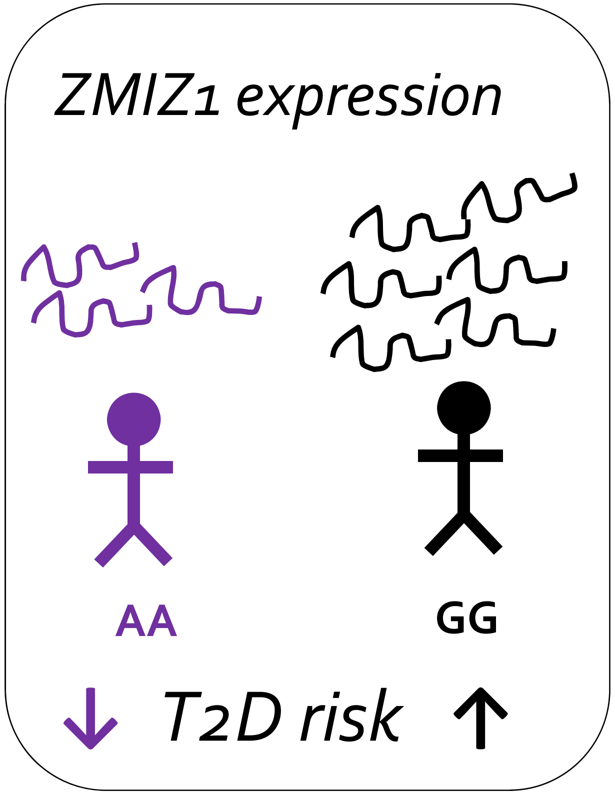
Gene with no previously known role
e.g ZMIZ1

n=118 human islet donors

van de Bunt et al PLoS Genetics 2015

T2D risk allele increases *ZMIZ1* expression





ZMIZ1 expression

AA

GG

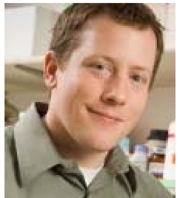
↓ *T2D risk* ↑

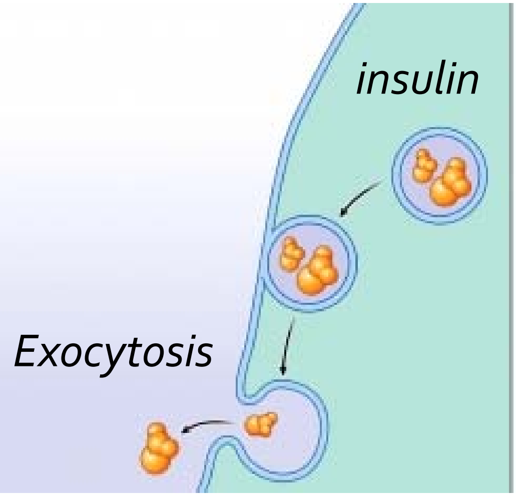
What is ZMIZ1 and what does it do in pancreatic islets?

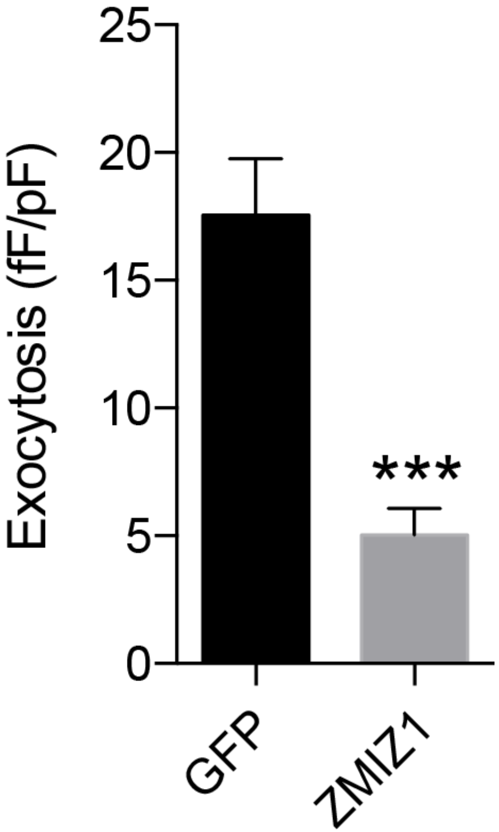
van de Bunt et al PLoS Genetics 2015

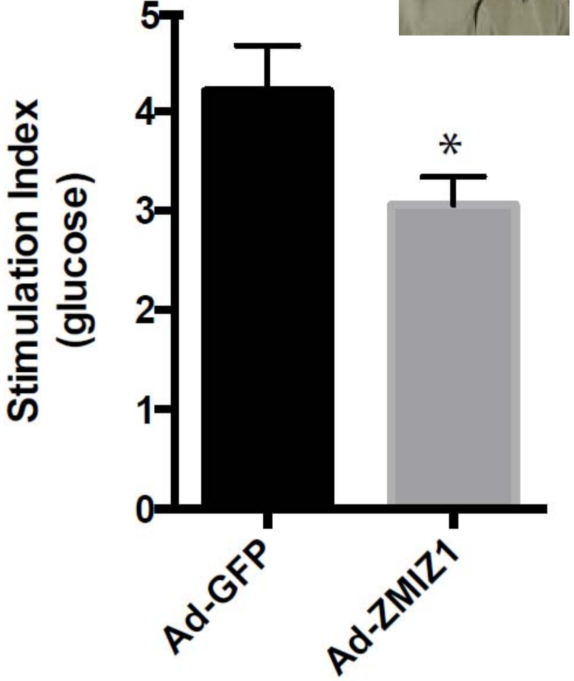
Novel Biology: *ZMIZ1* expression alters insulin secretion in human islets

Patrick MacDonald
Edmonton









Increasing ZMIZ1 in human islets reduces insulin secretion

van de Bunt et al PLoS Genetics 2015

GOLF-HIT project description

- Collaborative effort between **Lund** (Groop), **Oxford/Edmonton** (Gloyn/McCarthy/MacDonald), **Geneva** (Dermitzakis) and **FUSION** (Collins)
- Combined dataset of **429** human islet samples with RNA-seq and genotypes
- Islet physiology measurements (stimulatory index, insulin content, electrophysiology) available for subset of samples
- Additional data on ~**25 FACS-enriched** beta-cell samples

GOLF-HIT major aims

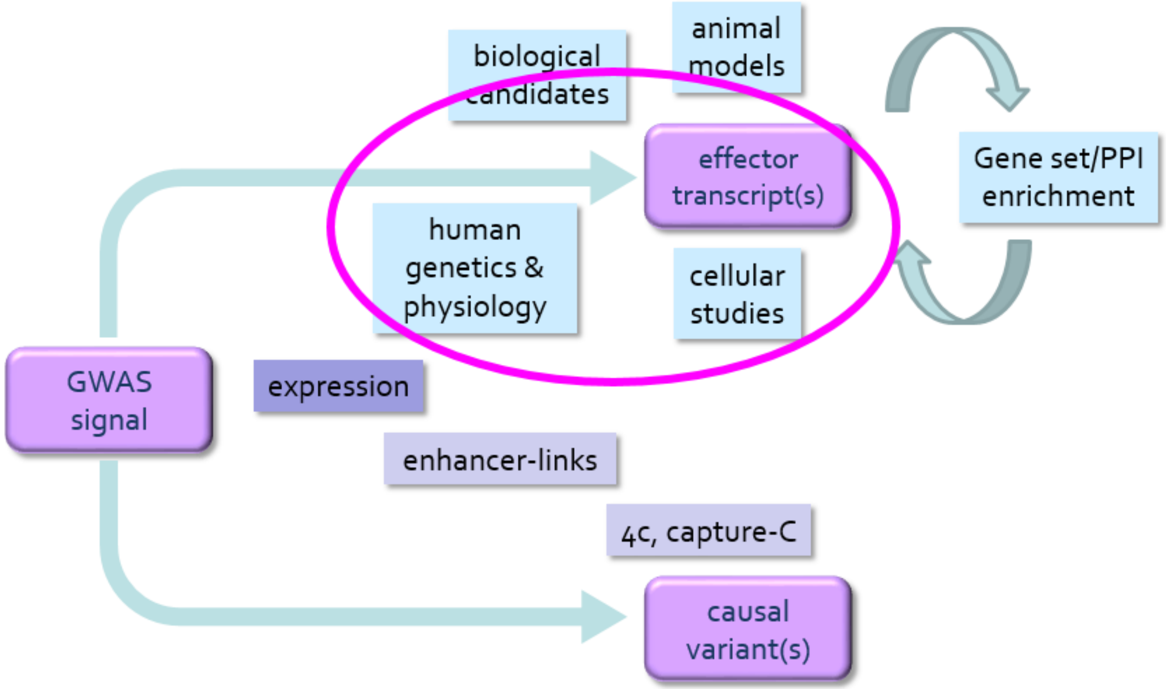
1. **Characterising the islet transcriptome (Coord: Geneva)**
   - Around 6,000 cis-eQTL genes (around 3x greater than previous studies)
   - First trans-eQTL analysis of human islets

2. **Characterising islet regulation (Coord: FUSION)**
   - Enrichment islet eQTLs in islet TSS regions marked by ATAC-seq and ChIP-seq states
   - Intersection with transcription factor footprints identified from islet ATAC-seq

3. **Interpreting trait-associated genetic variation (Coord: Oxford)**
   - 10 study-wide significant T2D GWAS variants with islet *cis*-eQTL
   - Complete replication of the 4 previous reported significant eQTL/T2D GWAS overlap

4. **Relationship of variation to islet cellular phenotypes (Coord: Lund)**
   - Largest GWAS for islet cellular phenotypes (up to 250 samples)
   - 17 genes that significantly correlate with T2D status

Can we identify causal variants through integrating genetic & genomic data?



biological candidates

animal models

effector transcript(s)

Gene set/PPI enrichment

human genetics & physiology

cellular studies

GWAS signal

expression

enhancer-links

4c, capture-C

causal variant(s)

Defining causal mechanisms using fine mapping & genomic annotation

39 GWAS loci

Metabochip
(dense scaffold across
T2D risk loci)
100,000 individuals

Impute with 1000G

Conditional analysis
(Bayesian)


Kyle Gaulton


Andrew Morris

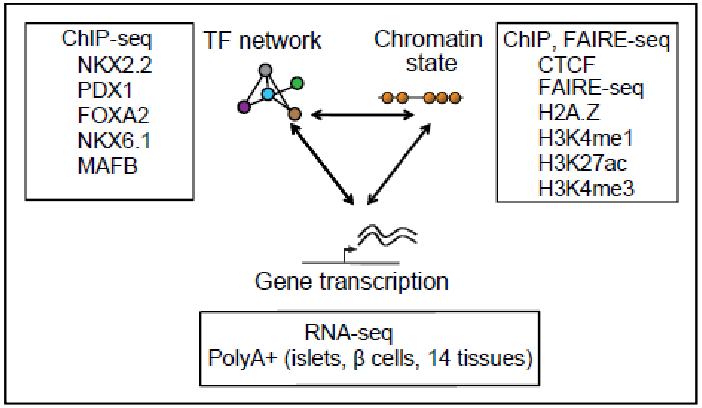

Mark McCarthy

49 distinct
signals

Variants in
credible sets

7 contain a non-synonymous variant
(*inc.HNF1A & HNF4A)*

Majority map to non-coding sequence implying effect on gene regulation



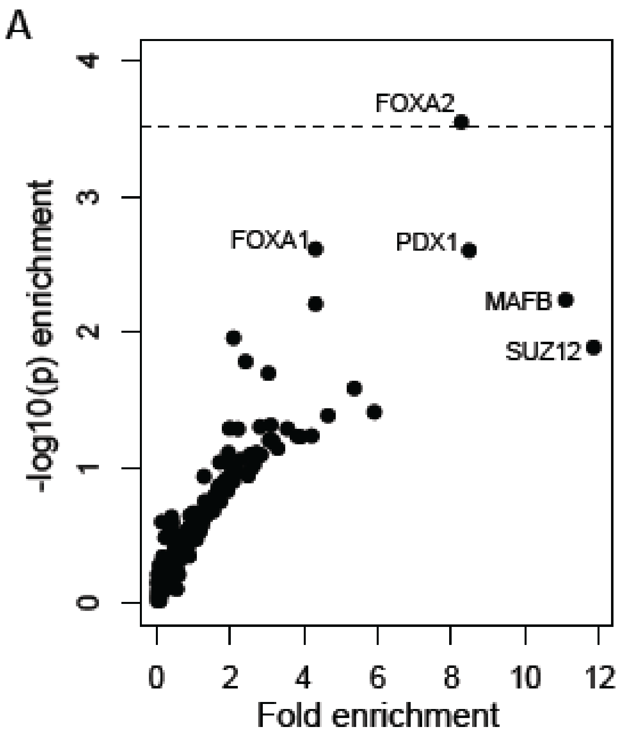
ChIP-seq data from human islet & liver cells

DIAGRAM consortium

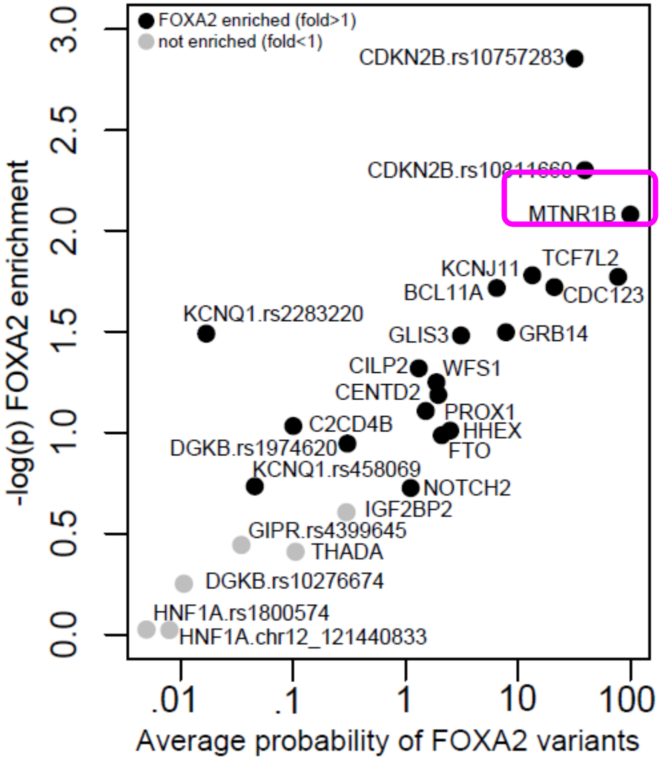
DIAbetes Genetics
Replication And Meta-analysis

Gaulton et al Nature Genetics 2015

Credible set T2D GWAS variants enriched for overlap with FOXA2 ChIP-seq binding sites



Enrichment at 19 signals



FOXA2 – proxy marker for altered transcription factor binding at islet enriched motifs

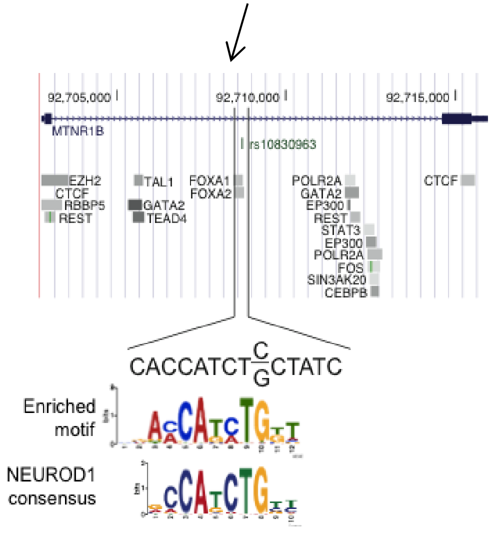
Gaulton et al Nature Genetics 2015

Causal variant & transcript at *MTNR1B* locus

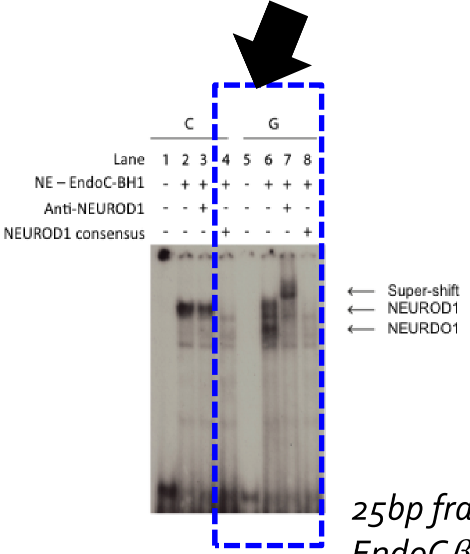
Anne Raimondo

Single causal variant

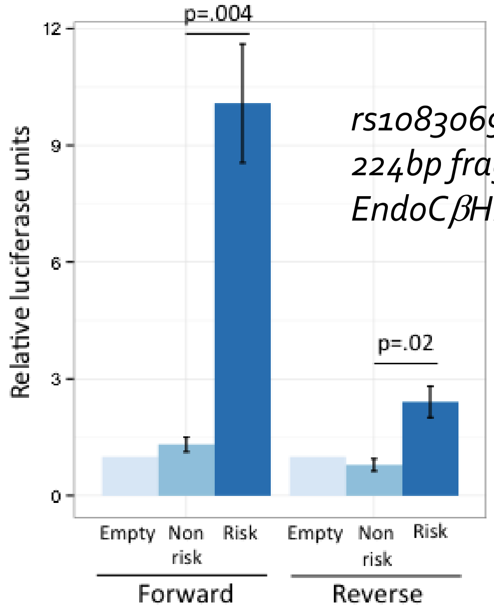
rs10830693



T2D Risk allele



25bp fragment
EndoCβH1 cells

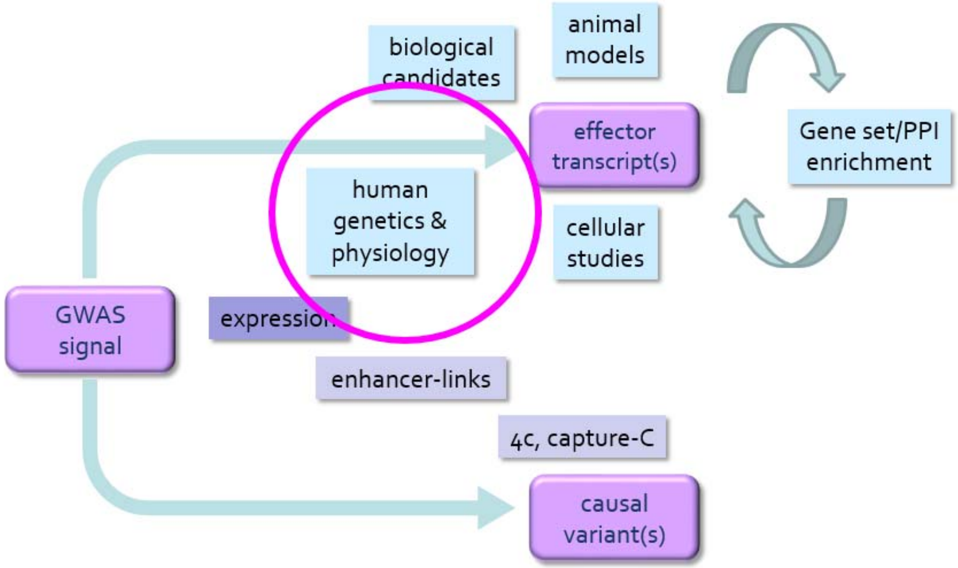


rs10830693
224bp fragment
EndoCβH1 cells

- T2D risk allele creates NEUROD1 consensus binding site
- ↑ FOXA2 bound enhancer activity with T2D risk allele at *MTNR1B* locus
- Consistent with islet eQTL – ↑ *MTNR1B* expression with T2D risk allele

Gaulton et al Nature Genetics 2015

Are coding variants associated with a glycaemic phenotype?



biological candidates

animal models

effector transcript(s)

Gene set/PPI enrichment

human genetics & physiology

cellular studies

GWAS signal

expression

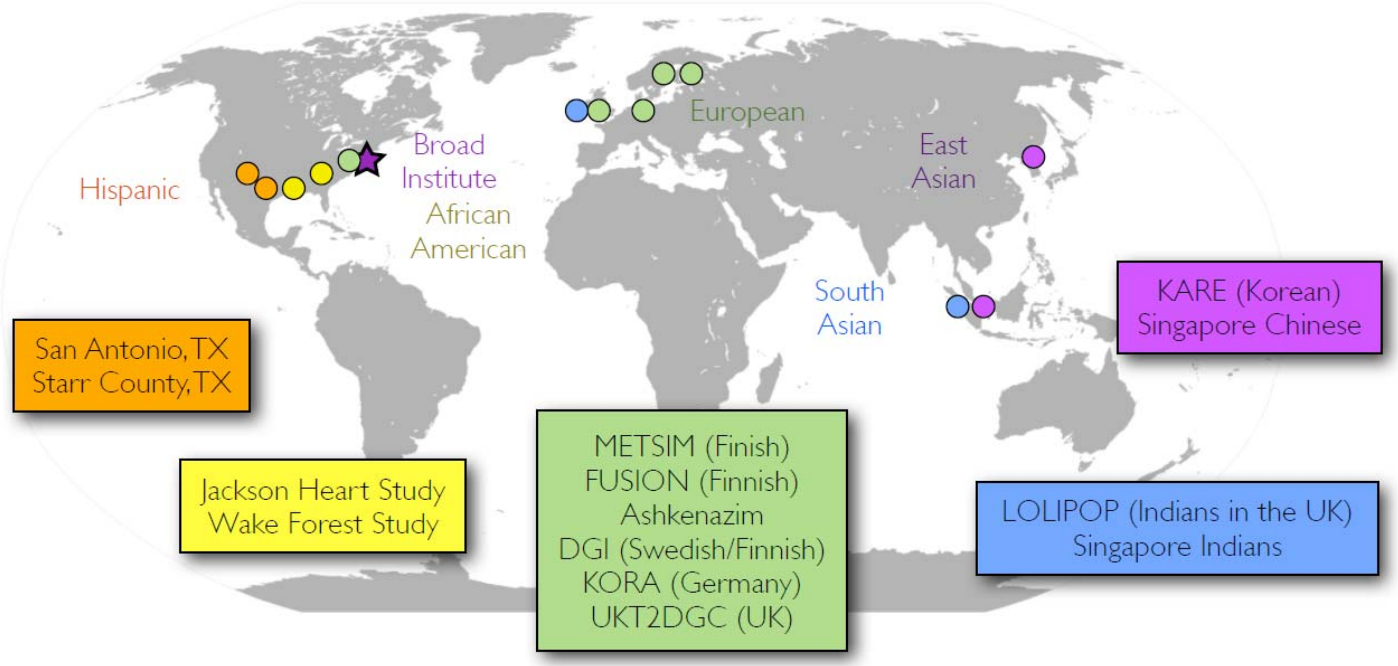
enhancer-links

4c, capture-C

causal variant(s)

David Altshuler, Mike Boehnke, Mark McCarthy

Getting a complete picture of exome variation



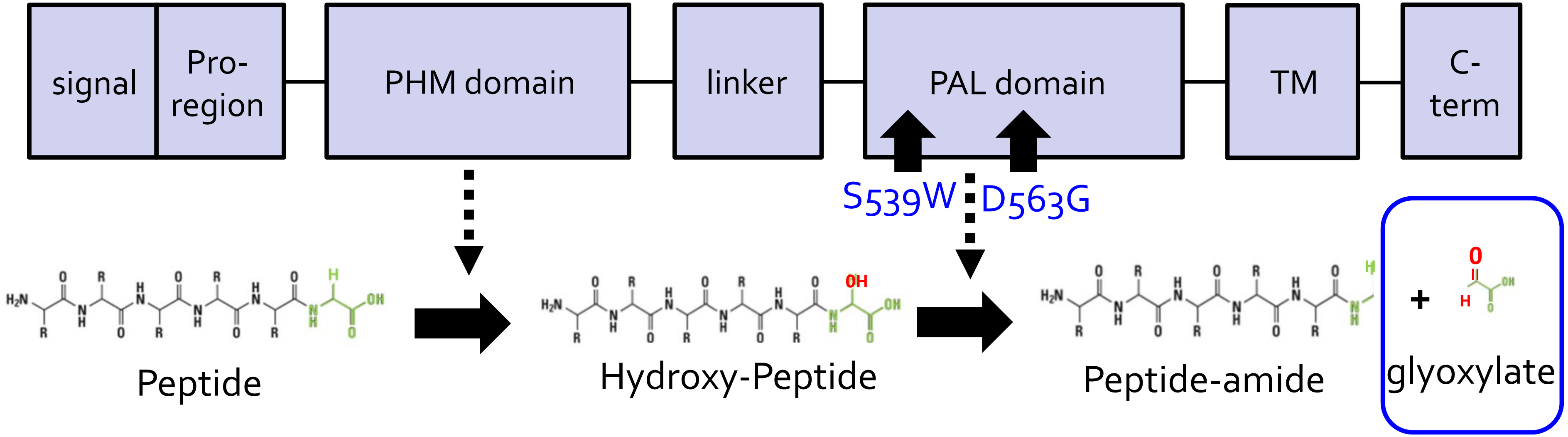
Exomes

~6500 T2D cases
~6500 controls
T2D-GENES
Multiethnic

+ 80,000 European samples typed for Illumina exome array
(+33,000 glycaemic traits)

Coding variants in *PAM* associated with T2D risk & beta-cell function

- Peptidylglycine alpha-amidating monooxygenase
  - Neuroendocrine secretory granules
  - Creates amide groups on glycine-extended peptide hormones ➔ biological potency
  - **1% variant [p.S539W] doubles T2D-risk**



signal

Pro-region

PHM domain

linker

PAL domain

TM

C-term

S539W

D563G

Peptide

Hydroxy-Peptide

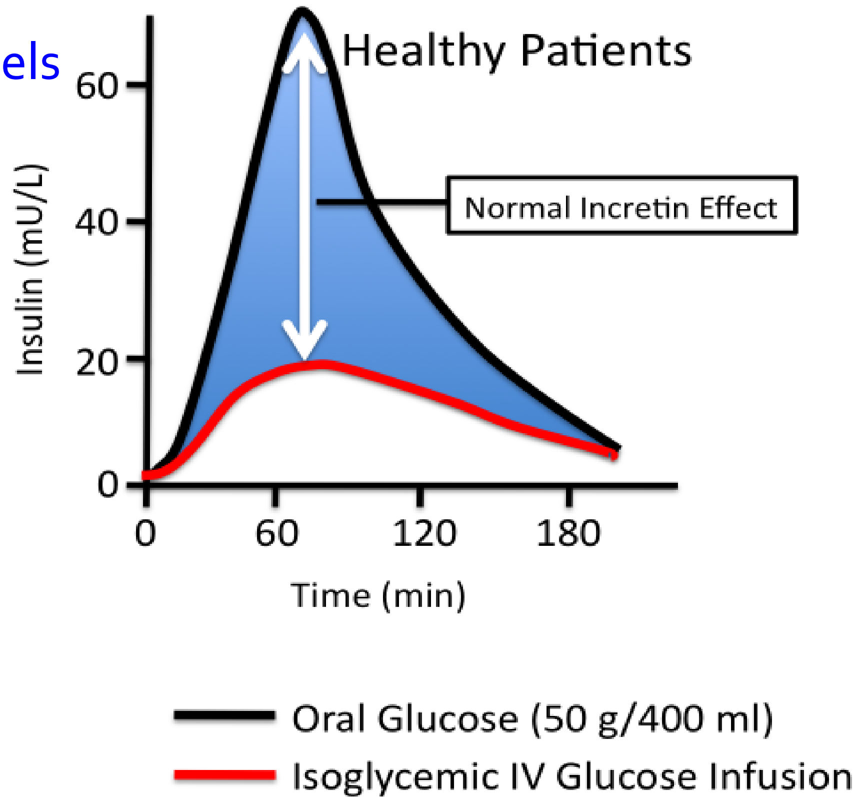
Peptide-amide

\+ glyoxylate

Hughye et al Nature Genetics 2013; Steinthorsdottir et al Nature Genetics 2014; T2DGENES Consortium in Revision

Extrinsic effects on beta-cell function

- Glucagon-like peptide 1 (GLP-1)
  - Lowers post-prandial blood glucose levels
  - Amidation increases GLP1 half life
- Recruit by Genotype (1% variant)

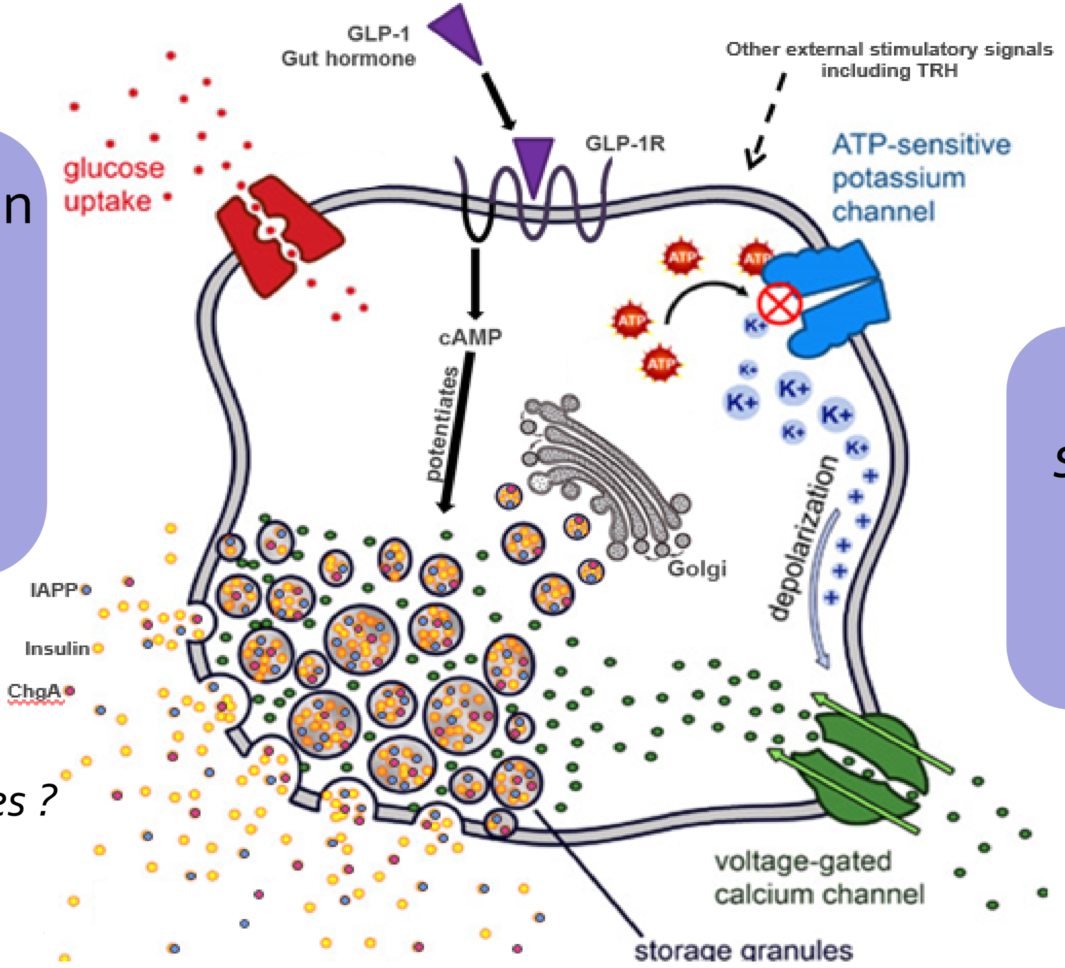


Torben Hansen, Jens Holst

Mechanisms for beta-cell dysfunction

Reduced amidated GLP-1 levels

Reduced insulin content and exocytosis & amidation of CgA



Reduced insulin packaging in granules ?

Patient stratification?

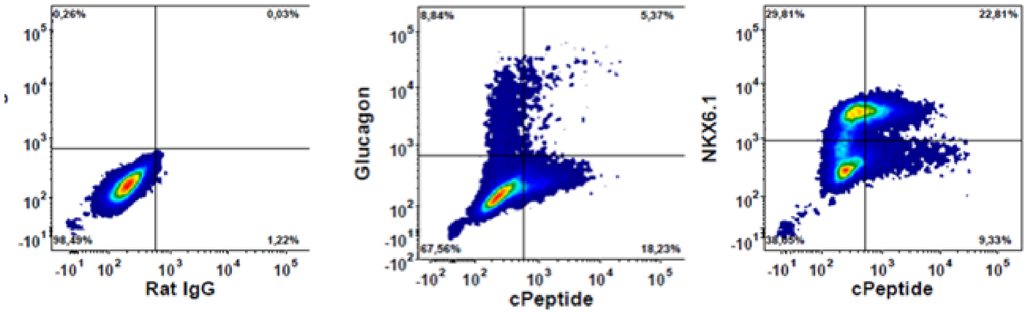
DPPIV inhibitors

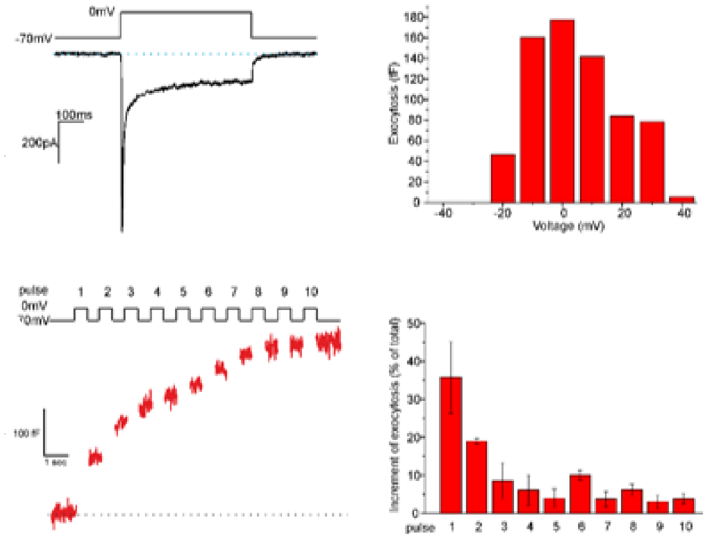
Image © 2004 Beta Cell Biology Consortium

New cell models new possibilities

stem BANCC
Stem Cells for Drug Discovery

- Human IPS cells
- Genome editing ➔ T2D risk alleles
- Differentiate ➔ “Beta like”-cells



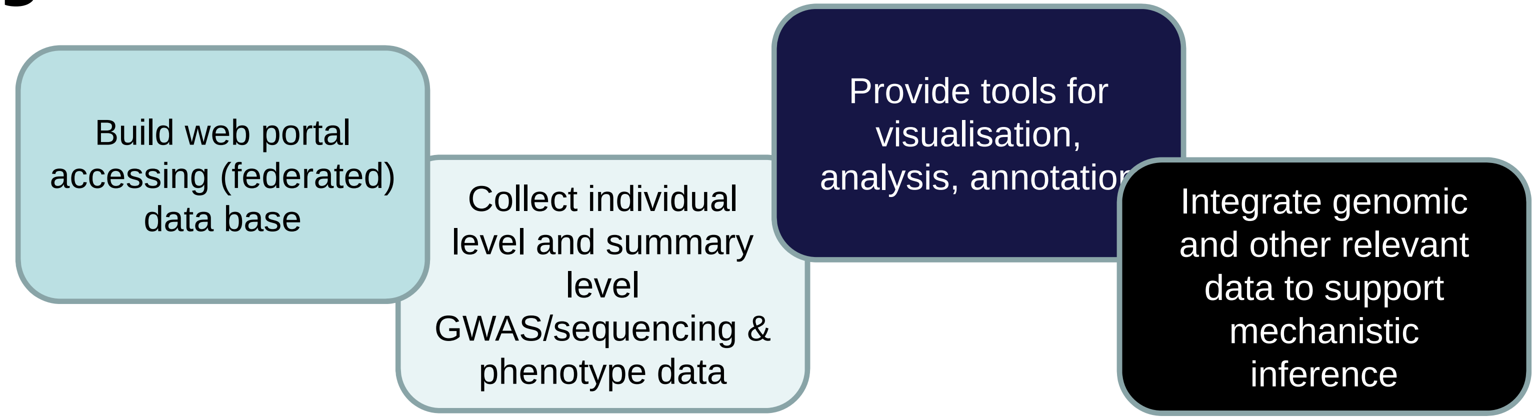


novo nordisk

Accelerating Medicines Partnership

- A partner for T2DSystems & for Rhapsody (IMI starting 2016)?
- Funded by NIH & pharma (precompetitive research partnership) ~$60M

- **Aim : to democratize access to T2D genetics & genomics data**



Build web portal accessing (federated) data base

Collect individual level and summary level GWAS/sequencing & phenotype data

Provide tools for visualisation, analysis, annotatio

Integrate genomic and other relevant data to support mechanistic inference

TYPE 2 DIABETES **GENETICS** beta

Explore data on a gene, variant, or region

examples: SLC30A8, rs13266634, chr9:21,940,000-22,190,000

Search for variants of interest

filter by p-value, odds ratio, predicted effect on protein, and more

View full GWAS results for a phenotype

type 2 diabetes

About the portal

This portal provides access to comprehensive results from 28 large human genetic association studies of type 2 diabetes and related traits that used sequencing or genotyping arrays. These data can shed light on biological functions of genes and point to possible drug targets. Click here to learn how to use the portal, or here to sign up for email alerts about new features. Click here to report bugs or provide feedback.

The following organizations provide funding and/or governance for this knowledge portal as part of the AMP T2D Program:

Funding and guidance are also provided by:

Please use the following citation when referring to data from this portal: AMP T2D-GENES Program, SIGMA; Year Month Date of Access; URL of page you are citing.

Projects

- So far ~20 awards (NIH and FNIH) for portal design, data aggregation, database federation, and genomic annotation from diabetes relevant tissues
- Oxford, San Diego, Boston/Montreal/UCLA, UNC/UMich, Wake Forest, Texas groups all involved in combining regulatory annotation data they have generated or are generating
- Developing an instance of the ENCODE database for data storage, which will (in time) be connected to the web portal
- Islet: RNA-Seq, ATAC-Seq, 3C/HiC, Massive parallel reporter & TF binding assays, ChIP-Seq.

Opportunities

- Some data will be available in both T2DSystems and AMP-T2D
- Some data will be represented in only one of these
- Opportunity to work towards interoperability of these
- T2DSystems would benefit in terms of additional genomic data, and access to most up-to-date genetic data.

Acknowledgements

NIH National Institute of Diabetes and Digestive and Kidney Diseases

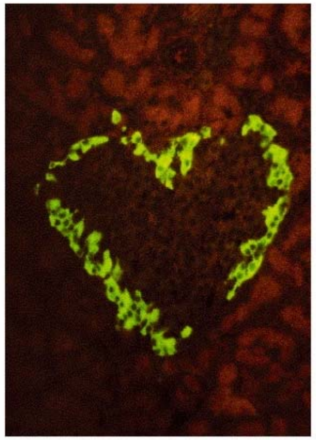
HORIZON 2020

Supported by wellcometrust

ISLET TRANSCRIPTOMICS

Martijn van de Bunt
Mark McCarthy
Kyle Gaulton
Amy Barrett
Amanda Bennett
Patrick MacDonald
Joss Manning-Fox
Patrik Rorsman
Lorna Whitty
David Buck
Manolis Dermitzakis
Ana Viñuela
Jorge Ferrer
Lorenzo Pasquali
Quin Wills
Chris Yau

Rory Bowden
Esther Mellado
Jim Hughes
Peter Donnelly
Doug Higgs
Ben Davies



DIAGRAM

Andrew Morris
Kyle Gaulton
Anne Raimondo
Nicola Beer
Jana Rundle
Søren Thomsen
Martijn van de Bunt
Michael Reschen
Chris O'Callaghan
Anil Chalisey
Mike Boehnke
David Altshuler
Mark McCarthy

imi Innovative Medicines Initiative

PAM

Anne Raimondo
Søren Thomsen
Benoit Hastoy
Anna Johnson
Torben Hansen
Oluf Pedersen
Jens Holst
Ewan Pearson
Fiona Gribble
Andrew Hattersley
Tim Frayling
Peter Ratcliffe
Norma Mason
Angus Jones
Mark McCarthy
Fredrik Karpe

MRC Medical Research Council

Oxford Biomedical Research Centre

DIABETES UK CARE. CONNECT. CAMPAIGN.

NHS National Institute for Health Research