Title
Integrative Genomics Analysis Reveals Tissue-specific Pathways and Gene Networks for Type 1 Diabetes

Permalink
https://escholarship.org/uc/item/0k8488ht

Author
Blencowe, Montgomery

Publication Date
2019

Peer reviewed|Thesis/dissertation
Integrative Genomics Analysis Reveals Tissue-specific Pathways and Gene Networks for Type 1 Diabetes

A thesis submitted in satisfaction of the requirements for the degree Master of Science in Physiological Science

by

Montgomery Charles Thomas Blencowe

2019
ABSTRACT OF THE THESIS

Integrative Genomics Analysis Reveals Tissue-specific Pathways and Gene Networks for Type 1 Diabetes

by

Montgomery Charles Thomas Blencowe

Master of Science in Physiological Science

University of California, Los Angeles, 2019

Professor Xia Yang, Chair

Type 1 diabetes (T1D) is a complex disease, involving a genetic predisposition that interacts with environmental triggers, leading to the loss of insulin producing beta cells in the pancreas. However, the molecular cascades underlying T1D are poorly understood and remain to be explored. We hypothesize that genetic risk factors of T1D perturb tissue-specific biological pathways and gene networks, which ultimately leads to the pathogenic end point in beta cells. We sought to identify the gene networks and key regulators for T1D by conducting a comprehensive, data-driven multi-omics analysis that integrates human genome-wide association studies (GWAS) of T1D, tissue-specific genetic regulation of gene expression in the form of expression quantitative trait loci (eQTLs), and tissue-specific gene network models using a computational pipeline Mergeomics. Our integrative genomics approach revealed immune pathways such as adaptive immune system, cytokines and inflammatory response, ZAP70 translocation, primary immunodeficiency and immunoregulatory interactions between a lymphoid and non-lymphoid cell, across various tissues. We also identified tissue-specific signals such as regulation of complement cascade in adipose tissue, macrophages, and monocytes, NOTCH signaling in adipose tissue and macrophages, protein folding, calcium signaling, chemotaxis and lysosomal pathways in the pancreas, adipose, and monocytes, and viral infection in macrophages and monocytes. Network modelling of these pathways highlights a number of key regulator genes such as GBP1, USP18, STAT1, RPL17, HLA genes (HLA-A, B, C, and -G), and
immunomodulatory genes (LCK, VAV1, ZAP-70), each of which has suggestive roles in the pathophysiology of T1D or other autoimmune disorders. Together, our integrative genomics approach offers comprehensive insights into the tissue-specific molecular networks and regulators as well potential between-tissue interactions underlying T1D, which may guide future development of therapeutic strategies targeting the disease.
The thesis of Montgomery Charles Thomas Blencowe is approved by:

Elaine Yih-Nien Hsiao

Stephanie Correa Van Veen

Xia Yang, Committee Chair

University of California, Los Angeles

2019
DEDICATION

This thesis is dedicated to my family and more particularly my parents for supporting my academic endeavours throughout.
# TABLE OF CONTENTS

ABSTRACT ................................................................................................................................. ii

COMMITTEE PAGE ................................................................................................................ iv

DEDICATION .............................................................................................................................. v

TABLE OF CONTENTS ................................................................................................................ vi

LIST OF FIGURES .................................................................................................................... vii

ACKNOWLEDGMENTS .............................................................................................................. viii

INTRODUCTION ...................................................................................................................... 1

MATERIALS AND METHODS .................................................................................................. 4

OVERVIEW OF STUDY DESIGN ............................................................................................... 4

T1D GWAS DATASETS ............................................................................................................ 5

MAPPING SNPS TO GENES ................................................................................................... 6

DATA-DRIVEN MODULES OF CO-EXPRESSED GENES ..................................................... 7

KNOWLEDGE BASED BIOLOGICAL PATHWAYS ................................................................ 8

MARKER SET ENRICHMENT ANALYSIS ............................................................................ 8

MERGING OVERLAPPING PATHWAYS INTO SUPERTS .................................................... 9

TISSUE SPECIFIC GENE REGULATORY NETWORKS AND KEY DRIVER ANALYSIS ............ 9

IN SILICO VALIDATION .......................................................................................................... 10

RESULTS .................................................................................................................................. 10

DISCUSSION ............................................................................................................................. 16

BIBLIOGRAPHY ....................................................................................................................... 33
LIST OF FIGURES

TABLE 1: TOP PATHWAYS ASSOCIATED WITH T1D IDENTIFIED ACROSS MULTIPLE TISSUES AT AN FDR < 5% .......................................................... 24

TABLE 2: T1D GENETIC SUPERSETS WITH KDs FOR BOTH BAYESIAN AND PPI NETWORKS ........................................................................................................ 25

FIG. 1: OVERVIEW OF STUDY DESIGN ................................................................................................................................. 26

FIG. 2: VENN DIAGRAM OF ENRICHED CANONICAL PATHWAYS AND CO- EXPRESSION MODULES FOR BOTH T1D GWAS COHORTS ................................................... 27

FIG. 3: HEATMAP OF THE TISSUE-SPECIFIC META-MSEA ENRICHMENT FROM THE COMBINED COHORT 1 AND COHORT 2 DATASETS FOR THE SUPERSETS DERIVED FROM THE CANONICAL PATHWAYS AND CO-EXPRESSION NETWORK MODULES ............................................................................................................................ 28

FIG. 4: TISSUE SPECIFIC KEY DRIVER SUBNETWORKS ................................................................................................................... 30
ACKNOWLEDGEMENTS

I would like to thank Dr. Xia Yang for the opportunity to join the lab as well as providing guidance and support throughout my master’s. I additionally would like to thank the lab members who have helped me develop my skillsets and my committee members for their time and input.
INTRODUCTION

Type 1 diabetes (T1D) is characterized as the autoimmune loss of pancreatic beta cells and resultant impairment of glucose homeostasis. Currently, T1D accounts for 5-10% of all diabetic individuals worldwide, approximately 40 million cases, with an increasing incidence rate of 2-3% per year (Maahs, West, Lawrence, & Mayer-Davis, 2010). Risk of developing T1D is increased by ~5.6% and ~50% with a diseased parent or diseased monozygotic twin, respectively, when compared to the general population. Parental heritability estimates predict that diabetic fathers confer an increased risk of T1D development of about 12% while mothers at around 6% (Steck et al., 2005). Thus, there is a strong genetic component predisposing T1D pathogenesis. However, T1D is not usually present in individuals with a family history, with only ~10% of patients having a first/second-degree relative with the disease, and a significant environmental contribution has been highlighted (Knip & Simell, 2012). Interestingly, there seems to be an increased disposition for developing T1D if one lives in regions of Northern Europe, independent of genetic background which is highlighted by an increase in T1D incidence in migrants living within these regions (Oilinki, Otonkoski, Ilonen, Knip, & Miettinen, 2012; Söderström, Åman, & Hjern, 2012). This environmental contribution has several potential manifestations, with alterations in gut microbiota (Kostic et al., 2015) or pre and post-natal dietary factors, including intake of gluten (Norris et al., 2003), vitamin D (Weets et al., 2004), and polyunsaturated fatty acids (Sørensen, Joner, Jenum, Eskild, & Stenc, 2012) as suggested factors. Furthermore, a longstanding hypothesis predicts that the exposure to viral infection may also be a causal factor (Gamble, Kinsley, FitzGerald, Bolton, & Taylor, 1969), particular enteroviruses (Coppeters, Wiberg, Tracy, & von Herrath, 2012) which seem to target pancreatic islet cells (Yeung, Rawlinson, & Craig, 2011). Therefore, both genetic and environmental components contribute to T1D incidence and progression, yet a large gap exists in understanding the complex genetic and environmental architectures as well as the interaction between the two.
From the genetics perspective, with the ease and power of Genome-Wide Association Studies (GWAS) we have uncovered ~60 T1D genetic risk loci. The main genes predisposing T1D patients are located within the HLA region on chromosome 6, encoding the major histocompatibility complex (MHC) which is critical for adaptive immunity. While HLA-encoding genes have the strongest association and account for up to 50% of the total genetic T1D risk (Lambert et al., 2004; Noble et al., 1996), loci outside of the HLA region including protein tyrosine phosphatase, non-receptor type 22 (PTPN22), the interleukin 2 receptor, alpha (IL2R-A), insulin gene (INS) and cytotoxic T-lymphocyte-associated protein 4 (CTLA4), have also been associated with disease development (Noble & Erlich, 2012).

While GWAS have played an essential role in identifying T1D-associated candidate genes, typically only the top loci reaching genome-wide association significance of p<5e-8 are reported, and these top loci cannot fully explain the total genetic heritability of T1D. Identifying the missing genetic risks, or the “dark matter”, is important to gain a full understanding of the genetic underpinnings of T1D pathogenesis. In addition, growing lines of evidence support an “omnigenic” disease model (Boyle, Li, & Pritchard, 2017), which states that a large proportion of genes on the genome may contribute to disease development through intimate gene-gene interactions in networks within and between tissues, and key network regulators likely play more central roles than other peripheral disease genes in the networks. More interestingly, GWAS loci for complex diseases have been found to be more concentrated in the periphery of gene networks and less likely to be network regulators. Therefore, simply focusing on the top GWAS hits will likely miss crucial regulatory genes, and understanding how T1D genetic risks interact in gene networks and identifying key regulators will offer novel insights
into T1D pathogenesis and help prioritize regulators for targeted therapeutic strategies (Chan et al., 2014; Chella Krishnan et al., 2018; Makinen et al., 2014; Shu et al., 2017).

Integration of GWAS data with functional data such as tissue-specific expression quantitative trait loci (eQTLs) and gene networks has proven to be a powerful tool to pinpoint causal genes and their associated pathogenic mechanisms and regulators within the context of particular tissue type (Chella Krishnan et al., 2018; Foroughi Asl et al., 2015; Hauberg et al., 2017; Makinen et al., 2014; Shu et al., 2017; Zhong, Yang, Kaplan, Molony, & Schadt, 2010). Our lab has previously developed a computational pipeline, Mergeomics, to facilitate such analysis (Shu et al., 2016). We hypothesize that T1D genetic risks with a wide spectrum of effect sizes (strong, moderate to subtle) interact and perturb tissue-specific gene networks through a select set of regulatory genes, resulting in variations in T1D susceptibility, and application of Mergeomics to T1D GWAS datasets in conjunction with other orthogonal functional genomics data will help reveal these networks and regulators.

Though the integration of the largest T1D GWAS studies to date with tissue-specific functional information such as genetics of gene expression and gene regulatory networks, our analysis confirmed the importance of immune pathways across various tissues, such as adaptive immune system, cytokines and inflammatory response, ZAP70 translocation, primary immunodeficiency, and immunoregulatory interactions between a lymphoid and non-lymphoid cell. Signals demonstrating certain tissue specificity include regulation of complement cascade (adipose, macrophage, and monocytes), NOTCH signalling (adipose tissue and macrophages), chemotaxis and lysosomal pathways (pancreas, adipose tissue, and monocytes), and genes involved in viral infection/interferon signaling (macrophages and monocytes). The finding related to infection suggests a potential interaction between genetic predisposition and environmental perturbation. Importantly, we found
key genes associated with pathways that are not directly associated with immune function, such as spliceosome, proteasome, calcium, Wnt, and cell cycle related pathways, however they may have modulatory or interactive effects on the immune system within the context of T1D. Of these, we find cell cycle to be enriched in the pancreatic tissue which highlights the potential for tissue specific dysregulation of this key cellular process directly leading to a local environment which promotes beta cell death. These results provide us with a novel area to further explore in disease progression and exploit as potential therapeutic targets.

MATERIALS AND METHODS

Overview of study design
We utilized an integrative genomics approach that leverages multiple large-scale human genetic and genomic datasets to elucidate the genetic networks and regulators of T1D pathogenesis (Figure 1). The datasets utilized included T1D GWAS from two independent cohorts, tissue specific eQTLs from diverse human tissues or cell types, various network models including gene coexpression networks, Bayesian gene regulatory networks and protein-protein interaction networks, and biological pathway information (detailed in subsequent sections). To address reproducibility, we ran the integrative analysis on each GWAS study independently and then focused on the findings that are consistent between the two cohorts. For each GWAS, we mapped the single nucleotide polymorphisms (SNPs) to genes using tissue/cell-specific eQTL data. The use of eQTLs helps inform on the most likely genes affected by GWAS SNPs based on functional evidence. Next, we grouped the genes based on whether they belong to the same biological pathways or show coexpression, which indicates functional relevance, in data-driven gene co-expression networks. We then assessed which pathways or gene coexpression modules (a module contains genes that show coexpression patterns) demonstrated
stronger genetic associations with T1D compared to randomly generated gene sets using a Marker Set Enrichment Analysis (MSEA). After carrying out the MSEA process for each T1D GWAS dataset, we subsequently used a Meta-MSEA to meta-analyze the two independent GWAS data sets to look for shared pathways/modules that showed significant T1D associations, which we further simplified into independent “supersets” to reduce redundancy between pathways/modules. Integrating these T1D supersets with gene regulatory networks (Bayesian) and protein-protein interaction networks, we carried out the weighted key driver analysis (wKDA) to identify key drivers (KD$s$), which are central network genes whose network neighborhoods are highly enriched for genes in the T1D pathways and coexpression modules. These KD$s$ were then visualized in tissue-specific networks. Furthermore, we carried out in silico validation via literature mining (T-HOD, PolySearch 2.0, DisGeNET) to seek if these key drivers are novel, suggestive, or known to be linked with T1D (Cheng et al., 2008; Dai, Wu, Tsai, Pan, & Hsu, 2013; Pinero et al., 2017).

**T1D GWAS datasets**

The summary statistics of GWAS for T1D was obtained from the JDRF/Wellcome Diabetes and Inflammation Laboratory, University of Oxford (Barrett et al., 2009; Cooper et al., 2017).

The study is comprised of 5913 T1D individuals of European descent (Barrett et al., 2009; Cooper et al., 2017), among which 3983 were genotyped using Illumina HumanHap550v3 (550k) Infinium Beadchip from the UK GRID and 1930 T1D individuals genotyped using Affymetrix 500K from the WTCCC. There were a total 8828 Controls, with 3999 genotyped using Illumina HumanHap550v3 (550k) Infinium Beadchip from the 1958 Birth Cohort (1958BC), 1490 genotyped using Affymetrix 500K (1958BC), 1455 genotyped using Affymetrix 500k from the UK Blood Services (UKBS) and 1884 genotyped using Affymetrix 500K from a cohort of Bipolar disorders.
Inclusion criteria for the UK GRID are: T1D diagnosed between 6 months to 16 years of age, insulin dependent for greater than 6 months, a UK resident and self-identified as white European (average diagnosis age = 7.8 years of age, SD = 3.9, 47% female). Inclusion criteria for the WTCCC are: T1D diagnosed less than 17 years of age, insulin dependent for greater than 6 months and self-identified white European (average diagnosis age = 7.2 years of age, SD = 3.8, 49% female). Control inclusion criteria for UKBS and 1958BC included being residents in the UK and self-identified white Europeans. For the Bipolar cohort, control individuals greater than 16 years old and resident in the UK and of European descent were included.

The above T1D and control individuals were partitioned into two independent cohorts based on matching genotyping platforms (i.e., Illumina or Affymetrix) between cases and controls. Cohort 1 was comprised of 1930 T1D patients and 4830 Controls. Cohort 2 was comprised of 3983 T1D patients and 3999 Controls.

SNPs genotyped were imputed to ~10 million SNPs (1000 Genomes Phase III) using IMPUTE2, and routine quality controls were conducted as described in Cooper et al. (Cooper et al., 2017). Statistical association between SNPs and T1D was carried out using a Bayesian analysis. All statistical association p values for all imputed SNPs that passed quality control, regardless of significance level for T1D association, were used in our downstream analyses.

**Mapping SNPs to genes**

In order to link GWAS SNPs to their potential target genes, tissue-specific eQTLs were used as they can provide functional insight for the role of SNPs in gene expression regulation within a given tissue.
Thirteen eQTLs sets were obtained from the GTEx database including subcutaneous adipose, visceral omentum adipose, blood, brain, colon, heart, liver, lymphocyte, muscle, pancreas, pituitary, spleen and stomach (Consortium, 2015). Additionally, we obtained macrophage and monocyte eQTLs from the Cardiogenics Consortium (Rotival et al., 2011). A broader spectrum of tissues was considered at this step to help objectively infer which tissues might be more informative for T1D association. GWAS was mapped to each tissue eQTL set separately to derive individual SNP-gene mapping sets reflecting tissue origins to allow assessment of tissue-specific signals.

A high degree of linkage disequilibrium (LD) was observed in the eQTL data, which may cause biases in the downstream analysis. For this reason, we removed redundant SNPs that had LD of $r^2 > 0.7$ with a chosen SNP. Briefly, a GWAS SNP was compared against other SNPs for LD and T1D association. If the SNP was in LD of $r^2 > 0.7$ with other SNPs, the one with the strongest T1D associations was chosen. This process was repeated until all remaining SNPs were not in LD based on the $r^2 > 0.7$ cut-off. These non-redundant SNPs were used for downstream analyses.

**Data-driven modules of co-expressed genes**

In order to assess whether T1D GWAS signals are enriched in specific gene subnetworks, we derived gene co-expression networks from transcriptomic data sets from genomic studies of subcutaneous adipose, visceral omentum adipose, blood and pancreas (Shu et al., 2016). These tissues were chosen due to their relevance to T1D. The WGCNA (Weighted Gene Correlation Network Analysis) package was used to reconstruct co-expression networks based on gene expression profiles (Langfelder & Horvath, 2008). Each tissue network contains multiple “modules” and each module is comprised of tens to hundreds or thousands of genes that show coexpression. A total of 272 co-expression modules were curated.
Knowledge-based biological pathways

We used a total of 1827 canonical pathways from Reactome (Version 45) (Croft et al., 2014), Biocarta (Nishimura, 2001) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) databases (Kanehisa & Goto, 2000). In addition to the knowledge-based pathways, we constructed a T1D positive control gene set based on candidate causal genes curated in GWAS catalog (p<5.0E-8) (MacArthur et al., 2017). Similar control gene sets for coronary heart disease (CHD), type 2 diabetes (T2D), and height were also constructed to compare with the T1D positive control set.

Marker Set Enrichment Analysis (MSEA)

To identify co-expression modules and pathways that show evidence for genetic association with T1D, we applied MSEA from the Mergeomics (Shu et al., 2016) package on each of the GWAS cohorts separately in conjunction with the eQTL sources. MSEA employs a chi-square-like statistic with multiple quantile thresholds to assess whether a co-expression module or pathway shows enrichment of functional disease SNPs (i.e., those likely regulate gene expression as captured in eQTLs) compared to random chance. 10,000 permuted gene sets were generated for each co-expression module and pathway. As detailed in Shu et al., the enrichment statistics from the permutations were used to approximate a Gaussian distribution from which enrichment P-values were determined. Benjamini-Hochberg (BH) false discovery rate (FDR) was estimated across all co-expression modules and pathways tested for each GWAS. Gene sets were considered to be statistically significant if FDR < 5% in at least one SNP-gene mapping set. To evaluate gene sets across both GWAS studies, we followed up with a meta-analysis at the module/pathway level using the meta-MSEA function in Mergeomics, to retrieve robust gene sets across both cohorts. Stouffer’s Z score method was used to calculate meta P-values based on the P-values from the multiple MSEA runs. Meta-FDR was
calculated using the Benjamini-Hochberg method, as described above.

Merging overlapping pathways into supersets

The curated pathways and gene coexpression modules may carry redundant information. For example, a KEGG pathway “insulin signalling” can have largely overlapping genes with a Reactome pathway “insulin receptor signalling”. To reduce redundancy, we compared the significant modules and pathways associated with T1D at FDR <5% and merged the overlapping ones using a merging algorithm in Mergeomics to produce independent, non-overlapping “supersets”. The algorithm employs an overlap ratio \( r \) between two gene sets \( A \) and \( B \) as \( r = (r_{AB} \times r_{BA})^{0.5} \), where \( r_{AB} \) is the proportion of genes in \( A \) that are also present in \( B \) and \( r_{BA} \) is the proportion of genes in \( B \) which are also in \( A \). The overlap ratio cut-off was set to \( r \geq 0.33 \) and Fisher’s exact test was used for assessing the statistical significance of gene overlap between modules/pathways. BH FDR < 5% was considered significant. Resultant supersets containing more than 500 genes were trimmed down to contain core genes shared among the overlapping gene sets.

Tissue-specific gene regulatory networks and key driver analysis (KDA)

Tissue-specific Bayesian gene regulatory networks of adipose, blood, and pancreas tissue were obtained (Shu et al., 2016) as well as protein-protein interaction (PPI) networks obtained from the Human protein reference database (Keshava Prasad et al., 2008). We chose to focus on these tissue networks due to our MSEA results showing the strongest statistical significance for these tissues. With these networks, we performed a key driver analysis using a KDA algorithm in Mergeomics to identify potential key drivers (KDs) whose network neighbors are enriched for genes within the T1D associated supersets uncovered by MSEA. The algorithm employed a chi-square like statistic similar to that described for MSEA, and FDR < 5% was used to focus on top robust KDs.
In silico validation of KDs

To determine the relevance of the top ranked KDs to T1D, we used literature-mining methods such as PolySearch 2.0 (Liu, Liang, & Wishart, 2015), DisGeNET (Pinero et al., 2017) and T-HOD (Dai et al., 2013).

RESULTS

Identification of T1D-associated knowledge-based biological pathways

We first assessed which knowledge-based biological pathways were enriched for T1D GWAS signals. The use of tissue-specific eQTLs served to guide SNP-to-gene mapping, allowing us to capture tissue specific results. From the 15 eQTLs sets used, we found the following six to be the most informative in terms of whether and how many significant pathways were identified: blood, macrophage, monocyte, pancreas, subcutaneous adipose and visceral omentum adipose. We therefore focused on reporting the results from these six informative tissues only, with other tissues serving to supplement the main results.

Out of the 1827 curated canonical pathways, we identified 224 pathways enriched from Cohort 1 at an FDR <5%. Of these 224 pathways, 154 were significant at FDR <5% in at least one of the six chosen tissues based on MSEA and the remaining 70 significant pathways passed FDR <5% when meta-analyzed across tissues using meta-MSEA. The top pathways enriched from Cohort 1 included T1D positive control group, CHD positive control, Adaptive Immune system, G-protein coupled receptor related processes, and positive control set for Height. In the remaining 70 pathways that collectively passed meta-MSEA significance across tissues, Influenza life cycle, Nuclear envelope breakdown and PLC-gamma-1 signalling showed top significance. From the Cohort 2 dataset we
found a total of 158 pathways enriched at an FDR <5%, with 129 pathways significant in at least one tissue in MSEA and the remaining 29 passed FDR<0.05 in meta-MSEA across tissues. The top pathways were confirmatory with Cohort 1, with T1D, CHD, and Height positive control sets all showing strong statistical significance. Additionally, Adaptive Immune system, interactions between lymphoid and non-lymphoid cells, ER-Phagosome pathways, Antigen processing, Wnt signaling, transfer in endocytosis and recycling, and SNARE interactions in vesicular transport were our top pathways.

When comparing Cohort 1 and Cohort 2 for shared pathways, as expected we found largely confirmatory and overlapping pathways heavily related to immune processes and importantly the T1D positive control gene set. Between the two datasets we found 111 pathways significant in both, 113 unique for Cohort 1 and 47 unique for Cohort 2 (Figure 2A). Some unique pathways for Cohort 1 include Influenza life cycle, HIV life cycle, and nuclear envelope breakdown. For Cohort 2 the unique pathways included purine and inositol phosphate metabolism and RNA degradation.

We found largely confirmatory results with the other tissue types tested across both Cohorts, with the spleen, adrenal, liver and lymphocyte revealing largely immune related signals. The pituitary gland also showed significance in Wnt signaling and insulin receptor signalling.

Identification of T1D associated gene co-expression modules

The above biological pathways were curated based on known knowledge. To increase the chance to uncover novel biology, we collected data-driven gene coexpression modules which reflect functionally related gene sets based on gene expression patterns. This approach has previously helped derived
novel biological insights into various diseases (Chella Krishnan et al., 2018; Makinen et al., 2014; Shu et al., 2017).

From our WGCNA networks we generated 272 co-expression modules, with which we ran our Cohort 1 and Cohort 2 data sets independently. Cohort 1 showed enrichment in a total of 40 unique modules of which 36 modules passed an FDR <5% in at least one tissue and four modules were deemed significant collectively across the six tissues (Figure 2B). The top annotated modules for Cohort 1 included interferon signaling, translocation of ZAP70, Stress related pathways, destabilization of mRNA, focal adhesion, and diabetes pathways. On the other hand, the Cohort 2 dataset showed significance in 46 modules (FDR <5%), with 37 significant in at least one tissue and 9 collectively across the six tissues (Figure 2B). The top modules included cytokine signaling, interferon signalling, antigen processing, DNA repair, metabolism, pre-NOTCH processing in golgi, and Pitx2 pathway.

When comparing the two cohorts we found that the majority of the modules reach significance in both, with a total of 30 overlapping modules. The top modules found in both were consistent with immune related pathways such as interferon signaling, and Antigen processing/presentation. In terms of unique pathways, we saw 16 unique modules for Cohort 2 including mTOR pathways, extracellular matrix (ECM) organization and DNA repair. In Cohort 1 we find 10 unique modules consisting of AKT signaling, complement pathway and ion transport (Figure 2B).

Comparing these 30 overlapping modules to the 111 canonical pathways uncovered in our pathway-based analysis, we see similar signals enriched across both. These signals include complement and
coagulation cascade, DNA repair, cell cycle, NOTCH signaling, metabolism of amino acids and derivatives, transcription, and translation.

**Merging of pathways and co-expression modules into independent supersets**

We chose to focus on the shared 141 significant gene sets uncovered in our meta-analysis between Cohort 1 and Cohort 2 (30 co-expression modules and 111 canonical pathways) as they reflect reproducible signals for T1D association (Figure 2A, 2B). As the co-expression modules and pathways were obtained from various sources, the gene sets may share a high number of overlapping gene members. To decrease the redundancy, we merged 83 overlapping gene sets into 12 independent supersets (Table 1). Interestingly, canonical pathways tend to merge with canonical pathways and coexpression modules tend to merge with coexpression modules, suggesting different biological properties of the two types of gene sets. The supersets represented diverse biological pathways including activation of adaptive immune system, antigen processing, ubiquitination & proteasome degradation, HIV infection, protein folding, RNA Polymerase I/III, mitochondrial transcription, signaling by GPCR, signaling by NOTCH, signaling by the B-cell receptor (BCR) and tRNA aminoacylation. The other 58 non-overlapping gene sets were kept intact, producing a total of 70 non-overlapping supersets.

**Second round of MSEA on the 70 supersets**

After using our merging algorithm of similar pathways, we ran a second round of MSEA to confirm whether our merged modules retained significance for T1D association. We confirmed 40 supersets derived from canonical pathways to show statistical significance in our combined Cohort 1 and Cohort 2 datasets (Figure 3A). Our results from the monocyte and macrophage-based analysis showed the strongest and most consistent enrichment of pathways including Natural Killer Cell mediated
cytotoxicity, lymphoid and non-lymphoid cell interaction, and mitotic G1/S phase. The adipose pathways included systemic lupus erythematosus, complement cascade, Natural Killer Cell mediated cytotoxicity and signaling by NOTCH. The blood tissue-derived pathways contained ABC transporters, Antigen processing and presentation, Secretin receptors, cell cycle and neuroactive ligand-receptor interaction. The pancreas showed pathways involved in Natural Killer Cell mediated cytotoxicity, secretin receptors, protein folding, calcium signaling and Wnt signaling. When looking at the remaining 30 supersets derived from our co-expression modules (Figure 3B). Again, monocyte and macrophage were informative across most of the listed supersets (Figure 3B). Pancreas and subcutaneous adipose tissue were also informative for 14 and 25, respectively, of the 30 supersets (FDR <5%). The blood appeared to be the least informative for these supersets, with only a few supersets related to DNA replication, interferon signaling, mitotic cell cycle, and TCR signalling being significant. Across all six tissues we found that translocation of ZAP70, calcium signaling, interferon signalling, and Stress pathways were significant.

Identification of central regulators for T1D via a weighted key driver analysis

To identify central regulatory genes, or key drivers (KDs), among the potential T1D associated supersets, we performed a weighted key driver analysis (wKDA) on our independent supersets using tissue-specific Bayesian gene regulatory networks as well as using the PPI networks (Figure 4).

The total number of unique KDs found from our Bayesian network included 299 across blood, monocyte, macrophage, adipose and pancreatic islet networks at an FDR <5% (Supplement Table). To focus on the most central regulators, for each superset in each tissue-specific network, we chose the top five ranked KDs satisfying an FDR <5%, to derive 60 KDs, of which 13 were shared among >=2 supersets and two (RPS29 and RPS18) were shared across the adipose, blood, monocyte and
macrophage networks. In terms of tissue specific KDs we found that the adipose network contained
the most with 45 KDs, and 13 unique KDs were shared between blood, monocyte and macrophage
networks (Supplement Table). There were no KDs discovered for our pancreatic tissue Bayesian
network, potentially due to the sparseness of this network. The key drivers uncovered can generally
be categorized under cell cycle, metabolism and immune related pathways. Interestingly, many of the
KDrs have been found to be involved in viral infections, autoimmune and childhood onset disease, all
of which are associated with T1D. These KDs include RPS29, RPLP2, CD19, OAS2, IFIH1 and
PTPN6.

To expand our search for KDs and look for consistencies between networks, we ran the PPI network
for adipose, blood, monocyte. macrophage and pancreatic tissues. Focusing on the top five ranked
KDrs in each tissue specific network within each superset, we found a total of 113 KDs, five specific
to adipose (CASP9, FLNA, BRF1, CD74, HLA-DRA), three specific to blood (PTPRC, MYC,
GNA11), one for the pancreas (JAK1), two for monocyte (AKT1, ESR1) and two for macrophage
(POLR2G, PROS1).

When comparing the PPI network with the Bayesian network we find 17 overlapping KDs for tissue
specific networks (15 for adipose, two for blood) including complement related genes such as CD19
and CD74, DNA replication genes such as MCM2 and MCM6, as well as those linked to autoimmunity
such as STAT1, SERPINE1 and GZMB. Those shared in the Top 5 KDs (FDR <5%) of the same
supersets for both the Bayesian and PPI networks include LCK, VAV1, PTPN6 for Natural Killer
Cell Cytotoxicity and F2 and PLG for Complement and Coagulation pathways (Table 2).

In silico validation of the Key Drivers
To determine the relevance and functional significance of the unique KDs from our Bayesian and PPI network analysis to T1D from our study, we cross-validated the genes by performing a comprehensive *in silico* analysis using three literature mining methods: PolySearch 2.0, DisGeNET, and T-HOD.

Our search yielded 6 KDs that are known T1D GWAS hits (*HLA-DRB1, IFIH1, PGM1, HLA-DQB1, HLA-B* and *HLA-DR-A*) as well as 43 KDs such as *GC, PLG, LCK, STAT1* and *PPARG* which were suggested to be associated with T1D due to previously being implicated in one or more studies to play a potential role in T1D pathogenesis.

Our study further uncovered a total of 107 potential novel KDs for T1D, many of which have viral infection associations (*OAS2, IFIH1, ISG15, SLC15A3* and *RTP4*). Additionally, we find large overlap of KDs with genes previously associated with autoimmune diseases such as Lupus Erythematosus, Psoriasis and Rheumatoid Arthritis (*PTPN6, MCM2, LCK, OAS2* and *CD3G*).

**DISCUSSION**

At present, genetic studies have uncovered >60 loci linked with T1D development, yet our understanding of the intricate mechanisms underlying these associations is still lacking. In addition, GWAS loci alone cannot provide the level of insights required for intervention. To this end, we utilised a multi-omics integrative approach to advance our understanding of T1D etiology and prioritize potential therapeutic targets among the large numbers of disease associated signals. Through the use of multiple T1D GWAS data sets, functional genomics data represented as tissue-specific eQTLs, knowledge-driven pathways, and data-driven networks, the perturbed pathways and key regulators which potentially drive the etiology of T1D were further elucidated.
This computational strategy confirmed known genes and biological pathways in T1D and revealed novel genes and their corresponding networks and biological processes. Among the biological networks which are perturbed to initiate or exacerbate T1D, are a significant number of immune and apoptosis related processes which are known to be involved in T1D, and several novel pathways such as viral infection, NOTCH signaling, Wnt signaling, protein folding, and calcium signaling. Our tissue specific analyses reveal consistent enrichment of antigen processing/presentation and INF-α/β/γ signalling across tissues, supporting the immune origin of T1D highlighted in the literature and emphasizing the presence of systemic dysregulation of immune function beyond insulitis. For the multi-tissue immune pathways for T1D, top KDs predicted by our analyses are HLA-DR/DQ alleles (Hu et al., 2015). Although HLA genes have been known to confer the greatest genetic risk for T1D, our network analysis uniquely highlights the critical regulatory roles of these genes in T1D development.

Of the immune related pathways, a notable process which shows perturbation in macrophage and monocyte is ubiquitination and proteasome degradation within the context of Class I MHC mediated antigen processing. Many of the key drivers which were highlighted in our analysis relate directly to the immunoproteasome and the maturation of the MHC I complexes, including GBP1 (Lundberg, Krogvold, Kuric, Dahl-Jorgensen, & Skog, 2016), P3MA6 (Sjakste et al., 2016), and P3MA4 (Jin et al., 2013), each of which have been previously implicated in T1D susceptibility. Additionally we identified several HLA genes, including HLA-A, -B (Nejentsev et al., 2007), -C (Genetic Analysis of Psoriasis et al., 2010), and -G (Eike, Becker, Humphreys, Olsson, & Lie, 2009), which have been previously associated with T1D or autoimmune disease susceptibility. Dysregulation of the MHC class I binding peptides produced by the cytosolic multicatalytic proteasome may play a fundamental role in the development of T1D causing autoimmunity. These MHC class I molecules are a major inducer
of the host immune response and alterations in MHC I peptide presentation have a putative role in autoimmune pathogenesis, with links to ankylosing spondylitis (Evans et al., 2011), multiple sclerosis (Guerini et al., 2012), and Crohn’s disease (Barrett et al., 2008). Previous evidence also supports the role of the immunoproteasome in the induction of T1D through the generation of aberrant polypeptide derived epitopes which act as β-cell neo-autoantigens (Thomaidou, Zaldumbide, & Roep, 2018), which is another potential avenue of disease development.

Furthermore, a number of the identified KDs relate to the modulation of immune response and autoimmunity, most notably IFIH1, a DEAD box RNA helicase which has been linked to T1D susceptibility. Additionally, we identify JAK1 via our PPI KDA, which has previously been shown to regulate pediatric autoimmunity and local inflammatory response via its interactions with STAT and INF signaling, thus acting as an intriguing target for further study. Moreover, the inhibition of JAK1 levels via AZD1480 blocked MHC class I upregulation via cytokine signaling within both mouse and human beta cells (Trivedi et al., 2017). Through this, the infiltration of immune cells into the pancreatic islets was dramatically slowed, as the T cells were no longer able to directly associate with beta cells, protecting NOD mice from insulitis. These results affirm the previous predictions made regarding the etiology of T1D, with significant emphasis on the failure of the immune system to recognize beta cell autoantigens as self, thus resulting in the destruction of these cells. Our results here further emphasize the potential of these immune related genes as central regulators and targets to ameliorate the pathogenesis of T1D.

Our results also support the importance of T cells and cytokine signalling in T1D pathogenesis. FYN (Figure 4A) is one of the most central key regulators in the T1D tissue networks and is connected a number of previously identified T1D associated GWAS loci (IL7R, MAPT, TYK2). Importantly, this
gene has been shown to be essential in T cell signaling and related processes, through its interaction with ZAP-70 (Michel, Grimaud, Tuosto, & Acuto, 1998) and VAV1 (Garcia-Bernal et al., 2005), which are both key components of T cell mediated immune response and have been identified as key drivers in our network analysis. The role of T-cell signaling and autoimmunity within the context of T1D progression is further highlighted by KD LCK (Stanley, Trivedi, Sutherland, Thomas, & Gurzov, 2017), which has also been previously shown to preclude autoimmunity through its interaction with DUSP22, serving as a negative regulator of T-cell activation. Also, the KDA identifies PTPN6, which has been shown to modulate cytokine signaling within pancreatic beta cells, in coordination with the action of PTPN1 to connect the beta cells with the immune system. Finally, several other KDs were noted using the Bayesian network blood tissue data, including SEMA4A (Chapoval, Vadasz, Chapoval, & Toubi, 2017) and TCF7 (Noble et al., 2003). SEMA4A, coding for Semaphorin 4A, has been shown to play a role in a number of autoimmune diseases including multiple sclerosis, Systemic lupus erythematosus, and Rheumatoid arthritis through its action in activating regulatory T cells (Delgoffe et al., 2013). Furthermore, it may have a potential role in inhibiting angiogenesis which may further perpetuate the T1D disease progression (Toyofuku et al., 2007). Similarly, polymorphisms within TCF7, a T cell transcription factor, have also been identified as being associated with T1D and may be an essential locus in the etiology of T1D through its downstream action of modulating immune response. Our analysis thus provides confirmation for the importance of T cell genes and pathways, as well as pinpointing the essential regulators to explore further within the context of T1D.

More globally, we identified several pathways which imply a pathogenic change in the molecular machinery coordinating protein production and processing, including spliceosome, mRNA metabolism, tRNA aminoacylation, gene expression, translation, and proteasome in our meta-MSEA. Kracht et al. reported that the production of a non-conventional products due to mRNA processing errors produces an autoimmune polypeptide that is detected by T cells in T1D patients (Kracht et al.,
It is also predicted that the formation of hybrid insulin peptides within beta cells activates CD4 T cells in NOD mice (Delong et al., 2016), thus further supporting the association between T1D and the failure of proteins to be processed correctly. It is plausible that variations within genes governing protein formation and processing components induce antigenic protein products within the pancreatic tissue itself or within immune cells, which results in the activation of an autoimmune response, beta cell death, and development of T1D.

Viral infection (Allen, Kim, Rawlinson, & Craig, 2018) and its association with T1D has been suggestive as the potential causal environmental trigger, particularly with regard to antenatal maternal infection and subsequent incidence of T1D. In support of this theory we found a number of pathways associated with HIV and influenza virus infection across multiple tissues tested. Given that these pathways are genetically perturbed as informed by our T1D GWAS datasets, our finding implies that genetic variants in genes involved in viral infection may confer vulnerability to infections and/or promote over-reactive viral response that induces autoimmunity, which explains how viral infection may trigger T1D pathogenesis. We have identified several KDs which may be susceptible to viral perturbation resulting in modulation of immune response, which includes PSMA7 (Apcher et al., 2003), TRAF6 (Yoboua, Martel, Duval, Mukawera, & Grandvaux, 2010), CD3G (Willard-Gallo, Furtado, Burny, & Wolinsky, 2001), PIK3R2 (Ylosmaki, Schmotz, Ylosmaki, & Sakse, 2015), and MCM6 (Gautier et al., 2009). Of these, proteasomal alpha subunit PSMA7 enables the human immunodeficiency virus-1 (HIV-1) Tat protein to disrupt proteasome function through its interaction with the subunit, thus highlighting a potential region which is susceptible to viral modulation (Apcher et al., 2003). The association with this protein and potential T1D pathogenesis goes further as it may modulate both cellular stress response (Li, Zhang, Zhang, & Wei, 2011) and antigen processing (Du et al., 2009), which may collectively impact beta cell survivability. Also, TRAF6, a TNF receptor, has
been previously shown to modulate viral load in hepatitis C virus patients (Pu et al., 2017) as well as the immunogenic inflammatory response in dengue virus infected patients (Motaleb, Nabih, Mohamed, & Elhalim, 2017), thus serving as a connection between viral infection and immune response (Fang et al., 2017). Additionally, there is evidence suggesting that viral infection alters the expression and local chromatin landscape of the T cell receptor complex subunit CD3G (Akl et al., 2007; H. Yu et al., 2015), with further reports associating this subunit with autoimmunity (Tokgoz et al., 2013). The other two predicted key driver genes, MCM6 (Tan et al., 2017) and PIK3R2 (Kim, Hollenbaugh, Kim, & Kim, 2011) are predicted to influence antiviral and inflammatory response after viral infection.

Another immune-related pathway is the SNARE interaction with vesicular transport pathway (Offenhauser et al., 2011), which is significant in that the protein complex seems to act as a negative regulator of macrophage activity and is associated with familial hemophagocytic lymphohistiocytosis type 4 (FHL-4) (zur Stadt et al., 2005), a hyper-inflammatory disease. Moreover, SNARE also regulates cross presentation of antigens within dendritic cells through the action of Sec22b (Cebrian et al., 2011) which enables the delivery of antigens from the phagosome into the cytosol, thus further corroborating this complex’s role within immune response.

Further, we identify NOTCH signaling as a potential contributor to T1D, which has been implicated in rheumatoid arthritis (Yabe, Matsumoto, Tsurumoto, & Shindo, 2005), another proposed autoimmune disease. There is evidence that Notch receptors are constitutively expressed on macrophages and dendritic cells, thus play a direct role in immunity (Zhang et al., 2012). NOTCH signaling has also been identified within the context of viral response (Ito et al., 2011), with its signal integral in influenza pathogenesis. These previous results highlight the integral role Notch signals play
within many of the proposed mechanisms of T1D initiation, thus our predictions serve as further evidence of this mechanism as a potentiation route to explore further mechanistically and therapeutically. Our KDA results also support this finding, with NOTCH1, 2, 3, and 4 all arising as significant KDs within blood, adipose, macrophage, and monocyte tissue/cell types. Furthermore, in a recent publication by Bartolome et al., chronic NOTCH activation is shown to impair beta cell function through a loss of maturity combined with increased proliferation rate during T2D related physiological stress (Bartolome, Zhu, Sussel, & Pajvani, 2018). Our results suggest that this may also contribute to the loss of beta cells function throughout T1D disease progression.

Finally, we highlight calcium signaling as a perturbed pathway in the pancreatic tissue. Dysregulation of this signaling process has been previously identified as a pathogenic mechanism in acute pancreatitis (Frick, 2012) and its perturbation is generally predicted to play a significant role in apoptotic pathways (Orrenius, Zhivotovsky, & Nicotera, 2003). Moreover, this pathway was identified within the macrophage and blood tissues which may further link calcium signaling and immune dysfunction. Previous reports show that modulation of the calcium signal transducer, calmodulin, mitigates the development of systemic lupus erythematosus in mice and suppresses INFγ production in human SLE cells (Ichinose, Juang, Crispín, Kis-Toth, & Tsokos, 2011). Furthermore, calcium alters B cell reactivity and anergy, thus critically regulating the development of autoreactivity (Bouillet et al., 1999) and inflammatory response (P. Yu et al., 2005). Interestingly, the B cell adaptor protein CD19 which integrates calcium signalling and B cell mediated immunity (Tedder, Haas, & Poe, 2002), was identified as a KD within the blood tissue, which has a protective role in slowing the decline in beta cell function within T1D models (Vonberg et al., 2018).
Overall, our integrative genomics analysis recapitulates previously known pathways, processes and genes associated with T1D pathogenesis, primarily components of the immune system, confirming the robustness of our study. We additionally uncover possible avenues to explore further with suggested pathways and key genes involving the pancreas to be potentially contributing significantly to the T1D outcome through genetic perturbations, including those that can interact with environmental factors such as viral infections. The KDs prioritized through our comprehensive integrative analyses may serve as putative T1D targets for therapeutic development. Future efforts to validate the predicted novel regulators and pathways are warranted.
Table 1. Top pathways associated with T1D identified across multiple tissues at an FDR <5%.


<table>
<thead>
<tr>
<th>Supersets</th>
<th>Module Size</th>
<th>Tissues</th>
<th>Pathways</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1D positive control</td>
<td>60</td>
<td>1, 2, 3, 4, 5, 6</td>
<td>Positive control gene set for T1DM</td>
</tr>
<tr>
<td>S1: Mitotic G1-G1/S phases</td>
<td>186</td>
<td>1, 2, 3, 5, 6</td>
<td>Mitotic G1-G1/S phases; Cell Cycle Checkpoints and Regulation; DNA Replication; Vpu mediated degradation of CD4; Disease; Antigen Processing-Cross presentation</td>
</tr>
<tr>
<td>S2: HIV Infection</td>
<td>222</td>
<td>2, 3</td>
<td>HIV Infection; Host Interactions of HIV factors</td>
</tr>
<tr>
<td>S3: Signaling by NOTCH</td>
<td>122</td>
<td>3, 5, 6</td>
<td>Signaling by NOTCH Pre-NOTCH Expression and Processing</td>
</tr>
<tr>
<td>S4: Protein folding</td>
<td>53</td>
<td>3, 4, 5</td>
<td>Protein folding; Chaperon-mediated protein folding; Association of TriC/CCT with target proteins during biosynthesis</td>
</tr>
<tr>
<td>S5: Signaling by B Cell Receptor (BCR)</td>
<td>170</td>
<td>2, 3</td>
<td>Signaling by the B Cell Receptor (BCR); Downstream Signaling Events of B Cell Receptor (BCR)</td>
</tr>
<tr>
<td>S6: RNA Polymerase I, RNA Polymerase III, and Mitochondrial Transcription</td>
<td>95</td>
<td>3, 5</td>
<td>RNA Polymerase I, RNA Polymerase III, and Mitochondrial Transcription; RNA Polymerase I Transcription; RNA Polymerase I Promoter Clearance; RNA Polymerase I Chain Elongation</td>
</tr>
<tr>
<td>S7: Antigen processing: Ubiquitination &amp; Proteasome degradation</td>
<td>211</td>
<td>2, 3</td>
<td>Antigen processing; Ubiquitination &amp; Proteasome degradation; Class I MHC mediated antigen processing &amp; presentation</td>
</tr>
<tr>
<td>S8: Adaptive Immune System</td>
<td>261</td>
<td>2, 3</td>
<td>Adaptive Immune System; Immune System</td>
</tr>
<tr>
<td>S9: tRNA Aminoacylation</td>
<td>46</td>
<td>2, 3, 5</td>
<td>tRNA Aminoacylation; Aminoacyl-tRNA biosynthesis; Mitochondrial tRNA aminoacylation</td>
</tr>
<tr>
<td>S10: Mitotic M-M/G1 phases</td>
<td>172</td>
<td>2, 3</td>
<td>Mitotic M-M/G1 phases; Mitotic Metaphase and Anaphase; Cell Cycle; Cell Cycle, Mitotic; M Phase; Mitotic Anaphase; Separation of Sister Chromatids</td>
</tr>
<tr>
<td>S11: Signaling by GPCR</td>
<td>230</td>
<td>3</td>
<td>Signaling by GPCR; GPCR downstream signaling GPCR ligand binding</td>
</tr>
<tr>
<td>S12: Activation of GABAB receptors</td>
<td>39</td>
<td>3</td>
<td>Activation of GABAB receptors; Activation of GABA B receptors</td>
</tr>
</tbody>
</table>
Table 2. T1D genetic supersets with KDs for both Bayesian and PPI networks. Only supersets with at least one KD (FDR <5%) were shown. For supersets with more than five KDs, only the top 5 KDs are shown.

<table>
<thead>
<tr>
<th>Supersets</th>
<th>Bayesian Network</th>
<th>PPI Network</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Adipose</td>
<td>Blood</td>
</tr>
<tr>
<td></td>
<td>RPL31, FAU, RPS29, RPS18, RPS27</td>
<td>RPL38, RPS29, RPL7A, RPLP2, RPL7</td>
</tr>
<tr>
<td>Metabolism of mRNA</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RPL31, RPS29, RPS18, RPS27</td>
<td></td>
</tr>
<tr>
<td>Signal Transduction</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene Expression</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Systemic Lupus Erythematosus</td>
<td>HIST1H2BM, HIST1H2BC</td>
<td>PTMA</td>
</tr>
<tr>
<td>Mitotic M-M/G1 phases</td>
<td>RAD51, MCM2, CDC8, MK167, FIGNL1</td>
<td>CDK2, PLK1, PCNA, CDKN1A, MCM7</td>
</tr>
<tr>
<td>Natural Kill Cell Cytoxicity</td>
<td>PTPN6, FERMT3, ARHGAP30, NCKAP1L, LCK</td>
<td>LCK, SHC1, SYK, PTPN6, VAV1</td>
</tr>
<tr>
<td>RNA Polymerase I, RNA Polymerase III, and Mitochondrial Transcription</td>
<td>HIST1H2BM, HIST1H2BC</td>
<td>TBP, ERCC3, CDK7, RB1, BRF1</td>
</tr>
<tr>
<td>Complement &amp; Coagulation</td>
<td>F2, FGG, PLG, AHSG, GC</td>
<td>F2, SERPINA5, PLG, C3, F10</td>
</tr>
<tr>
<td>Disease</td>
<td>RPS18, RPL27, MYO1F, ACTR3, FAU</td>
<td>PIK3R2, FGFR1, ERBB4, NOTCH1, STAT5B</td>
</tr>
<tr>
<td>Immunoregulatory interactions between a Lymphoid and a non-Lymphoid cell</td>
<td>CD3G, LCK</td>
<td></td>
</tr>
</tbody>
</table>
**Figure 1. Overview of Study.**

1) T1D GWAS SNPs are mapped onto genes using tissue specific eQTLs. Genes found are then linked to canonical pathways and co-expression modules. 2) MSEA of Cohort 1 and Cohort 2 are carried out independently for pathway/co-expression module enrichment. 3) Meta-MSEA of the combined cohorts for pathway/co-expression module enrichment. Similar modules are then categorized into independent supersets and input into a wKDA. 4) wKDA implemented using both protein-protein interaction (PPI) and Bayesian networks independently for key driver gene identification, followed by *in silico* validation.
Figure 2. Venn Diagram of enriched canonical pathways and co-expression modules for both T1D GWAS cohorts. A) Venn diagram of the independent and overlapping knowledge driven biological pathways for both cohorts (FDR <5%). B) Venn diagram of the independent and overlapping co-expression modules for both cohorts (FDR <5%).
Figure 3. Heatmap of the tissue-specific meta-MSEA from the combined Cohort 1 and Cohort 2 datasets for the supersets derived from the canonical pathways and co-expression network modules. A) Heatmap for the statistical significance of T1D genetic association across the supersets derived from the Canonical pathways (FDR <5%) in the tissue-specific and cross-tissue analyses. B) Heatmap for the statistical significance of T1D genetic association across the supersets derived co-expression modules (FDR <5%) in the tissue-specific and cross-tissue analyses.
Figure 3. **Tissue specific Key Driver Subnetworks.** A) Pancreas PPI network. B) Blood Bayesian Network. C) Adipose Bayesian Network.


Li, N., Zhang, Z., Zhang, W., & Wei, Q. (2011). Calcineurin B subunit interacts with proteasome subunit alpha type 7 and represses hypoxia-inducible factor-1α activity via the proteasome pathway. *Biochemical and biophysical research communications*, 405(3), 468-472.


