Identification of Differentially Expressed Genes in Pancreatic Regulatory T-Cell Survival

Jake Carrion* and Michael Carrion**

Department of Internal Medicine, Yale University School of Medicine, 300 Cedar Street, New Haven, CT, 06510, USA

*These authors contributed equally to this work

Abstract

Objective: Many current Diabetes treatments cause broad-based immunodeficiency, though future organ-specific therapies may be developed by exploiting genotypic differences found in regulatory T cell (Treg) subpopulations. This study goaled to determine genes preferentially expressed in pancreatic Tregs relative to other Treg subpopulations, tissues, and immune cells to serve as targets in Diabetes therapy.

Methods: Multi-step microarray analysis using GenePattern Software identified genes specific to pancreatic Tregs relative to Tregs in other tissues and to other immune-cell populations. Functional analysis of identified genes was performed using BioGrid and String Databases. Microarray datasets (n=2236) curated from the NCBI GEO database were processed (via RMA normalization, outlier-array exclusion, and global median transformation) to confirm the link of these genes to the pancreas, while further microarray comparisons (in R programming language via Map-5 normalization, log2 transformation, and “trimmed mean” algorithm) were performed to confirm the link of the identified genes to diabetes pathogenesis.

Results: Initial microarray analysis identified genes specific to pancreatic Tregs, with the top three genes (Clps, Pnliprp1, and Pla2g1b) expressed at values 23x, 12x, and 7x higher in pancreatic Tregs than in other Treg subpopulations (p<0.05). Further analysis revealed that the pancreatic gene expression of the identified genes was almost triple that of other tissues (n=29), while comparisons among other immune-cell types indicated that these genes were expressed 32x, 12.8x, and 7.5x higher, respectively, in pancreatic Tregs than in other immune cell-types, eliminating the possibility of augmenting the autoimmune response in future treatments. Diabetic vs. non-diabetic microarray analysis confirmed the link of these genes to Type 1, but not Type 2, Diabetes.

Conclusions: Future Diabetes therapies should target these genes to sole regulate pancreatic Treg levels, avoiding the broad-based immunosuppression caused by current therapies. However, gene knockout studies should be performed to further validate the functionality of the identified genes.

Keywords: Type 1 Diabetes; Differentially expressed genes; Regulatory T cells; Microarray; Autoimmunity

Introduction

Diabetes affects approximately 347 million individuals worldwide [1], and is only predicted to become more prevalent in the upcoming years, affecting potentially 1 in 3 individuals by 2050 [2]. Multiple micro- and macrovascular complications result from diabetes, including retinopathy, nephropathy, and cardiovascular disease [3], increasing the already staggering mortality rates; Diabetes is predicted to be the 7th leading cause of death in the world by 2030 [4]. Additionally, in 2012, Diabetes cost the United States $245 billion, marking a 41% increase in funding in 5 years [5].

Several immunosuppressive treatments, notably calcineurin inhibitors such as cyclosporine and tacrolimus, have been developed to help suppress effector T cells (Teffs), which trigger the autoimmune response in Type 1 Diabetes [6]. Yet immunosuppressive treatments are not specific to the autoimmune Teffs responsible for autoimmunity, and often result in complete immunosuppression and increased susceptibility to common diseases (Figure 1) [7,8].

In recent years, regulatory T cells (Tregs) have gained much popularity as a potential treatment option for immune disease due to their role in suppressing the autoimmune Teffs [9]. Tregs function through a variety of means, including via secretion of inhibitory cytokines (i.e. IL-10 and TGF-β), deprivation-mediated apoptosis of effector T cells through IL-2 cytokine depletion, and modulation of antigen-presenting cells through CTLA-4 protein interactions, to regulate autoreactive Teffs and thus prevent pancreatic beta cell apoptosis [10]. Yet contrary to widespread belief, a groundbreaking paper by Feurer and colleagues in 2010 revealed that Tregs are not a homogeneous population of T cells but rather have slightly different phenotypes based on differing anatomical locations [11]. The expression of distinct genes results in unique functions of each Treg subpopulation. For example, overexpression of Areg in muscle Tregs results in a unique ability to secrete the anti-inflammatory cytokine amphiregulin in response to acute injury [12]. By exploiting this relatively new discovery, it may be possible to stimulate only one subpopulation of Tregs by targeting a specific gene or protein that is unique to that type of Treg, while leaving the remaining Treg subpopulations undisturbed and thus preventing broad-based immunosuppression.

This study goaled to: (1) identify a means of selectively targeting pancreatic Tregs, without influencing Tregs in other anatomical
locations, (2) identify genes preferentially expressed within pancreatic Tregs, relative to other Treg subpopulations, other immune-cell types, and other tissues in the body, and (3) determine the importance of such preferentially expressed genes in modulating Treg suppressive ability.

Methodology

Initial gene screening

Normalized microarray data containing gene expression values for eight different populations of Tregs (pancreas, colon, adipose tissue, muscle, and corresponding splenic Tregs) were obtained from the Immunological Genome Project Consortium [13]. To eliminate genes (1) not capable of impacting cellular function, and (2) expressed on all Treg types, the dataset was filtered for genes possessing a significant expression value (>100) and a substantial fold change (>2) when compared to a tissue’s splenic counterpart [14] using Microsoft Excel (Version 14.0). The splenic counterparts provide a baseline Treg gene signature, as these splenic Tregs have not yet developed specific gene signatures resembling those of localized Treg subpopulations [12].

Selection refinement

Volcano plots displaying fold-change differences and p-values for the filtered genes were generated using Multiplot Studio (version 1.5.2) on GenePattern software (Version 3.8; Broad Institute). Data outputs were filtered according to established criteria (coefficient of variation <0.07, p<0.05, fold change >2) [14]. A heatmap was created with the remaining genes through Hierarchical Clustering (Version 6.0) on GenePattern. Heatmaps were used to identify genes with significant preferential expression in pancreatic Tregs compared to other Treg subpopulations (fold change >2 in pancreatic Tregs and fold change ≤0 in all other tissues) [15].

Verification of gene function

Functional analysis of the remaining genes (n=11) was performed using BioGrid (version 3.2) [16] and GeneMANIA (version 3.1) [14] databases and through previous literature review. Gene function was investigated for elimination under the following criteria: (1) genes not directly associated with Treg survival or suppressive ability, (2) genes not expressed in humans and only expressed in Mus musculus, and (3) genes not encoding for an extracellular protein and thus not targetable in a non-invasive manner.

Expression verification among all tissues

The NCBI GEO microarray database was manually curated for relevant microarray datasets, according to previously established criteria [15]: (1) Microarray data must be available from Homo sapiens, (2) Data must be obtained from normal (non-diseased) tissues, (3) Data must be obtained from tissues that are untreated and not-cultured in-vitro, and (4) Microarrays must be run on the HG-U133A microarray platform. The raw data sets were then categorized by tissue, and a Robust-Multi Average (RMA) normalization algorithm was applied to the data via RMAExpress Software (Version 1.1.0) to reduce technical variations across different microarray samples. The data was filtered for outlier arrays by calculating a resistant z-score for each probe in a given array, and then excluding those arrays where over 15% of probes had a resistant z-score greater than 3. A global-median transformation was then applied to the data, allowing comparisons between different microarray samples and tissues (Figure 2) [15,16].

Expression verification among other immune cell types

Microarray datasets for a variety of immune cells (B-cells, T-helper cells, Naive T-cells, Natural Killer Cells, Dendritic Cells, Monocytes, Macrophages, and Neutrophils) were obtained from the Immunological Genome Project Consortium [13]. Expression values for each of the remaining genes (n=3) were compared between each cell-type and pancreatic Tregs in order to determine preferential expression among pancreatic Tregs relative to other immune infiltrates found in the pancreas of a diabetic individual; Statistical significance was calculated via Student’s t-test, and a p-value <0.05 was considered statistically significant.

Diabetic vs. non-diabetic islet comparisons

Raw microarray data was downloaded from the NCBI GEO database for the islets of Type 1 Diabetic individuals, Type 2 Diabetic individuals, and normal-control patients. The data was then input into the Bioconductor Package (Version 3.0) in R programming software [17], and the data was normalized by applying a MAS-5 normalization algorithm and log2 transformation to the expression values [18]. A trimmed mean algorithm was then applied to the data to allow comparisons between different microarray samples.
Results

Initial gene screening and selection refinement

An original pool of 24,000 genes was surveyed. Initial screening identified 142 genes of interest (Figure 3A). The majority of genes were expressed in multiple Treg subpopulations (Figure 3B), with only 11 being significantly expressed among solely the pancreatic Tregs (Table 1). The 11 uniquely expressed genes were used in further functional analysis.

Preferential expression verification

Gene expression profiles for the 11 identified genes confirmed significant overexpression in the pancreatic Tregs relative to all other subpopulations (Figure 4). Clps, Try10, and Pnliprp1 possessed the highest relative expression values, with an average expression of 2,010 AU in pancreatic Tregs vs. an average of only 79 AU in the remaining Treg subpopulations (p<0.05).

An intensive literature search was then performed on these 11 remaining genes to determine which genes would serve as the best candidates for future drug targeting.

Identification of functionally significant genes

Functional analysis of the top 11 genes reveal that the genes Clps, Pnliprp1, and Pla2g1b likely influence the suppressive ability of Tregs with regards to Diabetes (Table 2).

Comparisons against other tissues

The relevant datasets curated from the NCBI GEO database (n=2236) were classified by tissue (n=29), and comparisons between pancreatic levels of the identified genes and other tissues confirmed the uniqueness of these genes to pancreatic Tregs (Figure 5).

The average expression levels of Clps were 217.10 AU in pancreatic Tregs vs. only 102.97 AU in the other tissues of the body (p<0.05). Similarly, the expression levels of Pnliprp1 were 202.66 AU in pancreatic Tregs vs. an average of 93.54 AU in other body tissues and Pla2g1b expression levels were 210.49 AU in the pancreas vs. 78.13 AU in other tissues (p<0.05).

Expression verification among other immune cell types

Clps, Pnliprp1, and Pla2g1b were then compared to other cell types found in the pancreas of diabetic individuals (B-cells, T-helper

---

Table 1: List of differentially expressed genes in pancreatic Tregs: Genes (n=11) were selected from the heat map if they possessed a fold-change > 2 in pancreatic Tregs and ≤ 0 in other Treg populations.

| Probe Set ID | Symbol | Gene Name                                      |
|--------------|--------|------------------------------------------------|
| 10412218     | Gzmk   | Granzyme K                                     |
| 10432652     | Cela1  | Chymotrypsin-like elastase family member 1     |
| 10449467     | Clps   | Colipase, Pancreatic                           |
| 10464313     | Pnliprp1| Pancreatic Lipase-related Protein 1             |
| 10497831     | Ccna2  | Cyclin A2                                      |
| 10518050     | Cela2a | Chymotrypsin-like elastase family member 2A    |
| 10518059     | Ctrc   | Chymotrypsin C (Caldecrin)                     |
| 10524698     | Pla2g1b| Phospholipase A2, Group 1B                     |
| 10537638     | Try10  | Trypsin 10                                     |
| 10537650     | Prss1  | Protease, Serine, 1 (Trypsin 1)                |
| 10544333     | Try5   | Trypsin 5                                      |

---

Figure 3: Initial identification of differentially expressed genes in pancreatic Tregs. Selected genes are unique to pancreatic Tregs relative to other Treg subpopulations (A) Treg genes (Boxed and highlighted red; n=142) are unique to pancreatic Tregs relative to splenic counterpart Tregs (p<0.05). (B) Try10 and Clps possess significant overexpression in the pancreas relative to all other Treg subpopulations (p<0.05).
cells, Naive T-cells, Natural Killer Cells, Dendritic Cells, Monocytes, Macrophages, and Neutrophils). Comparisons of gene expression values revealed that the three genes were expressed in very small amounts among other immune cells.

Whereas Clps possessed an average expression value of 4121.4 AU on the surface of pancreatic Tregs, its average expression value among the other cell types was only 125.6 AU, indicating a 32x greater expression in the pancreatic Tregs (Figure 6A). Similarly, Pnliprp1 had an average expression value of 313.1 AU on the surface of pancreatic Tregs as compared to an average expression value of 24.4 AU on the surface of the other cell types, marking a 12.8x difference (Figure 6B). Pla2g1b also possessed an average expression value of 658.8 AU in
pancreatic Tregs compared to an average expression value of only 87.6 AU on the other immune cells, a 7.5x greater fold change (Figure 6C).

**Diabetic vs. non-diabetic islet comparisons**

Microarray comparisons between Type 1 Diabetics and normal controls revealed that there is a significant difference in gene expression levels between the two groups, with all three selected genes significantly overexpressed in the diabetic islets (p<0.05; Figure 7A). However, microarray analysis revealed that there is no significant difference in expression levels between Type 2 diabetic islets and control islets (Figure 7B).

**Discussion**

Primary findings: The data identify three genes (Clps, Pnliprp1, and Pla2g1b) that are specific to pancreatic Tregs. The genes are not found on other Treg subpopulations nor in other tissues in the body, allowing for future therapies to reverse pancreatic autoimmunity while avoiding broad-based immunosuppression. Comparisons against other immune-cells confirm that the targeting of such genes will not impact any immune infiltrates that might be residing in the pancreas of a diabetic individual, and will thus not risk augmenting the autoimmune response. Further comparisons against diabetic islets reaffirm that the genes are implicated in Type 1 Diabetes pathogenesis, yet the insignificant differences between Type 2 Diabetic and normal islet controls suggest that the genes are not implicated in Type 2 Diabetes development. However, this link between the genes and Type 2 Diabetes is expected, as Type 2 Diabetes is caused by obesity and insulin resistance, not autoimmunity, and thus the immune composition of both type 2 diabetic islets and normal islet controls should be relatively similar.

**Potential mechanism of gene-treg interactions**

Pnliprp1, Clps, and Pla2g1b impact diabetes development and insulin-resistance in individuals, likely influencing Treg function with regards to diabetes.

Pnliprp1 knockout mice experienced increased weight gain, higher fat mass, and severe insulin resistance compared to wild-type controls [19]. This most likely results from altered triglyceride hydrolysis in the absence of Pnliprp1. Similarly, arginine-to-cysteine mutations in Clps have been associated with increased Type 2 Diabetes susceptibility [19]. Such mutations also hinder triglyceride hydrolysis, resulting in inefficient fat digestion and a decrease in free floating fatty acids [20-23]. Conversely, Pla2g1b may foster diabetes development in individuals [19], as previous results indicate that Pla2g1b-/- mice are protected from insulin resistance, diet-induced obesity, and hyperlipidemia [20,21]. This is a result of inadequate phospholipid hydrolysis, which hinders lipid digestion and fat absorption, but also results in reduced numbers of fatty acids [24,25].

The decreased numbers of free floating fatty acids resulting from alterations in Pla2g1b, Pnliprp1, and Clps may alter the plasma membrane of Tregs, either through affecting receptor molecules on the surface of these cells or disrupting the intracellular signaling processes that drive Treg cell survival or suppressive processes (Figure 8) [26,27].

This conclusion linking Pnliprp1, Clps, and Pla2g1b to the
functioning of regulatory T cells has not previously been established, though it is in accordance with (1) the genes’ extremely high expression on Tregs relative to other cell types and (2) their role in lipid hydrolysis and free fatty acid modulation, which has previously been shown to impact Treg function.

Thus Pla2g1b, Pnliprp1, and Cpls are attractive candidate genes for future therapies due to their functional importance in Treg suppressive ability and preferential expression among pancreatic Tregs.

Applications

The data confirm the classification of phenotypically different Treg subpopulations based on anatomical location, and can improve the current understanding of phenotypic differences among distinct subpopulations of regulatory T cells. The targeting of identified genes may allow future immunosuppression therapies to be made specific to a particular Treg population and thus a specific region in the body, consequently avoiding the harms associated with the broad-based immunosuppression used today [8]. Malignancies associated with transplantation, including allograft rejection or graft-versus-host disease, may potentially be treated by activating specific subpopulations of Tregs found in the location of transplantation. The methodology used in the current study can be utilized in other organ-specific autoimmune diseases as well, including Grave’s Disease or Autoimmune hepatitis, where the identification of phenotypic differences among Tregs located in the thyroid and liver, respectively, can lead to more effective treatments [11].

Future investigations

Additional studies should be conducted to confirm the validity of the results and set the stage for the development of therapeutic treatments. For example, flow cytometric analysis should be run using the pancreata of 25-30-week-old wild type mice to confirm the expression of the desired proteins on a cell's surface [28,29], as RNA quantities obtained from microarray data do not always correlate with protein expression and occasionally yield false positives. Cell-type-specific gene knock-out experiments can also be performed, either through in-vivo mouse models or in-vitro methods including standard siRNA procedures [30,31]. The use of these procedures can directly confirm the significance of the chosen genes in pancreatic Treg function and survival [32,33]. Collectively, these experiments can validate the significance of the identified genes in Treg suppressive function and serve as the primary steps in the development of future diabetes therapies.

Conflict of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgements

The authors thank Dr. Serena McCalla for critical reading of the manuscript.

References

1. Danai E, Finucane M, Lu Y, Singh G, Cowan M, et al. (2011) National, regional, and global trends in fasting plasma glucose and diabetes prevalence since 1980: systematic analysis of health examination surveys and epidemiological studies with 370 country-years and 2.7 million participants. The Lancet 378: 31-40.
2. Boyle J, Thompson T, Gregg E, Barker L, Williamson D (2010) Projection of the year 2050 burden of diabetes in the US adult population: dynamic modeling of incidence, mortality, and prediabetes prevalence. Population Health Metrics 8: 29.
3. Fowler M (2008) Microvascular and Macrovascular Complications of Diabetes. Clinical Diabetes 26: 77-82.
4. Rathmann W, Giani G (2004) Global Prevalence of Diabetes: Estimates for the Year 2000 and Projections for 2030. Diabetes Care 27: 2568-2569.
5. Narayan K, Boyle J, Geiss L, Saaddeen J, Thompson T (2006) Impact of Recent Increase in Incidence on Future Diabetes Burden: U.S., 2005-2050. Diabetes Care 29: 2114-2116.
6. Schade D, Valentine V (2002) To Pump or Not to Pump. Diabetes Care 25: 2100-2102.
7. Corthay A (2009) How do Regulatory T Cells Work? Scandinavian Journal of Immunology 70: 328-336.
8. Taylor A, Watson C, Bradley J. (2005) Immunosuppressive agents in solid organ transplantation: Mechanisms of action and therapeutic efficacy. Critical Reviews in Oncology/Hematology 56: 23-46.
9. Sakaguchi S, Yamaguchi T, Nomura T, Ono M (2008) Regulatory T Cells and Immune Tolerance. Cell 133: 775-787.
10. Schmidt A, Oberle N, Krammer P (2012) Molecular Mechanisms of Treg-Mediated T Cell Suppression. Frontiers in Immunology 3: 1-20.
11. Feurer M, Hill J, Kretschmer K, von Boehmer H, Mathis D et al. (2010) Genomic definition of multiple ex vivo regulatory T cell subphenotypes. Proceedings of the National Academy of Sciences 107: 5919-5924.
12. Burzyn D, Kuswanto W, Kolodin D, Shadrach J, Cerletti M, et al. (2013) A Special Population of Regulatory T Cells Potentiates Muscle Repair. Cell 155: 1282-1295.
13. Shay T, Kang J (2013) Immunological Genome Project and systems immunology. Cell 34: 602-609.
14. Warde-Farley D, Donaldson S, Cones O, Zuberi K, Badrawi R, et al. (2010) The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. Nucleic Acids Research 38: W214-W220.
15. Wang L, Srivastava A, Schwartz C (2010) Microarray data integration for genome-wide analysis of human tissue-selective gene expression. BMC Genomics 11: S15.
16. Stark C (2006) BioGRID: a general repository for interaction datasets. Nucleic Acids Research 34: D535-D539.
17. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M et al. (2004) Bioconductor: open software development for computational biology and bioinformatics. Genome Biology 5: R80.
18. Piccolo SR, Sun Y, Campbell JD, Lenburg ME, Bild AH, et al. (2012) A single-sample microarray normalization method to facilitate personalized-medicine workflows. Genomics 100: 337-344.
19. Hollie N, Konaniah E, Goodin C, Hui D (2014) Group 1B phospholipase A2 inactivation suppresses atherosclerosis and metabolic diseases in LDL receptor-deficient mice. Atherosclerosis 234: 377-380.

20. Giller T, Buchwald P, Blum-Kaelin D, Hunziker W (1992) Two novel human pancreatic lipase related proteins, hPLRP1 and hPLRP2. Differences in colipase dependence and in lipase activity. The Journal of Biological Chemistry 267: 16509-16516.

21. Labonte E, Kirby R, Schildmeyer N, Cannon A, Huggins K, et al. (2006) Group 1B Phospholipase A2-Mediated Lysophospholipid Absorption Directly Contributes to Postprandial Hyperglycemia. Diabetes 55: 377-380.

22. de Jong A, Kloppenburg M, Toes R, Ioan-Facsinay A (2008) Fatty Acids, Lipid Mediators, and T-Cell Function. Frontiers in Immunology 9: 1-15.

23. Gust T, Neubrandt L, Merz C, Aadullah K, Zagel U, et al. (2008) RNA interference-mediated gene silencing in murine T cells: in vitro and in vivo validation of proinflammatory target genes. Cell Commun. Signal 6: 3.

24. Vignal D, Collison L, Workman C (2008) How regulatory T cells work. Nature Reviews Immunology 8: 523-532.

25. Craig F, Foon K (2008) Flow cytometric immunophenotyping for hematologic neoplasms. Blood 111: 3941-3967.

26. Nicolaou A, Mauro C, Urquhart P, Marelli-Berg F (2014) Polyunsaturated Fatty Acid-Derived Lipid Mediators and T Cell Function. Frontiers in Immunology 5: 1-15.

27. Vignal D, Collison L, Workman C (2008) How regulatory T cells work. Nature Reviews Immunology 8: 523-532.

28. Craig F, Foon K (2008) Flow cytometric immunophenotyping for hematologic neoplasms. Blood 111: 3941-3967.

29. Hanley MB, Lomas W, Mittar D, Maino V, Park E (2013) Detection of Low Abundance RNA Molecules in Individual Cells by Flow Cytometry. PLOS One 8: 1-8.

30. von Herrath M, Peakman M, Roep B (2013) Progress in immune-based therapies for type 1 diabetes. Clin. Exp. Immunol 172: 186-202.