ANDRÉ RAMOS VAQUERO

**Estudo comparativo de redes gênicas de expressão
de genes associados à diabetes mellitus tipo 2 (DM2)
e genótipos de risco da doença**

Dissertação apresentada à Faculdade de Medicina
da Universidade de São Paulo para obtenção do
título de Mestre em Ciências

Programa de Ciências Médicas

Área de concentração: Distúrbios Genéticos de
Desenvolvimento e Metabolismo

Orientador: Dr. Alexandre da Costa Pereira

São Paulo

2013

Dados Internacionais de Catalogação na Publicação (CIP)

Preparada pela Biblioteca da
Faculdade de Medicina da Universidade de São Paulo

©reprodução autorizada pelo autor

Vaquero, André Ramos

Estudo comparativo de redes gênicas de expressão de genes associados à diabetes mellitus tipo 2 (DM2) e genótipos de risco da doença / André Ramos Vaquero. -- São Paulo, 2013.

Dissertação(mestrado)--Faculdade de Medicina da Universidade de São Paulo. Programa de Ciências Médicas. Área de concentração: Distúrbios Genéticos de Desenvolvimento e Metabolismo.

Orientador: Alexandre da Costa Pereira.

Descritores: 1.Proteína fator de transcrição 7 semelhante ao 2/genética 2.Diabetes mellitus tipo 2 3.Locos de características quantitativas 4.Células do músculo liso 5.Expressão gênica 6.Redes reguladoras de genes

USP/FM/DBD-050/13

Dedicatória

Aos meus pais, Efigênia e Rui, pelo exemplo e suporte nos momentos em que precisei, pelo carinho e compreensão, por todo apoio obtido, o qual propiciou eu escolher o meu caminho e não ser escolhido por ele. Sem eles realmente meus estudos e minha formação (pessoal e profissional) não seriam possíveis. Vocês são minha base.

Agradecimentos

Aos meus professores

Mesmo aqueles que não obtêm o título, mas seriam merecedores. Alexandre da Costa Pereira (Ale), Célia Koiffman e Maria Cristina Arias, os quais tenho o prazer de chamar de professores (para mim o maior elogio que posso fazer). Obrigado pelo direcionamento, oportunidade, compreensão e confiança depositada.

Aos meus colegas/amigos de laboratório

Sei que é padrão agradecer por agradecer aos colegas de trabalho, mas nesse caso diria que não faço isso por obrigação. Realmente tenho um ambiente dos melhores para trabalhar e agradeço a todos que contribuíram para esse trabalho, seja desenvolvendo o projeto comigo ou simplesmente por participar do meu dia-a-dia. Paulo, Júlia, Théo, Menino Baependi, Amanda, Carolzinha, Carla, Cíntia, Carol, Noely, Lívia, Léa, Lívia S., o pessoal mais novo do grupo, o pessoal da extração, o pessoal do HF, o pessoal da cultura de células, o pessoal da BioInfo, o pessoal do laboratório geral, o pessoal da secretaria....no dia-a-dia vocês sabem o quanto gosto de vocês.

À minha Família

Agradeço a eles por apenas existirem e gostarem de mim, acima de tudo e independentemente de tudo. Pais, Kátia, Luiz, Luízinha e Sassa. É recíproco com certeza.

À minha segunda família

Família da Sassa, que sempre me acolheu muito bem e por quem tenho muito carinho também. Pelos momentos divertidos que foram e estão por vir. Em especial Rosa, Giovana, Leleco e Edson.

Aos meus amigos Diogo e Judas, saibam que vocês são muito importantes para mim, pela amizade, pelos momentos partilhados e conversas sempre produtivas, seja qual for o assunto.

Aos meus amigos queridos de faculdade, os quais são parte da minha família hoje com certeza. Sem eles eu não seria o que sou hoje. Vocês me tornam uma pessoa melhor.

Aos meus amigos do Biosal, com os quais me diverti muito, passei por muitos momentos bons e ruins juntos e aprendi a trabalhar em grupo. São muito importantes na minha formação pessoal e pude aprender muito com vocês.

À Capes e ao Departamento de Clínica Médica, pelo apoio financeiro que permitiu a realização do mestrado.

*“It's Not Who I Am Underneath,
But What I Do That Defines Me”*

*“Não é quem eu sou por dentro
mas o que eu faço que me define”*

*“But it ain't about how hard you
hit. It's about how hard you can
get hit and keep moving forward”*

*“Mas isso não é sobre o quão
forte você bate. Isso é sobre o
quanto você agüenta apanhar e
continuar seguindo em frente”*

Esta dissertação ou tese está de acordo com as seguintes normas, em vigor no momento desta publicação:

Referências: adaptado de *International Committee of Medical Journals Editors (Vancouver)*

Universidade de São Paulo. Faculdade de Medicina. Divisão de Biblioteca e Documentação. Guia de apresentação de dissertações, teses e monografias. Elaborado por Anneliese Carneiro da Cunha, Maria Julia de A. L. Freddi, Maria F. Crestana, Marinalva de Souza Aragão, Suely Campos Cardoso, Valéria Vilhena. 3a ed. São Paulo: Divisão de Biblioteca e Documentação; 2011.

Abreviaturas dos títulos dos periódicos de acordo com *List of Journals Indexed in Index Medicus*.

SUMÁRIO

Lista de siglas

Resumo

Abstract

1. INTRODUÇÃO	1
1.1 Doenças cardiovasculares e diabetes mellitus tipo 2	1
1.2 Relação de variantes genéticas com DM2	3
1.3 Estudo das variantes genéticas para entendimento de fenótipos celulares	4
1.4 Bases genéticas da variação na expressão gênica	5
1.5 Tipos de células usadas em análises de expressão gênica relacionados ao DM2	6
1.6 Análise de expressão gênica	7
1.7 Outras aplicações de dados de expressão gênica	8
1.8 Tipos de redes de expressão gênica	8
1.9 Justificativa	9
2. OBJETIVOS	11
2.1 Objetivo geral	11
2.2 Objetivos específicos	11
3. MÉTODOS	13
3.1 Amostras do estudo e Termo de Consentimento Livre e Esclarecido	13
3.2 Genotipagem de 92 indivíduos e separação em dois grupos genotípicos	13
3.3 Análise da expressão gênica	17
4. RESULTADOS	26
4.1 Análises de efeitos ambientais	26
4.2 Comparações de medianas de expressão e de redes de interações protéicas entre os grupos genotípicos	30
4.3 Comparações de padrões de co-expressão gênica entre os grupos genotípicos	34
5. DISCUSSÃO	44
6. CONCLUSÃO	52
7. REFERÊNCIAS	53

LISTA DE SIGLAS

ANOVA	Análise de variância
AVC	Acidente vascular cerebral
C	Citosina
CMLs	Células de músculo liso
CRCCP	Ciclos de replicação celular em cultura primária
DAC	Doença arterial coronariana
DM2	Diabetes mellitus tipo 2
DNA	Ácido desoxiribonucleico
eQTLs	Estudo quantitativo de expressão gênica associado a um <i>locus</i>
FM-USP	Faculdade de Medicina da Universidade de São Paulo
GWA	<i>Genome-wide association</i>
HRM	<i>High Resolution Melting</i>
HWE	Equilíbrio de Hardy-Weinberg
I2D	Banco de dados de interações protéicas
IMC	Índice de massa corpórea
InCor	Instituto do Coração
PCR	Reação em cadeia da polimerase
RNA	Ácido ribonucléico
RT-PCR	Transcrição reversa em tempo real da PCR
SNP	Polimorfismo de base única
T	Timina
Wnt	<i>Wingless</i>

LISTA DE TABELAS

Tabela 1 - Características clínicas das 92 amostras do estudo	14
Tabela 2 - Formas de <i>splicing</i> do <i>TCF7L2</i> e ensaios realizados	19
Tabela 3 - Lista de todos os genes quantificados neste estudo	20
Tabela 4 -Lista de <i>primers</i> utilizados nos ensaios de RT-PCR em tempo real	21
Tabela 5 - Efeitos ambientais nas medidas de expressão gênica	27
Tabela 6 - Diferenças de medianas de expressão gênica entre os grupos de diabéticos e não diabéticos de 40 genes	29
Tabela 7 - Diferenças de medianas de expressão gênica entre os grupos de diabéticos e não diabéticos das formas de <i>splicing</i> do <i>TCF7L2</i>	30
Tabela 8 - Diferenças de medianas de expressão gênica entre os grupos selvagem e de risco das formas de <i>splicing</i> do <i>TCF7L2</i>	31
Tabela 9 - Diferenças de medianas de expressão gênica entre os grupos selvagem e de risco de 40 genes	32
Tabela 10 - Diferenças significativas de coeficientes de correlação das formas do gene <i>TCF7L2</i> entre os grupos selvagem e de risco	34
Tabela 11 - Comparações entre os grupos genotípicos de coeficientes de correlação gênica das formas do <i>TCF7L2</i> com os demais genes	35
Tabela 12 - Comparações entre os grupos genotípicos de coeficientes de correlação gênica dos genes com maiores números de diferenças de coeficientes de correlação	38
Tabela 13 - Demais comparações entre os grupos genotípicos de coeficientes de correlação gênica significativamente diferentes	39
Tabela 14 - Comparações entre os grupos genotípicos de coeficientes de correlação gênica com apenas os genes diferentemente expressos	41

LISTA DE FIGURAS

Figura 1 - Resumo dos objetivos específicos 12

Figura 2 - Ensaaios com formas de *splicing* do *TCF7L2* realizados 19

Figura 3 - Diferenças de expressão e co-expressão das formas de *splicing* do gene *TCF7L2* entre os grupos genotípicos 31

Figura 4 - Redes de interação protéica com medianas de expressão gênica dos grupos selvagem e de risco 33

Figura 5 - Redes de coexpressão com os genes *RXR α* , *CALM1* e *CALR* nas quais as faixas de valores de coeficientes de correlação de Spearman são $\leq -0,4$ e $\geq 0,4$ 40

Figura 6 - Redes de coexpressão com os genes diferentemente expressos nas quais as faixas de valores de coeficientes de correlação de Spearman são $\leq -0,4$ e $\geq 0,4$ 43

RESUMO

Vaquero, A.R. Estudo comparativo de redes gênicas de expressão de genes associados à diabetes mellitus tipo 2 (DM2) e genótipos de risco da doença [Dissertação]. São Paulo: Faculdade de Medicina, Universidade de São Paulo; 2013.

INTRODUÇÃO: O polimorfismo dentro do gene *TCF7L2*, rs7903146, é, até o momento, o marcador genético mais significativamente associado ao risco de diabetes mellitus tipo 2, sendo também associado à doença arterial coronariana. Contudo, pouco ainda se conhece sobre o papel funcional desse polimorfismo na patologia dessas doenças. O objetivo desse projeto foi investigar esse papel funcional, no fenótipo de células vasculares de músculo liso de 92 indivíduos, usando abordagens de comparação de níveis de expressão gênica e de comparação de correlações de expressão gênica, de modo que tais comparações fossem representadas visualmente como redes de interação gênica. **MÉTODOS:** Inicialmente, foram comparados os níveis de expressão de 41 genes (genes que possuem ou estão perto de variantes genéticas associadas ao diabetes mellitus tipo 2 e outros genes relacionados às vias de sinalização de diabetes mellitus tipo 2 ou às vias de proliferação celular) entre indivíduos com o alelo associado ao risco de diabetes mellitus tipo 2 (CT e TT) e indivíduos sem o alelo de risco (CC) do rs7903146. Com a finalidade de se observar se os genes estavam se relacionando de modo diferente entre os grupos genotípicos, foram comparados os padrões de correlação de expressão dos 41 genes. **RESULTADOS:** Quanto às comparações de níveis de expressão entre os grupos, cinco formas de *splicing* do gene *TCF7L2* e os genes *CDKAL1*, *IGF2BP2*, *JAZF1*, *CDKN2B*, *CAMK1D*, *JUN*, *CDK4*, *ATP2A2*, e *FKBP1A* apresentaram níveis de expressão significativamente diferentes. Quanto às comparações de correlação de expressão entre os grupos, os genes *RXRα*, *CALM1*, *CALR* e *IGF2BP2* foram os que mostraram os mais diferentes padrões de correlação com os outros genes. **CONCLUSÃO:** Deste modo, o alelo de risco analisado é apontado como tendo influência em *cis* na regulação da expressão de determinadas formas de *splicing* do gene *TCF7L2* em células vasculares de músculo liso; além de parecer influenciar nas expressões e nas interações de genes relacionados à homeostase glicolítica e/ou proliferação celular. Sendo assim, através de nossas análises identificaram-se possíveis candidatos-alvos no tratamento de redução do risco em indivíduos com alto risco de desenvolvimento de diabetes mellitus tipo 2 e de doença arterial coronariana, especialmente os indivíduos que possuem os genótipos de risco analisados do gene *TCF7L2*.

Descritores: 1. Proteína fator de transcrição 7 semelhante ao 2/genética 2. Diabetes mellitus tipo 2 3. Locos de características quantitativas 4. Células do músculo liso 5. Expressão gênica 6. Redes reguladoras de genes

ABSTRACT

Vaquero, A.R. Comparative study of gene networks of genes associated with type 2 diabetes mellitus (DM2) and the risk genotypes for the disease [dissertation]. São Paulo: "Faculdade de Medicina, Universidade de São Paulo"; 2013.

INTRODUCTION: The SNP within the *TCF7L2* gene, rs7903146, is, to date, the most significant genetic marker associated with type 2 diabetes mellitus risk, well as being associated with coronary artery disease. Nonetheless, its functional role in these diseases pathology is poorly understood. The aim of the present study was to investigate this role, in vascular smooth muscle cells from 92 patients undergoing aortocoronary bypass surgery, using expression levels and expression correlation comparison approaches, which were visually represented as gene interaction networks. **METHODS:** Initially, the expression levels of 41 genes (seven *TCF7L2* splice forms and other 40 relevant genes) were compared between rs7903146 wild-type (CC) and type 2 diabetes mellitus risk (CT + TT) genotype groups. Next, the expression correlation patterns of the 41 genes were compared between genotypic groups in order to observe if the relationships between genes were different. **RESULTS:** Five *TCF7L2* splice forms and *CDKAL1*, *IGF2BP2*, *JAZF1*, *CDKN2B*, *CAMK1D*, *JUN*, *CDK4*, *ATP2A2* and *FKBP1A* genes showed significant expression differences between groups. *RXR α* , *CALM1*, *CALR* and *IGF2BP2* genes were pinpointed as showing the most different expression correlation pattern with other genes. **CONCLUSION:** Therefore, type 2 diabetes mellitus risk alleles appear to be influencing *TCF7L2* splice form's expression in vascular smooth muscle cells; besides it can be influencing expression and interactions of genes related to glucose homeostasis and/or cellular proliferation. Thereby, through our analysis were identified possible treatment target candidates for risk reduction in individuals with high-risk of developing type 2 diabetes mellitus and coronary artery disease, especially individuals harboring *TCF7L2* risk genotypes.

Descriptors: 1. *TCF7L2* 2. Type 2 diabetes mellitus 3. eQTL 4. Vascular smooth muscle cells 5. Gene expression 6. Regulatory gene networks

1. INTRODUÇÃO

1.1 Doenças cardiovasculares e diabetes mellitus tipo 2

As doenças cardiovasculares afetam o coração e os vasos sanguíneos, sendo a principal causa de mortalidade no mundo¹. Tais doenças compreendem diagnósticos como doença arterial coronariana, doença arterial periférica, infarto do miocárdio, acidente vascular cerebral (AVC), hipertensão arterial, cardiopatia congênita e insuficiência cardíaca; sendo que algumas delas possuem como evento primário um espessamento da parede arterial, podendo causar uma patologia conhecida como aterosclerose, a qual leva a uma redução do fluxo sanguíneo e à isquemia².

O diabetes é apontado como um dos principais fatores de risco para doenças cardiovasculares^{3,4}, sendo associado também à aceleração do processo de aterosclerose⁵, e conseqüentemente, podendo ocasionar aos indivíduos complicações crônicas (tardias) macrovasculares (como macroangiopatias e hipertensão arterial) e microvasculares (como retinopatia, nefropatia e neuropatia)⁶. Segundo dados da Organização Mundial de Saúde¹, 346 milhões de pessoas no mundo têm diabetes. O risco total de morte de pessoas com diabetes é pelo menos o dobro do risco de pessoas que não possuem a doença – e isso se deve, em parte, por essa doença aumentar o risco de problemas cardíacos e AVC, de modo que 50% das pessoas com diabetes morrem de doenças cardiovasculares¹. Mais de 80% dessas mortes por diabetes ocorrem em países de renda baixa e média, e estima-se que o total de mortes por diabetes tende a duplicar no período entre 2005 e 2030¹.

Os tipos principais de diabetes são o diabetes mellitus tipo 1 e o diabetes mellitus tipo 2 (DM2), sendo o segundo de maior freqüência na população geral, afetando mais de 300 milhões de pessoas pelo mundo¹. O DM2 apresenta como principais características: altos níveis de glicose no sangue; resistência à sinalização da insulina e, na fase tardia, diminuição da secreção de insulina pelas células β (beta) do pâncreas^{7,8}.

O DM2 é causado por fatores de risco ambientais e genéticos (doença multifatorial complexa). Um ponto importante para a prevenção da doença e para a diminuição das complicações crônicas que esta acarreta é a identificação rápida de pessoas com alto risco de desenvolvê-la. Com esta finalidade, alguns dos fatores típicos de risco de DM2 analisados comumente são: aumento no índice de massa corpórea (IMC, o peso em quilogramas dividido pelo quadrado da altura em metros); aumento nos níveis de glicose sanguínea; e histórico familiar de DM2⁹⁻¹².

A obesidade, geralmente relacionada ao aumento no IMC, é um fator ambiental muito importante associado ao risco de resistência à sinalização de insulina¹³. Embora a resistência à insulina ocorra na maioria dos tecidos de indivíduos obesos, o DM2 geralmente é prevenido através do aumento na produção de insulina (o qual pode ocorrer por meio do aumento no número de células-beta do pâncreas e/ou pelo aumento da secreção de insulina por cada célula-beta)^{14,15}. Quando o aumento na produção de insulina deixa de ocorrer, a resistência à insulina leva ao aumento na concentração de glicose no sangue e ao DM2¹⁶.

Desse modo, o nível elevado de glicose vem a ser outro importante fator que contribui para o desenvolvimento de complicações crônicas do DM2⁶. Essa característica, assim como a obesidade, também contribui para a resistência à insulina e para a progressão da doença¹⁷. Os mecanismos pelos quais a glicose induz a resistência à insulina ainda não estão completamente claros, mas imagina-se que a desregulação no metabolismo de glicose possa ter papel importante no desenvolvimento dessa resistência^{18,19}. Interessantemente, uma possível consequência descrita da resistência à insulina, é que o aumento de níveis de glicose no sangue poderia estar afetando o processo de aterosclerose, seja através da ativação do gene *NFK β* ou através do aumento na taxa de proliferação e migração de células vasculares de músculo liso (CML). Concentrações de glicose elevadas por si só tem sido mostradas como ativando o gene *NFK β* , que foi proposto como regulando a expressão de vários genes inflamatórios relacionados ao processo de aterosclerose^{20,21}. A hiperglicemia é também

descrita como induzindo diretamente a aterosclerose, processo o qual envolve uma complexa série de eventos incluindo a proliferação e migração anormais de células vasculares de músculo liso²², os quais contribuem de modo importante para a formação organizada da placa aterosclerótica²³.

1.2 Relação de variantes genéticas com DM2

Apesar dos importantes papéis etiológicos do estilo de vida e de fatores ambientais, há anos tem sido reconhecido que fatores genéticos são importantes para o desenvolvimento do DM2. As primeiras evidências de um componente genético na patologia da doença vêm de estudos com famílias grandes, estudos com gêmeos e irmãos não gêmeos, e de estudos de adoção^{9,24-28}.

Nos últimos anos, tem sido um desafio identificar variantes genéticas que expliquem o excesso de risco de DM2 associado com o histórico familiar da doença. Recentemente, estudos de associação entre variantes genéticas e doenças multifatoriais complexas (como os *genome-wide association studies*, GWAs) têm conseguido muitos resultados, com muitos novos *loci* identificados associados a doenças complexas, com níveis de probabilidade nunca antes obtidos²⁹.

Inicialmente, de uma longa lista de genes candidatos, apenas três variantes (as quais estavam inseridas nos genes *TCF7L2*, *KCNJ11* e *PPAR γ* , respectivamente) foram consistentemente associadas com DM2³⁰⁻³². Contudo, em 2007, novas variantes genéticas nos genes *CDKAL1*, *IGF2BP2*, *FTO*, *SLC30A8*, *WFS1*, em um *locus* perto dos genes *CDKN2A* e *CDKN2B* e em um *locus* perto do gene *HHEX*, foram associadas à maior susceptibilidade de DM2³³⁻³⁹. Mais adiante, uma meta-análise identificou seis novas variantes polimórficas associadas com DM2 nos genes *JAZF1*, *THADA*, *ADAMTS9*, *NOTCH2*, em um *locus* perto dos genes *CDC123* e *CAMK1D* e em um *locus* perto dos genes *TSPAN8* e *LGR5*⁴⁰. Até o momento, mais de 38 variantes comuns já foram consistentemente associadas ao DM2^{41,42}.

Apesar dos múltiplos *loci* associados ao DM2 em estudos de GWA⁴³, a variante rs7903146, a qual se situa em um *intron* do gene *TCF7L2*, é, até o momento, o marcador genético mais significativamente associado com o risco da doença, sendo consistentemente associada com DM2 em diferentes populações, incluindo a brasileira^{32,37,44-46}. Interessantemente, um estudo realizado por nosso grupo de pesquisa obteve uma associação significativa do polimorfismo rs7903146 com a gravidade da doença arterial coronariana (DAC) e com mortalidade em indivíduos não-diabéticos⁴⁶. Outro estudo, realizado por Muendlein *et al*, também foi capaz de identificar uma associação significativa da variante rs7903146 com DAC, porém, em pacientes com DM2²¹.

Alguns estudos realizados atentaram para a investigação da influência do gene *TCF7L2* na homeostase glicolítica e, conseqüentemente, no DM2^{47,48}. Já o estudo realizado por Muendlein *et al*, hipotetizou a influência desse gene no processo de aterosclerose, tanto por intermédio do gene *NFK β* , quanto diretamente afetando a regulação da proliferação de células vasculares de músculo liso²¹. De fato, o gene *TCF7L2* é descrito como envolvido na via de sinalização *Wingless* (Wnt), a qual tem seus produtos finais de sinalização atuando sobre genes relacionados ao ciclo celular⁴⁹ e, conseqüentemente, à proliferação². Contudo, o papel funcional do *locus* do gene *TCF7L2* associado ao DM2 e à DAC ainda são pouco conhecidos. Desse modo, mostram-se necessários mais estudos para entender a influência do rs7903146 na patologia dessas doenças.

1.3 Estudo das variantes genéticas para entendimento de fenótipos celulares

Uma maneira de se entender a contribuição de variantes genéticas em doenças é estudar como o genótipo pode afetar fenótipos celulares. Existem mais de vinte milhões de polimorfismos de único nucleotídeo (SNPs)⁵⁰ e outras variantes na seqüência do genoma humano, como polimorfismos de número de cópias⁵¹. Embora a maioria dessas variantes seja aparentemente neutra, algumas são funcionais. Assim, um dos desafios atualmente tem sido

identificar essas variantes funcionais, pois os efeitos funcionais de um polimorfismo do ácido desoxirribonucléico (DNA) em doenças multifatoriais podem ser mediados por inúmeros mecanismos.

Apesar dos polimorfismos que alteram a função da proteína poderem ter efeitos muito importantes, não foram observados muitos resultados significantes de efeitos de SNPs não-sinônimos conhecidos em doenças complexas⁵². Atualmente, uma das hipóteses que tem se destacado é que a variação na expressão gênica é um dos mecanismos mais importantes por trás da suscetibilidade de doenças complexas, sendo que polimorfismos em regiões regulatórias podem estar diretamente modificando a abundância de um produto gênico²⁹.

1.4 Bases genéticas da variação na expressão gênica

Uma das mais importantes consequências da obtenção de marcadores genéticos associados a variações de níveis de expressão gênica é a ligação que ela pode fornecer sobre a causa e a consequência de doenças.

Está bem estabelecido que os níveis de expressão gênica, além de serem controlados por fatores ambientais, são controlados por uma combinação de reguladores *cis* e *trans* ativos* (por exemplo, a ligação de fatores *trans*, como fatores de transcrição, a seqüências regulatórias alvo em *cis*)⁵³. Resultados de mapeamentos genéticos sugerem que, em humanos, os efeitos individuais de variantes em *cis* são normalmente mais fortes do que os efeitos de variantes em *trans* – e isso deve ocorrer provavelmente, pois genes são normalmente influenciados por muitos reguladores em *trans*. Considerando que existe normalmente um ou poucos reguladores em *cis* que afetam o gene alvo, seus efeitos são maiores na expressão gênica e teoricamente mais fáceis de se observar⁵³.

Em humanos, uma das dificuldades para observar os efeitos de polimorfismos na expressão gênica é a de obter tecidos para esses estudos;

* As definições dos determinantes como *cis* e *trans* apenas fazem referência à distância do determinante da regulação em relação ao gene alvo, sendo que denominar reguladores ou seqüências regulatórias em *cis* remete ao fato que eles estão perto dos genes alvo, e denominar como *trans* que estão mais distantes dos genes alvo, deste modo essa definição não tem uma significância funcional.

de modo que uma grande quantidade de amostras seria necessária para se identificar significativamente pequenos efeitos de variantes genéticas na variação da expressão gênica.

1.5 Tipos de células usadas em análises de expressão gênica relacionados ao DM2

Um dos pontos de partida para a análise da expressão gênica é definir quais tipos de tecidos e de células usar – e isso depende da disponibilidade de amostras e do objetivo do projeto.

Os tipos de células usadas em estudos de expressão gênica são as células não manipuladas e as células cultivadas. Células não manipuladas de indivíduos têm o problema de carregarem consigo as exposições sofridas pelos doadores (como medicação e/ou dieta), as quais são difíceis de controlar. As exposições (fatores ambientais) podem ter influência significativa na expressão gênica, portanto, podem diminuir a capacidade de influências genéticas serem identificadas na expressão gênica⁵⁴. Já as células cultivadas podem crescer sob condições controladas e seu uso pode minimizar a influência ambiental na expressão gênica⁵³.

A análise genética de fenótipos e de doenças humanas tem encontrado muitos empecilhos, tanto pelo tamanho e pela complexidade do genoma humano, quanto pelo fato de humanos não serem organismos experimentais. O uso de tecidos humanos acarreta em muitos problemas que precisam ser resolvidos, pois, amostras de tecidos normais e com doença, além de serem difíceis de obter, requerem a atenção para questões éticas, legais e sociais.

Até o momento, os tecidos relacionados ao DM2 mais estudados, através de análises de expressão gênica, foram: o adiposo, o pancreático, o hepático e o intestinal^{48,55,56}, sendo que, para esses estudos, não foram obtidos números amostrais grandes. Outros tecidos foram pouco estudados nessa área específica; assim, a literatura apresenta grande lacuna referente a dados de expressão de outros tecidos de relevância para a patologia do DM2, como, por exemplo, o músculo liso vascular.

Embora seja difícil coletar amostras de tecidos e células humanas, estudos futuros de expressão gênica precisam continuar a identificar meios para que sejam analisados novos tipos celulares e para que sejam obtidos maiores tamanhos amostrais. Desse modo, tais estudos podem resultar em um mapa mais detalhado de mecanismos e de variantes regulatórias que influenciem a expressão gênica humana.

1.6 Análise de expressão gênica

As técnicas de análise de níveis de transcritos são as que mais têm sido utilizadas para se estudar a expressão gênica e suas respectivas variações entre indivíduos. Essas técnicas, se comparadas com as de análises de níveis de proteínas, possuem maior custo-eficiência com relação à sensibilidade aumentada e a reprodutibilidade dos experimentos⁵⁷⁻⁵⁹.

Um estudo que pode ser feito de maneira simples utilizando a análise de níveis de transcritos, e que está entre as estratégias mais difundidas para se estudar fenótipos de transcrição celular, é o estudo quantitativo de expressão gênica associada a um *locus* (eQTLs). Nesse tipo de análise, efeitos genéticos nos níveis de transcritos são mapeados, ou associados, a polimorfismos genéticos⁵⁵. Usualmente, para esse tipo de estudo são aplicadas análises testando-se a significância estatística da diferença de expressão de determinado gene entre dois ou mais grupos genotípicos⁶⁰.

Em humanos, estudos iniciais de análises de expressão foram feitos de modo modesto com pequenos números de amostras. A maioria desses estudos de expressão usou medições de expressão de transcritos obtidas pelas técnicas de microarray e de transcrição reversa em tempo real de reações em cadeia da polimerase (RT-PCR em tempo real). Ambas as técnicas apresentam vantagens como rapidez, rendimento e alto grau de automatização, quando comparadas com técnicas como as de análise por Northern-blot e ensaios com ribonuclease^{57,59}. Porém, quando essas técnicas são comparadas entre si, observamos que o RT-PCR em tempo real possui maiores sensibilidade e especificidade, o que pode ser traduzido na necessidade de menor número amostral⁶¹.

1.7 Outras aplicações de dados de expressão gênica

A genética e a biologia celular tradicionais se baseiam na suposição de que um simples estímulo (ex. variante de DNA), quando aplicado em uma célula (ou gene), teria um resultado simples, afetando um ou poucos genes. A realidade é que mesmo um estímulo simples poderá induzir mudanças na transcrição de muitos genes, os quais interagem em redes gênicas complexas, com uma resposta que afeta muitos transcritos diferentes e muitos processos.

Recentemente, análises de redes gênicas incorporando dados de expressão gênica forneceram novas e importantes percepções sobre mecanismos causadores de doenças multifatoriais⁶²⁻⁶⁴. Com base nisso, estudos recentes têm ido além dos testes de diferenças de expressão de um único gene e, comparando condições ou tipos biológicos diferentes, começaram a procurar por diferentes padrões de expressão e de interação gênica, considerando vários genes⁶⁵⁻⁶⁷. A identificação desses diferentes padrões entre grupos pode fornecer melhor entendimento de como os genes se comportam em grupo e das diferenças entre interações gênicas, permitindo a inferência sobre as causas centrais de desregulação e suas respectivas consequências em certa via ou sistema durante determinada patologia⁶⁸.

1.8 Tipos de redes de expressão gênica

A fim de considerar os genes dentro de um contexto celular (via ou sistema), alguns métodos têm sido desenvolvidos para testar mudanças múltiplas de expressão em uma via ou em um grupo de genes previamente definidos^{65,69,70}, assim como para identificar sub-redes expressas diferencialmente dentro de redes de interações protéicas⁷¹. Acredita-se que grupos de genes ou vias nos quais os níveis de expressão estão subindo ou descendo, se comparados com grupos controle, estejam associados com algum fenótipo de patologia.

Contudo, alguns conhecidos genes causadores de doenças não estão expressos diferencialmente durante o processo patológico. Por exemplo:

mutações em regiões codificadoras podem afetar a função do gene sem afetar seu nível de expressão; ou algumas alterações pós-tradução podem afetar a atividade da proteína independentemente do nível de expressão estar alterado. Desse modo, é importante tentar entender também mudanças no modo com que os genes se relacionam. Para alcançar esse objetivo, algumas abordagens têm sido desenvolvidas, de modo a procurar identificar como as relações entre atividades gênicas mudam entre dois tipos biológicos.

A análise de correlação de expressão gênica é uma das possibilidades para se tentar identificar tais relações gênicas, podendo ser usada para elucidar redes de regulação disfuncionais em determinada doença⁶⁸. Baseando-se em correlações de expressão gênica, podem ser desenhadas conexões entre genes altamente correlacionados⁷², fornecendo informação sobre como cada gene está conectado a outros na rede e identificando genes que estão mais conectados do que outros.

Deste modo, através de um diagrama resultante da conectividade, pode-se examinar o grupo inteiro de genes altamente correlacionados, ao invés de se examinar apenas relações de genes dois-a-dois.

Observar como relações de co-expressão de genes mudam entre duas redes é um meio potencialmente poderoso de se obter a compreensão geral de como as relações gênicas mudam entre dois estados, ou tipos genotípicos, sendo que um dos quais está associado a uma doença⁷³.

Estudos de eQTL, utilizando abordagens de redes gênicas, têm se mostrado uma estratégia importante para ajudar a identificar os reguladores e os alvos nas redes^{64,74,75}. Com esse presente trabalho, utilizando tais abordagens, esperamos ajudar a entender como o polimorfismo rs7903146 pode afetar os fenótipos de DM2 e DAC.

1.9 Justificativa

O diabetes mellitus tipo 2 representa um dos mais sérios problemas de saúde pública da atualidade, tanto pelo grande número de pessoas acometidas como também pelas complicações crônicas, que podem levar a

doenças cardiovasculares, causando incapacitações e mortalidade prematura. Conseqüentemente, o DM2 e as doenças cardiovasculares ocasionam alto custo com os seus tratamentos e os de suas co-morbidades.

A etiologia destas doenças é conhecida; entretanto, os mecanismos pelos quais os fatores genéticos de risco influenciam no desenvolvimento das mesmas ainda não foram claramente elucidados. Estudos apontam a presença de polimorfismos, em/ou perto de genes nucleares, que estão associados às doenças e que poderiam estar modulando a expressão gênica em células de alta relevância para as patologias – como as células vasculares de músculo liso. Contudo, estudos posteriores ainda são necessários para esclarecer o modo de contribuição desses polimorfismos para as doenças.

Com base no conhecimento atual, tem se mostrado cada vez mais importante estudar o comportamento em conjunto dos genes dentro de sistemas. Deste modo, nos propusemos a elaborar um diagrama resultante da conectividade entre genes relacionados ao DM2 e às doenças cardiovasculares, afim de que se examinasse o comportamento de um grupo de genes e o modo como os mesmos podem interagir, de acordo com um estímulo genético (rs7903146). A elaboração desse diagrama pode nos propiciar um avanço no entendimento da patologia das doenças.

2. OBJETIVOS

2.1. Objetivo geral

O objetivo do presente trabalho foi entender melhor o comportamento de grupos de genes relacionados ao DM2 e à DAC, e de suas respectivas interações, de acordo com o tipo genotípico (com e sem alelo associado ao risco de DM2 e DAC). Para isso, elaboramos e analisamos, utilizando células vasculares de músculo liso de 92 indivíduos, redes de expressão e de correlações de expressão gênica com os genes: *TCF7L2*, *PPAR γ* , *CDKAL1*, *JAZF1*, *IGF2BP2*, *WFS1*, *CDKN2A*, *CDKN2B*, *CAMK1D* e *HHEX* (os quais possuem ou estão perto de variantes genéticas associadas ao DM2); e outros 31 genes relacionados às vias de sinalização de DM2 ou às vias de proliferação celular. A construção dessas redes teve por finalidade observar de modo global se as relações de expressão e co-expressão dos genes mudavam entre dois tipos genotípicos (com e sem alelo associado ao risco de DM2 e de DAC, para o polimorfismo rs7903146).

2.2. Objetivos específicos (Figura 1)

- Determinar o perfil genotípico do polimorfismo rs7903146 associado ao DM2 e a DAC em uma amostra de 104 indivíduos e separar os indivíduos pelos genótipos (indivíduos com e sem alelo associado ao risco das doenças);
- Cultivar células vasculares de músculo liso da artéria mamária dos 104 indivíduos e extrair o RNA dessas células;
- Verificar, nas células dos indivíduos, o perfil de expressão dos genes *TCF7L2*, *PPAR γ* , *CDKAL1*, *JAZF1*, *IGF2BP2*, *WFS1*, *CDKN2A*, *CDKN2B*, *CAMK1D*, *HHEX*, e de outros 31 genes relacionados ao DM2 e à DAC;
- Formar redes de expressão gênica e comparar os padrões entre os grupos genotípicos
- Formar redes de correlação de expressão e comparar os padrões entre os grupos genotípicos.

Figura 1 – Resumo dos objetivos específicos. **A.** Obtenção de artérias mamárias e de sangue dos pacientes do estudo. **B.** Genotipagem e divisão dos pacientes por grupo genotípico (grupo sem alelo de risco X grupo com alelo de risco). **C.** Obtenção de células de músculo liso. **D.** Células de músculo liso tiveram a expressão gênica medida por RT-PCR em tempo real. **E e F.** Formação de redes de expressão gênica e de co-expressão gênica, e comparação dessas redes entre os grupos (indivíduos com alelo de risco X sem alelo de risco). **G e H.** Comparação dos padrões de expressão e co-expressão gênica entre os grupos.

3. MÉTODOS

3.1 Amostras do estudo e termo de consentimento livre e esclarecido

Inicialmente, selecionamos células cultivadas de tecido muscular liso da artéria mamária de 104 pacientes do Instituto do Coração, Faculdade de Medicina da Universidade de São Paulo (InCor/FM-USP). Os segmentos de artéria mamária humana foram obtidos de pacientes submetidos à cirurgia de revascularização miocárdica no InCor, sendo o critério de inclusão apenas a disponibilidade de células. Foram também obtidas amostras de sangue periférico destes pacientes para genotipagem.

O presente projeto posteriormente utilizou o ácido ribonucléico (RNA) extraído das amostras de células de músculo liso e o DNA extraído de leucócitos para ensaios de quantificação de expressão gênica e para os ensaios de genotipagem, respectivamente. Todos os indivíduos participantes leram e assinaram o respectivo termo de consentimento para participação do estudo, o qual foi revisado e aprovado pelo comitê de ética local (SDC 2454/04/074 - CAPPesq 638/04). O inquérito estava em conformidade com os princípios enunciados na declaração de Helsinque (ver Cardiovascular Research 1997; 35:2-4).

3.2 Genotipagem de 92 indivíduos e separação em dois grupos genotípicos

3.2.1 Extração do DNA

Foram coletados 10mL de sangue periférico em tubo de hemograma contendo EDTA para o processo de genotipagem. A extração do DNA foi realizada como descrito por Miller *et al*⁷⁶. Resumidamente, o sangue foi hemolisado com tampão contendo NH₄Cl (0,144M) e NH₄CO₃ (0,001M). Em seguida, os leucócitos foram *lisados* utilizando as soluções: Tris (0,01M), NaCl (0,4M), EDTA (0,5M pH 8,0), e SDS (10%). Posteriormente, as proteínas foram precipitadas em solução saturada de NaCl, e o DNA foi precipitado em etanol absoluto para depois ser diluído em TE 1X (Tris-HCl 10mM, EDTA 1mM). A concentração da solução de DNA obtida foi

determinada com leitura em espectrofotômetro à 260nm. No próximo passo, a solução de DNA foi diluída em TE, obtendo-se uma concentração final de 20ng/mL para uso. Por fim foram obtidas 92 amostras de DNA dos pacientes, as quais foram utilizadas para posterior genotipagem. Os dados desses pacientes estão dispostos na tabela 1.

Tabela 1 - Características clínicas das 92 amostras do estudo

CARACTERÍSTICAS	PACIENTES (n=92)	GRUPO SELVAGEM (n=48)	GRUPO DE RISCO (n=44)
IDADE [ANOS] (MÉDIA ± DP)	66.4 ± 9.8	66.1 ± 9.9	66.7 ± 9.7
SEXO FEMININO (%)	23.9	36.4	63.6
DM2 OU GJA* (%)	41.4	41.2	58.8
CRCCP 1 (%)	7.6	42.9	57.1
CRCCP 2 (%)	37.0	61.8	38.2
CRCCP 3 (%)	52.2	45.8	54.2
CRCCP 4 (%)	3.3	66.7	33.3

CRCCP = Ciclo de replicação celular em cultura primária; DM2 = Diabetes mellitus tipo 2; DP = Desvio padrão; e GJA = Glicemia de jejum alterada. * O "n" do número de indivíduos para a variável DM2/GJA é 82.

3.2.2 Seleção dos oligonucleotídeos para genotipagem

A seqüência do gene *TCF7L2*, que possui o polimorfismo rs7903146, foi obtida no endereço eletrônico <http://genome.ucsc.edu/> da Universidade Santa Cruz da Califórnia (UCSC)⁷⁷, a fim de se desenhar os oligonucleotídeos iniciadores (*primers*) da amplificação do fragmento de interesse. Tais *primers* foram criados com o auxílio do programa *primer3*⁷⁸, desenvolvido pelo *Howard Hughes Medical Institute*, EUA, de modo que suas seqüências são: 5' CCG TCA GAT GGT AAT GCA GA 3' e 5' GTA GCA GTG AAG TGC CCA AG 3' para o PCR/HRM, e 5' CTG AAC AAT TAG AGA GCT AAG CA 3' e 5' ACT AAG GGT GCC TCA TAC GG 3' para o seqüenciamento. Posteriormente os *primers* foram reconstituídos em 500µL de TE 1X e utilizados na concentração de 5pM.

3.2.3 Padronização da amplificação dos fragmentos de interesse

3.2.3.1 Reação em cadeia da polimerase

A região do polimorfismo rs7903146 foi amplificada a partir do DNA genômico com auxílio do método de PCR. De maneira geral, o protocolo desenvolvido por amostra constou de: 2µL de DNA (20ng/ml), 0,02µL da solução de desoxirribonucleotídeos trifosfatados (10mM), 1µL de 10x Assay Buffer (Ultra Chem®), 2µL de MgCl² (10 mM), 0,4µL de oligonucleotídeos iniciadores senso e 0,4µL de oligonucleotídeos iniciadores antisenso (5mM), e 0,06µL de enzima Taq polimerase (Easy Taq Invitrogen; 0,625 U), sendo adicionada água milliQ no final para totalizar um volume de 10µL. As reações de PCR foram realizadas em um termociclador (MiniCycler MJ Research), seguindo as seguintes condições para o referido par de *primers*: 94°C por 5 minutos de desnaturação inicial; 35 ciclos de 94°C por 30 segundos, 59° por 30 segundos (anelamento), 72°C por 40 segundos; e extensão final a 72°C por 5 minutos.

3.2.3.2 Eletroforese em gel de agarose

Os produtos de amplificação obtidos pela PCR foram analisados em gel de agarose 1,0%. Para tanto, 1,0g de agarose foi suspensa em 100ml de tampão tris-acetato-EDTA e fervida em forno de microondas até dissolução completa. Brometo de etídeo (25mM) foi utilizado na proporção de 50µl para 100mL de gel antes de sua polimerização. O tampão utilizado no preparo das amostras era composto de glicerol 30%, azul de bromofenol 0,05% e xileno-cianol 0,05%. A eletroforese foi realizada aplicando-se aproximadamente 120 Volts em uma cuba horizontal com tampão tris-acetato-EDTA. O marcador de peso molecular utilizado foi o DNA “Ladder” de 100 pb (Gibco-BRL). Os fragmentos amplificados foram confirmados de acordo com o tamanho e visualizados pela exposição do gel à luz ultravioleta no equipamento EagleEye II (Stratagene).

3.2.4 Detecção dos polimorfismos do DNA por High Resolution Melting (HRM)

Após a padronização do PCR em um termociclador normal, realizamos a reação de HRM no Rotor Gene 6000 (Qiagen, Courtaboeuf, France), como

descrito por Santos *et al*⁷⁹. A reação de HRM trate-se de uma reação de PCR com o acréscimo de 1,5 uM de SYTO 9® (Invitrogen, Carlsbad, USA). O SYTO 9® é um corante que se intercala apenas à dupla fita de DNA e emite fluorescência quando associado à molécula. Após cada ciclo a temperatura aumenta de 0,1 em 0,1 grau, partindo de 70°C até 94°C. À medida que a temperatura aumenta, as duplas fitas do DNA se desfazem e o corante deixa de emitir fluorescência. Uma simples troca de nucleotídeos é suficiente para gerar uma curva de dissociação diferente devido ao sutil aumento na temperatura, e desse modo pudemos identificar a presença de genótipos diferentes de acordo com o polimorfismo.

Os diferentes padrões de curvas de dissociação, relativos aos três genótipos possíveis (CC, CT e TT), foram posteriormente identificados por seqüenciamento.

3.2.5 Seqüenciamento automático para genotipagem

De acordo com o polimorfismo rs7903146, foram selecionados três indivíduos de cada padrão de curva de dissociação para serem seqüenciados.

A purificação foi realizada com o produto ExoSAP-IT (USB Corporation), o qual utiliza duas enzimas hidrolíticas, a Exonuclease I e a Shrimp Alkaline Phosphatase (SAP), para remover sobras de oligonucleotídeos e dNTPs, respectivamente. A Exonuclease I hidrolisa os oligonucleotídeos de fita simples residuais e qualquer outra fita simples de DNA “estranha” produzida no PCR. A SAP hidrolisa os dNTPs remanescentes da mistura de PCR que poderiam interferir na reação de seqüenciamento. Adiciona-se a enzima diretamente no produto de PCR em qualquer volume, obedecendo-se sempre a proporção de 2µl de ExoSAP para 5µl de produto. Leva-se a mistura ao termociclador a 37°C por 15 minutos para ativação e ação das enzimas, seguido de um período a 85°C por 15 minutos de inativação. Assim, as amostras estarão prontas para o seqüenciamento.

Todas as amostras purificadas foram submetidas ao seqüenciamento. O equipamento utilizado foi o ABI-3100 (Applied Biosystems) em conjunto com o kit ABI PRISM® BigDye® terminator V 3.1 cycle sequencing (Applied Biosystems). Amostras de 2,5µl dos DNAs amplificados foram adicionadas a 2,5µl de *primers* senso ou antisenso (5pmol) juntamente com 1µl do kit BigDye® terminator, 3µl de BigDye® terminator v1.1 / 5X sequencing buffer v3.1, e 1µl de água, totalizando um volume de 10µl.

A mistura de reação foi purificada por precipitação com a adição de 30µl de etanol absoluto 100% (Merck) em cada poço. A placa das amostras foi coberta com papel alumínio, deixando-a em temperatura ambiente por 10 minutos. Foi realizada a primeira etapa de centrifugação (4000rpm por 45 minutos a 4°C), sendo o sobrenadante desprezado na seqüência. Posteriormente, foi feita uma rápida centrifugação com a placa invertida (300rpm), para depois ser adicionado 100µl de etanol (70%) em cada poço da placa. Dando seqüência ao processo, as amostras foram novamente centrifugadas (4000rpm por 15 minutos a 4°C), a placa foi virada cuidadosamente para descartar o etanol, e as amostras passaram por uma nova etapa de rápida centrifugação com a placa invertida a 300rpm para a retirada do excesso de etanol. No próximo passo, as amostras de DNA foram colocadas no termociclador a 80°C por 2 minutos para que o etanol restante evaporasse, sendo em seguida o precipitado de DNA ressuspenso com 10µl de formamida HI-Di em cada poço. As amostras, depois de passarem por uma nova etapa de rápida centrifugação, foram desnaturadas no termociclador por 3 minutos a 96°C. Após a desnaturação, a placa foi acondicionada imediatamente em gelo por um período de 3 a 5 minutos. Finalizando o processo, as amostras foram colocadas no seqüenciador conforme os parâmetros recomendados pelo fabricante.

3.2.6 Divisão dos indivíduos em dois grupos genotípicos: grupo sem o alelo de risco e grupo com o alelo de risco para DM2 e DAC

De acordo com o polimorfismo rs7903146, os indivíduos foram separados em dois grupos: grupo selvagem (homozigoto sem o alelo de

risco, CC), e grupo de risco (heterozigoto e homozigoto com o alelo de risco, CT e TT) para DM2 e DAC.

3.3 Análise da expressão gênica

3.3.1 Cultura primária de células de músculo liso humanas

Células musculares lisas de artéria mamária humana foram obtidas pelo protocolo descrito a seguir. Resumidamente, a camada de endotélio foi removida por atrito mecânico e pequenos fragmentos (0,09 cm²) de vasos foram colocados em placas de cultura, que possuíam seis poços contendo gelatina a 3%. Os fragmentos foram cultivados com meio Dulbecco Eagle modificado contendo 20% de soro fetal bovino, 100U/ml de penicilina, e 100mg/ml de estreptomicina. Após cerca de duas semanas as CMLs derivadas de fragmentos de vasos foram isoladas e expandidas. No final, as células cultivadas permaneceram em cultura por um período de dois a três meses (1 a 4 passagens/gerações) antes da extração do RNA.

Efeitos da influência de diferentes passagens/gerações celulares na expressão gênica foram posteriormente corrigidos estatisticamente (ler mais abaixo). As CMLs foram caracterizadas pelo padrão de crescimento e pela técnica de coloração por imuno-fluorescência de α -actina de músculo liso, método de caracterização que tem sido utilizado por nosso grupo⁸⁰.

3.3.2 Seleção de formas de *splicing* do gene *TCF7L2* e de outros genes para análise de expressão e co-expressão

Com a finalidade de se decidir quais genes quantificar, como ponto de partida nós selecionamos 10 dos genes que possuem ou estão perto de SNPs que tem os mais baixos p-valores de associação com DM2^{55,81}. Primeiramente, com o propósito de analisar a possível regulação em *cis* do rs7903146 no gene *TCF7L2*, foram quantificadas sete grupos de transcritos (formas de *splicing*)⁷⁷, que já foram estudadas em outros trabalhos⁸²⁻⁸⁵ (Tabela 2 e Figura 2). Selecionamos também genes que, de acordo com o banco de dados de vias de sinalização KEGG⁸⁶, são relacionados a vias de DM2 (por exemplo, vias de sinalização de insulina ou vias de sinalização de DM2); e genes que poderiam ter a expressão afetada por uma alteração da

expressão do gene *TCF7L2*, como os genes relacionados à proliferação celular^{23,87}. A influência da expressão do gene *TCF7L2* na expressão de genes que possuem um papel regulatório do ciclo celular, e conseqüentemente da proliferação celular, é conhecida há algum tempo; e a proliferação de CMLs é descrita como altamente afetada durante a patologia do DM2 e de DAC^{15,21,49}. Usando dados do I2D⁸⁸, uma ferramenta da rede que acessa vários bancos de dados de interação protéica (MINT, BIND, IntACT, HPRD, e etc.), selecionamos também genes com relação protéica muito próxima (relações preditas ou descritas na literatura entre proteínas) aos 10 genes primeiramente selecionados, os quais possuem ou estão próximos de SNPs associados ao DM2. A lista de todos os genes quantificados encontra-se na tabela 3.

Tabela 2 - Formas de *splicing* do *TCF7L2* e ensaios realizados

ENSAIOS	FEITOS PARA QUANTIFICAR:
A	Transcritos com os <i>exons</i> 15 e 17 e sem o <i>exon</i> 16
B	Transcritos com o <i>exon</i> 9 mais longo
C	Transcritos com os <i>exons</i> 14 e 16 e sem o <i>exon</i> 15
D	Transcritos com os <i>exons</i> 14 e 17 e sem os <i>exons</i> 15 e 16
E	Transcritos que não possuem os <i>exons</i> 13, 14, 15 e 16
F	Transcritos com os <i>exons</i> 4 e 5
G	Transcritos com os <i>exons</i> 13 e 14

Figura 2 - Ensaios com formas de *splicing* do *TCF7L2* realizados. Variantes de transcritos do *TCF7L2* e ensaios realizados das formas de *splicing*.

Tabela 3 - Lista dos genes quantificados neste estudo

GENES QUANTIFICADOS	
Genes com ou perto de variantes genéticas associadas ao DM2	Genes relacionados à proliferação celular
<i>CAMK1D</i>	<i>ATP2A2</i>
<i>CDKAL1</i>	<i>BDKRB2</i>
<i>CDKN2A</i>	<i>BMP2</i>
<i>CDKN2B</i>	<i>CALR</i>
<i>HHEX</i>	<i>CDKN1A</i>
<i>IGF2BP2</i>	<i>CHTF18</i>
<i>JAZF1</i>	<i>CSF3</i>
<i>PPARγ</i>	<i>CSRP1</i>
<i>TCF7L2</i>	<i>CSRP2</i>
<i>WFS1</i>	<i>CXCL3</i>
Genes relacionados à vias de DM2	<i>CYBA</i>
<i>AKT1</i>	<i>DTNB</i>
<i>BAD</i>	<i>FKBP1A</i>
<i>CALM1</i>	<i>IGFBP2</i>
<i>ERBB4</i>	<i>MAP4K4</i>
<i>IL1R1</i>	<i>NKAP</i>
<i>NRG1</i>	<i>PLN</i>
<i>RXRα</i>	<i>SPTBN1</i>
<i>TNFα</i>	<i>TOMM20</i>
<i>TXNIP</i>	
Genes selecionados pelo I2D	
<i>JUN</i>	
<i>CDK4</i>	
<i>EP300</i>	

3.3.3 Extração de RNA de CMLs e quantificações de expressão por RT-PCR em tempo real.

O RNA total foi isolado com TRIzol Reagent (Invitrogen) de acordo com as instruções do fabricante, e a síntese do cDNA foi realizada com o kit SuperScriptTM III Reverse Transcriptase (Invitrogen). Para evitar que o DNA genômico fosse amplificado, nós desenhamos *primers* em *exons* diferentes, os quais possuíam um *intron* grande entre eles, desse modo o DNA genômico não seria amplificado em nosso protocolo⁵⁹. Em todos os nossos ensaios de expressão foram observadas bandas específicas para os produtos amplificados. Os *primers* utilizados nos ensaios de RT-PCR em tempo real estão descritos na tabela 4.

Tabela 4 - Lista de *primers* utilizados nos ensaios de RT-PCR em tempo real

Gene	Primer F	Primer R	Isoformas detectadas	Tamanho (pb)
<i>AKT1</i>	GCAGCACGTGTACGAGAAGA	GGTGTCACTCTCCGACGTG	3 de 3	67
<i>ATP2A2</i>	CGTGGGCTCCATCTGCCTGT	GCGGGCCACAACTTGAGCG	2 de 2	169
<i>BAD</i>	CCCAGAGTTTGAGCCGAGTG	CCCATCCCTTCGTCGTCT	2 de 2	249
<i>BDKRB2</i>	CGACATGCTCAATGTCACCT	CTGGATGGTGTGAGCCAG	1 de 1	103
<i>BMP2</i>	GCCAGCCGAGCCAACAC	GGTTGTTTCCCACTCGTTTC	1 de 1	81
<i>CALM1</i>	TCTTCCTCCCTTCGCTCGCACCAT	CCCAGTGACCTCATGACAGTTCCA	1 de 1	144
<i>CALR</i>	CCACCCAGAAATTGACAACCCCG	TGCTGCCTTTGTTACGCCCA	1 de 1	190
<i>CAMK1D</i>	AGCCCTCAACAAAAACATCC	CTGCTCGAAACACTTGCAAT	2 de 2	165
<i>CDK4</i>	CCCGTGGAGCCTTTCCCCC	GCAGCAGCTGTGCTCCCGA	1 de 1	84
<i>CDKAL1</i>	TTTATGAAAGGGCAGCCAG	GCAGCACCATTCTGGAACT	1 de 1	187
<i>CDKN1A</i>	GACTCTCAGGGTCGAAAA	GGATTAGGGCTTCCTCTTGG	4 de 4	93
<i>CDKN2A</i>	ACCAGAGGCAGTAACCATGC	CCTGTAGGACCTTCGGTGAC	4 de 4	129
<i>CDKN2B</i>	TCCCTGGATCACGACGCTGC	TAGTCTGCCTGCGGAACCCG	2 de 2	99
<i>CSF3</i>	CTGGACGTCGCCGACTTT	CCATTCCCAGTCTTCCATCTG	4 de 4	58
<i>CSRP1</i>	GAGCCAGCTGCCAGAATG	CCTTCGCACTGAACCTCTTC	2 de 4	95
<i>CSRP2</i>	GGTGAAATCTATTGTAAAGGATGC	TACTGGGCATGAACAAGAGC	1 de 1	95
<i>CXCL3</i>	CATCCCCATGGTTCAGAAA	AAGTGTCAATGATACGCTGATAAGCT	1 de 1	101
<i>CYBA</i>	CGCTTCACCCAGTGGTACTT	GTGGAGCCCTTCTTCTCTT	1 de 1	98
<i>DTNB</i>	AGCACCAGATGAAGGAGCAT	TGGGTACACACCCCAAAGAT	8 de 8	87
<i>EP300</i>	GGACCGCTTTGTCTACACCT	CTGCTGGTTGTGCTCTCAT	1 de 1	172
<i>ERBB4</i>	TGCCCTACAGAGCCCCAACTA	GCTTGCGTAGGGTGCCATTAC	2 de 2	105
<i>FKBP1A</i>	CCCACCACATGCCACTCTCGT	TGTGCACATGTCTGGAGGCACC	2 de 3	129
<i>HHEX</i>	AGAATCGACGCGCTAAATG	TTCTGTTCCTGAGGCAAATC	1 de 1	121
<i>IGF2BP2</i>	CGTGCCTCGTGACCAAACGC	TGCTGGTGCCTGTGGGAGC	2 de 2	189
<i>IGFBP2</i>	GAGGGCACTTGTGAGAAGC	TTCATACCCGACTTGAGGG	1 de 1	176
<i>IL1R1</i>	GTGAGCCCAGCTAATGAGACA	ACTCAACTGGCCGGTGACA	1 de 1	81
<i>JAZF1</i>	GAAGCCTTTTGCTGCCCAGTT	CGCAGGCCCTGAGCTGTCTT	1 de 1	159
<i>JUN</i>	ACGGCCAACATGCTCAGGGA	AGCCCCGACGGTCTCTCTT	1 de 1	127

Continua

Conclusão Tabela 4

Gene	Primer F	Primer R	Isoformas detectadas	Tamanho (pb)
<i>MAP4K4</i>	CAAAAAGTCTGGTGGACATCGA	CAACCACTTCCACCAGCTCAA	5 de 5	75
<i>NKAP</i>	CAAAGGCTGAAGAACCATCA	AGCCAATTTACCTCTTCGT	1 de 1	171
<i>NRG1</i>	CGTGGAATCAAACGAGATCATCA	GCTTGTCCTCCAGTGGTGGATGT	11 de 17	145
<i>PLN</i>	AAACTCCCCAGCTAAACACCCGT	AGCTTTTGACGTGCTTGTGAGGC	1 de 1	254
<i>PPARγ</i>	TGAGAGCGACTTGGCAATA	CAGGGTGGTTCAGCTTCAG	4 de 4	140
<i>CHTF18</i>	TCTTCTATCGCTCCCATCG	CTGGGCACGGATTTTAATG	1 de 1	180
<i>RXRα</i>	CTTCCTTACCAAGCACATCT	ACCCCTCGCAGCTGTACACT	1 de 1	81
<i>SPTBN1</i>	GATCGACAAGTGGGAAGACC	CTCTCGGCTGGATAGGTACG	2 de 2	127
<i>TCF7L2(A)</i>	GTCGGGCACTGTTCGGGCTT	TGTGACTTGGCGTCTCGGGG	1, 7 e 8	190
<i>TCF7L2(B)</i>	TGCTTCCATGTCCAGCTTTCTGTCT	ATGCCCGTCGTGTGTAGCGT	7	81
<i>TCF7L2(C)</i>	GCGCTCCTAAGAAATGCCGAGC	CCAAATTCAAAGACTGCAAGGGCCG	9 e 10	80
<i>TCF7L2(D)</i>	GCGCTCCTAAGAAATGCCGAGC	TCTCCTGCAAGGGCCGCAC	2 e 5	71
<i>TCF7L2(E)</i>	CAGCCGGGAGAGACCAATGGAG	TCTGAAGAGGGTGGGCTGAGGC	3, 6, 12,13	85
<i>TCF7L2(F)</i>	GGCAGCACACATTACTCTGC	TGAGGGCTTGTCTACTCTGG	1 e 13	121
<i>TCF7L2(G)</i>	GCCTTTCACTTCTCCGATT	GCCGCACCAGTTATTCTGTT	2, 7, 9, 10	89
<i>TNFα</i>	GAAAACAAGGGCTTTTGACAC	GGAATGGTGTCTGGAACCT	3 de 3	113
<i>TOMM20</i>	AGCAAGGAAAGAATGGGCT	CAACGCCTTTTCAAGCATT	1 de 1	171
<i>TXNIP</i>	TGGTGATCATGAGACCTGGA	CGGTCAAGAAAAGCCTTCAC	1 de 1	130
<i>WFS1</i>	GACGAGCTGGCGGGGAAGA	AGCGCGTTGATGTGGTGCG	2 de 2	167

pb = pares de bases.

Resumidamente, cinco nanogramas de cDNA foram utilizadas para a reação de RT-PCR em tempo real (SYBRw Green PCR Master Mix-PE Applied Biosystems), realizada em um aparelho ABI Prism 7700 Sequence Detection System (Applied Biosystems). Para cada indivíduo, foram realizados ensaios em duplicata ou triplicata e a média de expressão de cada gene foi utilizada como a medida padrão. As expressões de transcritos-alvo foram normalizadas pela expressão de mRNA do gene ribossomal 28S (gene referência), usando o método comparativo C_t^{89} . C_t é o número de ciclos no qual uma quantidade de cDNA do gene alvo chega a um limiar fixado, e ΔC_t é a diferença entre o ciclo limiar do gene alvo e do gene de

referência. Posteriormente, o ΔCt linearizado (i.e. $2^{-\Delta Ct}$) foi utilizado para comparações⁹⁰. No final, obtivemos a expressão dos 41 genes para 104 indivíduos.

3.3.4 Efeitos ambientais

Por terem sido utilizadas células cultivadas, passando por algumas gerações em cultura, esperava-se que os efeitos ambientais fossem minimizados e houvesse um predomínio de fatores genéticos na variabilidade da expressão gênica.

Os dois grupos genotípicos foram analisados de acordo com a presença de DM2/GJA (glicemia de jejum alterada), para que se avaliasse uma possível associação entre essa variável e os grupos. Foram também avaliados efeitos de número de ciclos de replicação celular em cultura primária (passagens/gerações), de idade, de DM2/GJA e de sexo nos valores de expressão gênica. Para as análises de efeitos ambientais na expressão gênica, nós utilizamos os dados de expressão do nosso banco de dados total de CMLs, o qual consistia de 104 pacientes.

3.3.5 Montagem de redes de interação protéica

Utilizando dados do I2D, foi montada uma rede de interação protéica para ser utilizada pelos dois grupos genotípicos com: os 41 genes quantificados (uma das formas de *splicing* do gene *TCF7L2* representando o gene, e os outros 40 genes), outros genes que possuem ou estão perto de SNPs associados ao DM2, e genes que tem interações protéicas (preditas ou descritas na literatura) conectando os dois primeiros grupos de genes. Deste modo, cada interação protéica dos genes em questão, descrita na literatura e presente em bancos de dados, foi representada graficamente por uma ligação entre dois nós (círculos). Para cada grupo genotípico, o comportamento em grupo dos genes foi graficamente visualizado utilizando escalas de cor para representar medidas de expressão gênica. Nessa representação gráfica os genes que possuem maiores níveis de expressão são representados em vermelho e os genes com os menores níveis de

expressão são representados em branco. Para manter os valores de expressão gênica em uma mesma faixa de valores nessas representações gráficas, tais valores foram centrados e ajustados pelo desvio padrão da distribuição de expressão gênica. O programa Cytoscape v.2.8.0 e o protocolo descrito por Cline *et al* foram utilizados para a montagem das redes^{91,92}.

3.3.6 Comparação de padrões de co-expressão gênica entre grupos e representação dos padrões em redes

Com a finalidade de se observar se as correlações de expressão gênica eram estatisticamente diferentes entre os dois grupos genotípicos, analisamos as correlações de expressão dos 41 genes dois a dois (uma das formas de *splicing* do gene *TCF7L2* representando o gene, e os outros 40 genes), e comparamos todos os coeficientes de correlação entre os dois grupos. O objetivo principal dessas análises foi observar os genes que possuíam o maior número de correlações diferentes com significância estatística entre os dois grupos genotípicos. Com o intuito de se observar diferenças de relações entre as formas de *splicing* do gene *TCF7L2* analisadas, montamos redes de co-expressão com apenas as formas de *splicing* para cada grupo genotípico. As redes de correlação de expressão montadas com os dados dos 41 genes foram construídas e visualizadas utilizando o programa Cytoscape v.2.8.0. Assim, foram construídas redes com a utilização de faixas de limiares de coeficientes de correlação. Ex. Redes de correlação com faixas de limiares de $\pm 0,4$ ilustram genes que possuem valores de correlação de expressão com outros genes acima de 0,4 ou abaixo de -0,4, fazendo com que apenas coeficientes de correlação com alto poder estatístico (acima de 80%) sejam visualizados. Deste modo, em tais redes uma ligação foi desenhada entre genes que possuem valores de correlação acima dos valores positivos e abaixo dos valores negativos. As redes dos dois grupos foram comparadas topologicamente utilizando o

plug-in NetworkAnalyzer v.2.7⁹³, com o qual se obteve os genes que possuíam o maior número de correlações de expressão diferentes.

3.3.7 Análises estatísticas

Primeiramente, utilizamos o teste de Qui-quadrado para estimar o equilíbrio de Hardy-Weinberg (HWE) para os genótipos, assim como para testar a associação entre DM2/GJA com os grupos genotípicos.

Análises de influência das variáveis idade, sexo e DM2/GJA nas medições de expressão, foram realizadas através de ANOVA de um caminho, com o auxílio do programa SPSS v.13 para o Windows. Para os genes que mostraram sofrer influência de efeitos ambientais, nós realizamos ajustes por regressão linear dos valores de expressão ($2^{-\Delta Ct}$) para determinada variável e posteriormente trabalhamos com os resíduos desse ajuste.

Observamos que muitos dos genes analisados não apresentaram distribuição normal dos dados de quantificação de expressão, deste modo, as comparações de expressão gênica entre os grupos genotípicos e entre grupos com presença ou ausência de DM2/GJA foram feitas através da comparação de valores medianos de expressão. Tais comparações foram realizadas através do teste de Mann-Whitney-U, utilizando o programa SPSS v.13 para o Windows, e consideramos significantes p-valores de comparações de expressão iguais ou menores que 0,05, ou iguais ou menores que 0,01.

As correlações de expressão gênica dois a dois foram obtidas através de correlações de Spearman. Posteriormente, os coeficientes de correlação de cada par de genes foram comparados entre os grupos através de testes para se avaliar a significância da diferença entre dois coeficientes de correlação⁹⁴. Para essas análises consideramos significantes p-valores iguais ou menores que 0,05, ou iguais ou menores que 0,01, para diferença entre os coeficientes de correlação.

4. RESULTADOS

4.1 Análises de efeitos ambientais

Analisando os 92 indivíduos genotipados (doadores das linhagens celulares), obtivemos as seguintes frequências genótípicas para o rs7903146: CC = 52.2% (n = 48); CT = 38.0% (n = 35); TT = 9.8% (n = 9). Deste modo, não foi observado afastamento do HWE (P-valor $\cong$ 0,5). Trinta e quatro indivíduos apresentavam DM2 ou glicemia de jejum alterada (GJA), sendo que quatorze destes (41,18%) eram do grupo sem o alelo de risco e vinte (58,82%) eram do grupo com o alelo de risco (p-valor 0,1) (Tabela 1).

Quando testamos os efeitos de idade, sexo, e DM2/GJA nas medidas de expressão gênica, observamos que o sexo dos indivíduos estava influenciando na expressão do gene *MAP4K4*, e a idade estava influenciando na expressão das formas de *splicing* D e F do gene *TCF7L2* (Tabela 5). Observamos também que o número de ciclos de replicação celular em cultura primária (CRCCP) estava influenciando na expressão de vários genes (Tabela 5).

Trabalhando com os valores ajustados para CRCCP, quando separamos os indivíduos em grupos de diabéticos e não diabéticos, e comparamos os níveis medianos de expressão das sete formas de *splicing* do gene *TCF7L2* e dos outros 40 genes analisados, observamos que apenas os genes *JUN* e *ERBB4* tiveram diferenças significativas em expressão (P-valor $\leq$ 0,05) (Tabelas 6 e 7).

Tabela 5 - Efeitos ambientais nas medidas de expressão gênica

	CRCCP (n=104)	SEXO (n=104)	IDADE (n=104)	DM2/GJA (n=93)
GENES	P-Valor	P-Valor	P-Valor	P-Valor
<i>AKT1</i>	0.731	0.326	0.267	0.269
<i>ATP2A2</i>	0.000*	0.944	0.487	0.357
<i>BAD</i>	0.486	0.354	0.897	0.802
<i>BDKRB2</i>	0.516	0.975	0.908	0.726
<i>BMP2</i>	0.011*	0.207	0.881	0.483
<i>CALM1</i>	0.318	0.278	0.637	0.132
<i>CALR</i>	0.061	0.104	0.406	0.084
<i>CAMK1D</i>	0.214	0.671	0.286	0.311
<i>CDK4</i>	0.003*	0.272	0.473	0.433
<i>CDKAL1</i>	0.000*	0.975	0.274	0.226
<i>CDKN1A</i>	0.488	0.439	0.289	0.268
<i>CDKN2A</i>	0.478	0.543	0.414	0.353
<i>CDKN2B</i>	0.005*	0.267	0.115	0.175
<i>CHTF18</i>	0.737	0.338	0.660	0.136
<i>CSF3</i>	0.041*	0.111	0.238	0.769
<i>CSRP1</i>	0.014*	0.874	0.748	0.831
<i>CSRP2</i>	0.417	0.485	0.723	0.816
<i>CXCL3</i>	0.011*	0.119	0.946	0.456
<i>CYBA</i>	0.900	0.587	0.713	0.496
<i>DTNB</i>	0.342	0.861	0.716	0.667
<i>EP300</i>	0.000*	0.159	0.383	0.484
<i>ERBB4</i>	0.093	0.881	0.962	0.394
<i>FKBP1A</i>	0.002*	0.606	0.601	0.655
<i>HHEX</i>	0.679	0.369	0.699	0.376
<i>IGF2BP2</i>	0.000*	0.670	0.399	0.210
<i>IGFBP2</i>	0.002*	0.366	0.518	0.640
<i>IL1R1</i>	0.000*	0.431	0.509	0.655
<i>JAZF1</i>	0.000*	0.654	0.340	0.384
<i>JUN</i>	0.009*	0.769	0.931	0.201

Continua

			Conclusão	Tabela 5
	CRCCP (n=104)	SEXO (n=104)	IDADE (n=104)	DM2/GJA (n=93)
GENES	P-Valor	P-Valor	P-Valor	P-Valor
<i>MAP4K4</i>	0.396	0.045*	0.155	0.585
<i>NKAP</i>	0.742	0.347	0.466	0.175
<i>NRG1</i>	0.357	0.581	0.752	0.667
<i>PLN</i>	0.008*	0.630	0.371	0.457
<i>PPARγ</i>	0.490	0.316	0.786	0.421
<i>RXRα</i>	0.001*	0.543	0.739	0.689
<i>SPTBN1</i>	0.764	0.151	0.143	0.368
<i>TCF7L2 (A)</i>	0.001*	0.876	0.084	0.369
<i>TCF7L2 (B)</i>	0.000*	0.959	0.111	0.299
<i>TCF7L2 (C)</i>	0.000*	0.601	0.221	0.650
<i>TCF7L2 (D)</i>	0.000*	0.637	0.030*	0.342
<i>TCF7L2 (E)</i>	0.000*	0.669	0.078	0.427
<i>TCF7L2 (F)</i>	0.000*	0.251	0.032*	0.210
<i>TCF7L2 (G)</i>	0.063	0.390	0.969	0.283
<i>TNFα</i>	0.438	0.253	0.850	0.949
<i>TOMM20</i>	0.601	0.682	0.779	0.169
<i>TXNIP</i>	0.104	0.484	0.723	0.549
<i>WFS1</i>	0.000*	0.558	0.934	0.823

*Influência significativa (p-valor $\leq$ 0.05). CRCCP = Ciclo de replicação celular em cultura primária; DM2 = Diabetes mellitus tipo 2; e GJA = Glicemia de jejum alterada.

Tabela 6 - Diferenças de medianas de expressão gênica entre os grupos de diabéticos e não diabéticos de 40 genes

EXPRESSÃO				EXPRESSÃO			
GENE	GRUPO	MEDIANA	P-VALOR	GENE	GRUPO	MEDIANA	P-VALOR
<i>AKT1</i>	GND	0.00006	0.817	<i>EP300</i>	GND	-0.00204	0.274
	GD	0.00004			GD	-0.00137	
<i>ATP2A2</i>	GND	-0.08860	0.122	<i>ERBB4</i>	GND	0.00026	0.041*
	GD	-0.06315			GD	0.00005	
<i>BAD</i>	GND	0.00254	0.432	<i>FKBP1A</i>	GND	-0.17727	0.218
	GD	0.00400			GD	-0.08206	
<i>BDKRB2</i>	GND	0.00203	0.701	<i>HHEX</i>	GND	0.00009	0.774
	GD	0.00245			GD	0.00011	
<i>BMP2</i>	GND	-0.00031	0.906	<i>IGF2BP2</i>	GND	-0.00473	0.177
	GD	-0.00045			GD	-0.00238	
<i>CALM1</i>	GND	0.00185	0.931	<i>IGFBP2</i>	GND	-0.00017	0.197
	GD	0.00103			GD	-0.00023	
<i>CALR</i>	GND	0.00531	0.465	<i>IL1R1</i>	GND	-0.00001	0.512
	GD	0.00339			GD	-0.00003	
<i>CAMK1D</i>	GND	0.00879	0.330	<i>JAZF1</i>	GND	-0.00324	0.282
	GD	0.01028			GD	-0.00177	
<i>CDK4</i>	GND	-0.02490	0.272	<i>JUN</i>	GND	-0.01015	0.016*
	GD	-0.02202			GD	-0.00425	
<i>CDKAL1</i>	GND	-0.00056	0.263	<i>MAP4K4</i>	GND	0.00023	0.535
	GD	-0.00050			GD	0.00021	
<i>CDKN1A</i>	GND	0.01995	0.334	<i>NKAP</i>	GND	0.00002	0.121
	GD	0.01833			GD	0.00005	
<i>CDKN2A</i>	GND	0.00046	0.280	<i>NRG1</i>	GND	0.00003	0.186
	GD	0.00041			GD	0.00006	
<i>CDKN2B</i>	GND	-0.02951	0.162	<i>PLN</i>	GND	-0.00077	0.534
	GD	-0.01302			GD	-0.00066	
<i>CHTF18</i>	GND	0.00044	0.808	<i>PPARγ</i>	GND	0.02278	0.978
	GD	0.00048			GD	0.02328	
<i>CSF3</i>	GND	-0.00054	0.487	<i>RXRα</i>	GND	-0.00067	0.994
	GD	-0.00137			GD	-0.00076	
<i>CSRP1</i>	GND	-0.00216	0.863	<i>SPTBN1</i>	GND	0.00364	0.520
	GD	-0.00225			GD	0.00417	
<i>CSRP2</i>	GND	0.00020	0.604	<i>TNFα</i>	GND	0.00015	0.421
	GD	0.00019			GD	0.00022	
<i>CXCL3</i>	GND	-0.00065	0.771	<i>TOMM20</i>	GND	0.00036	0.887
	GD	-0.00260			GD	0.00038	
<i>CYBA</i>	GND	0.00203	0.500	<i>TXNIP</i>	GND	0.06037	0.338
	GD	0.00197			GD	0.04450	
<i>DTNB</i>	GND	0.00001	0.956	<i>WFS1</i>	GND	-0.00034	0.270
	GD	0.00002			GD	-0.00067	

GD = Grupo diabético; GND = Grupo não diabético. * diferenças significativas (p-valor $\leq$ 0.05).

Tabela 7 - Diferenças de medianas de expressão gênica entre os grupos de diabéticos e não diabéticos das formas de *splicing* do *TCF7L2*

FORMAS DE <i>SPLICING</i> <i>TCF7L2</i>	GRUPO	EXPRESSÃO MEDIANA	P-VALOR
A	Grupo não diabético	-0.00219	0.128
	Grupo diabético	-0.00147	
B	Grupo não diabético	-0.00109	0.153
	Grupo diabético	-0.00067	
C	Grupo não diabético	-0.00026	0.060
	Grupo diabético	-0.00012	
D	Grupo não diabético	-0.00342	0.136
	Grupo diabético	-0.00160	
E	Grupo não diabético	-0.00584	0.136
	Grupo diabético	-0.00241	
F	Grupo não diabético	-0.00445	0.293
	Grupo diabético	-0.00388	
G	Grupo não diabético	0.00080	0.153
	Grupo diabético	0.00053	

4.2 Comparações de medianas de expressão e de redes de interações protéicas entre os grupos genotípicos

Primeiramente, buscamos identificar se os genótipos de risco do gene *TCF7L2* estavam afetando a expressão das formas de *splicing* deste gene (influência em *cis* na regulação), e considerando as expressões das sete formas do *TCF7L2* analisadas, observamos que cinco (formas A, B, C, D e E) tinham diferenças de expressão significativas de acordo com o grupo genotípico (P-valores $\leq 0,05$ e $\leq 0,01$) (Tabela 8) (Figura 3). Comparamos também entre os grupos genotípicos os níveis medianos de expressão dos outros 40 genes e observamos que nove (*CDKAL1*, *IGF2BP2*, *JAZF1*, *CDKN2B*, *CAMK1D*, *JUN*, *CDK4*, *ATP2A2*, e *FKBP1A*) apresentavam diferenças significantes estatisticamente (p-valor 0,05), sendo que seis destas diferenças se mantinham utilizando o p-valor 0,01 (*IGF2BP2*, *JAZF1*, *CDK4*, *JUN*, *ATP2A2*, e *FKBP1A*) (Tabela 9).

Através de uma análise inserindo dados de expressão gênica em redes de interação protéica e comparando as redes dos dois grupos genotípicos, notamos que o gene *TCF7L2* (aqui representado pela forma de *splicing* B, uma das cinco que foram observadas como tendo diferença de expressão entre os grupos) e outros 28 genes possuem maiores níveis medianos de

expressão no grupo de risco (sendo 10 com significância estatística, p-valor 0,05) (Figura 4).

Tabela 8 - Diferenças de medianas de expressão gênica entre os grupos selvagem e de risco das formas de *splicing* do *TCF7L2*

FORMAS DE <i>SPLICING</i> <i>TCF7L2</i>	GRUPO	EXPRESSÃO MEDIANA	P-VALOR
A	Grupo Selvagem	-0.00241	0,002*
	Grupo de risco	-0.00134	
B	Grupo Selvagem	-0.00153	0,005*
	Grupo de risco	-0.00048	
C	Grupo Selvagem	-0.00030	0,008*
	Grupo de risco	-0.00012	
D	Grupo Selvagem	-0.00343	0,010*
	Grupo de risco	-0.00059	
E	Grupo Selvagem	-0.00649	0,009*
	Grupo de risco	-0.00253	
F	Grupo Selvagem	-0.00520	0,056
	Grupo de risco	-0.00261	
G	Grupo Selvagem	0.00069	0,522
	Grupo de risco	0.00088	

* Diferenças significativas (p-valor $\leq$ 0.05).

A

B

C

Figura 3 - Diferenças de expressão e co-expressão das formas de *splicing* do gene *TCF7L2* entre os grupos genotípicos. **A.** Comparações dos níveis medianos de expressão das formas do gene *TCF7L2* entre os grupos. **B.** Rede de correlações do grupo selvagem. **C.** Rede de correlações do grupo de risco. * diferenças significativas entre expressões gênicas. Linhas sólidas representam os coeficientes de correlação significativamente diferentes.

Tabela 9 - Diferenças de medianas de expressão gênica entre os grupos selvagem e de risco de 40 genes

GENE	GRUPO	EXPRESSÃO		P-VALOR	GENE	GRUPO	EXPRESSÃO		P-VALOR
		MEDIANA					MEDIANA		
AKT1	Selvagem	0.00003		0.504	EP300	Selvagem	-0.00130		0.323
	Risco	0.00007				Risco	-0.00116		
ATP2A2	Selvagem	-0.11028		0.004*	ERBB4	Selvagem	0.00018		0.743
	Risco	-0.03372				Risco	0.00016		
BAD	Selvagem	0.00326		0.448	FKBP1A	Selvagem	-0.20144		0.003*
	Risco	0.00251				Risco	-0.04673		
BDKRB2	Selvagem	0.00206		0.330	HHEX	Selvagem	0.00006		0.386
	Risco	0.00334				Risco	0.00009		
BMP2	Selvagem	-0.00026		0.558	IGF2BP2	Selvagem	-0.00566		0.005*
	Risco	-0.00032				Risco	-0.00052		
CALM1	Selvagem	0.00159		0.060	IGFBP2	Selvagem	-0.00016		0.667
	Risco	0.00288				Risco	-0.00017		
CALR	Selvagem	0.00443		0.267	IL1R1	Selvagem	-0.00002		0.682
	Risco	0.00644				Risco	-0.00002		
CAMK1D	Selvagem	0.00945		0.024*	JAZF1	Selvagem	-0.00382		0.003*
	Risco	0.01454				Risco	-0.00099		
CDK4	Selvagem	-0.02856		0.004*	JUN	Selvagem	-0.00857		0.003*
	Risco	-0.01385				Risco	-0.00396		
CDKAL1	Selvagem	-0.00076		0.013*	MAP4K4	Selvagem	0.00022		0.617
	Risco	-0.00027				Risco	0.00023		
CDKN1A	Selvagem	0.01716		0.842	NKAP	Selvagem	0.00002		0.221
	Risco	0.01995				Risco	0.00006		
CDKN2A	Selvagem	0.00041		0.833	NRG1	Selvagem	0.00004		0.584
	Risco	0.00043				Risco	0.00005		
CDKN2B	Selvagem	-0.03108		0.017*	PLN	Selvagem	-0.00084		0.117
	Risco	-0.01258				Risco	-0.00066		
CHTF18	Selvagem	0.00035		0.645	PPAR γ	Selvagem	0.02356		0.360
	Risco	0.00058				Risco	0.01792		
CSF3	Selvagem	-0.00050		0.963	RXR α	Selvagem	-0.00096		0.088
	Risco	-0.00085				Risco	-0.00066		
CSRP1	Selvagem	-0.00237		0.194	SPTBN1	Selvagem	0.00385		0.705
	Risco	-0.00222				Risco	0.00414		
CSRP2	Selvagem	0.00023		0.254	TNF α	Selvagem	0.00015		0.545
	Risco	0.00016				Risco	0.00019		
CXCL3	Selvagem	-0.00065		0.981	TOMM20	Selvagem	0.00031		0.848
	Risco	-0.00189				Risco	0.00039		
CYBA	Selvagem	0.00191		0.313	TXNIP	Selvagem	0.06253		0.634
	Risco	0.00351				Risco	0.04585		
DTNB	Selvagem	0.00002		0.836	WFS1	Selvagem	-0.00068		0.059
	Risco	0.00002				Risco	-0.00046		

* Diferenças significativas (p-valor $\leq$ 0.05).

Figura 4 - Redes de interação proteica com medianas de expressão gênica dos grupos selvagem e de risco. **A.** Rede do grupo selvagem. **B.** Rede do grupo de risco. **C.** As cores das proteínas estão de acordo com os níveis medianos de expressão gênica. Vermelho representa maiores valores de expressão gênica e branco menores valores. Cinza representa genes que não tiveram os níveis medianos de expressão gênica medidos.

4.3 Comparações de padrões de co-expressão gênica entre os grupos genotípicos

Primeiramente, analisamos as correlações de expressão entre formas de *splicing* do gene *TCF7L2* e observamos que as formas B e G apresentam respectivamente 4 e 5 diferenças de correlações entre os grupos (Tabela 10). Construindo redes de correlação de expressão para as formas de *splicing* do gene *TCF7L2*, podemos observar melhor que as diferenças em correlações estão principalmente concentradas em duas formas de *splicing* (B e G) (Figura 3).

Tabela 10 - Diferenças significativas de coeficientes de correlação das formas de *splicing* do gene *TCF7L2* entre os grupos selvagem e de risco

		COEFICIENTES DE CORRELAÇÃO		P-VALOR
FORMAS DE <i>SPLICING</i> <i>TCF7L2</i>	FORMAS DE <i>SPLICING</i> <i>TCF7L2</i>	GRUPO SELVAGEM	GRUPO DE RISCO	
B	A	0,89	0,72	0,021
B	C	0,95	0,89	0,029
B	D	0,93	0,84	0,033
B	G	-0,04	0,50	0,006
G	A	-0,04	0,43	0,020
G	C	-0,06	0,40	0,024
G	D	0,06	0,47	0,038
G	E	0,05	0,46	0,041

NOTA: As comparações apresentadas mostram apenas diferenças significativas entre os coeficientes de correlação (P-valor $\leq 0,05$).

Posteriormente, comparamos entre os grupos os coeficientes de correlação de expressão gênica das formas de *splicing* com os outros 40 genes, e observamos que a forma de *splicing* B possui a maior quantidade de coeficientes de correlação de expressão diferentes (nove) (Tabela 11). Nota-se também ao comparar os grupos que os genes *RXRα*, *CSF3*, *CXCL3* e *CALR* possuem muitas diferenças de correlações significantes com as formas de *splicing* (6, 6, 6 e 5, respectivamente, de 7 possíveis) (Tabela 11). Interessantemente, a forma de *splicing* G apresenta diferenças de coeficientes de correlação de expressão com os genes (*ATP2A2*, *CDK4*, *CDKAL1*, *IGF2BP2* e *JAZF1*), enquanto as outras formas de *splicing* não apresentam nenhuma diferença de correlação com genes diferentemente expressos entre os grupos (Tabela 11).

Tabela 11 - Comparações entre os grupos genotípicos de coeficientes de correlação gênica das formas de *splicing* do *TCF7L2* com os demais genes

FORMAS DE SPlicing TCF7L2	GENES	COEFICIENTES DE CORRELAÇÃO		P-VALOR
		GRUPO SELVAGEM	GRUPO DE RISCO	
A	<i>BMP2</i>	-0.32	0.13	0.034
A	<i>CALR</i>	-0.10	0.33	0.038
A	<i>CSF3</i>	-0.51	0.04	0.004
A	<i>CXCL3</i>	-0.47	0.13	0.002
A	<i>IGFBP2</i>	-0.05	0.40	0.029
A	<i>RXRα</i>	0.54	-0.01	0.004
A	<i>TXNIP</i>	0.00	0.49	0.014
B	<i>CALM1</i>	0.00	0.44	0.029
B	<i>CALR</i>	-0.10	0.44	0.009
B	<i>CSF3</i>	-0.50	0.01	0.011
B	<i>CYBA</i>	-0.01	0.44	0.026
B	<i>CXCL3</i>	-0.46	0.04	0.013
B	<i>MAP4K4</i>	-0.09	0.32	0.048
B	<i>IL1R1</i>	0.17	0.64	0.006
B	<i>RXRα</i>	0.54	0.11	0.023
B	<i>TXNIP</i>	0.03	0.45	0.035
C	<i>CALR</i>	-0.14	0.30	0.034
C	<i>CSF3</i>	-0.48	-0.01	0.016
C	<i>CXCL3</i>	-0.46	-0.07	0.048
C	<i>IL1R1</i>	0.14	0.58	0.015
C	<i>RXRα</i>	0.55	0.07	0.010
D	<i>CSF3</i>	-0.42	0.05	0.020
D	<i>CXCL3</i>	-0.38	0.13	0.014
D	<i>IGFBP</i>	0.01	0.44	0.035
D	<i>RXRα</i>	0.56	0.04	0.006
E	<i>CALR</i>	-0.03	0.39	0.045
E	<i>CSF3</i>	-0.46	0.00	0.021
E	<i>CXCL3</i>	-0.42	0.08	0.015
E	<i>RXRα</i>	0.57	0.05	0.006
E	<i>TXNIP</i>	0.08	0.48	0.038
F	<i>CALR</i>	0.06	0.45	0.049
F	<i>CSF3</i>	-0.32	0.09	0.049
F	<i>CYBA</i>	0.08	0.49	0.038
F	<i>CXCL3</i>	-0.27	0.18	0.035
F	<i>RXRα</i>	0.49	0.00	0.013
G	<i>ATP2A2</i>	0.01	0.47	0.017
G	<i>BDKRB2</i>	0.82	0.63	0.048
G	<i>CDK4</i>	-0.05	0.39	0.034
G	<i>CDKAL1</i>	0.08	0.54	0.016
G	<i>IGF2BP2</i>	0.02	0.47	0.023
G	<i>JAZF1</i>	0.04	0.50	0.019

NOTA: As comparações apresentadas mostram apenas diferenças significativas entre os coeficientes de correlação (P-valor $\leq 0,05$).

Comparando os coeficientes de correlação referentes aos 41 genes (820 comparações) entre os dois grupos genotípicos, utilizando a forma de *splicing* B para representar a expressão do gene *TCF7L2*, obtivemos que o gene *RXRα* foi o gene com o maior número de coeficientes de correlação diferentes (onze) (Tabela 12). Na sequência, observamos os genes *CALM1*, *CALR* e *IGF2BP2* com 10 coeficientes de correlação diferentes entre os grupos (Tabela 12). Juntamente com a forma de *splicing* B do gene *TCF7L2*, esses quatro genes listados contabilizam por mais de 50% (43 de 83) das diferenças de correlação detectadas. As demais comparações de correlações estão mostradas na tabela 13.

Analisando as diferenças de correlações significantes usando um limiar de $\pm 0,4$ e $\pm 0,5$ para pelo menos uma das correlações comparadas, observamos que os genes *RXRα*, *CALM1* e *CALR* permanecem como os genes que possuem o maior número de correlações significativamente diferentes (10, 10 e 10 respectivamente para o limiar $\pm 0,4$; e 8, 9 e 9 respectivamente para o limiar $\pm 0,5$) (Tabela 12). Ilustrando as diferenças significativas em padrões de correlação de expressão entre os grupos, usando um limiar de $\pm 0,4$ para pelo menos uma das correlações comparadas, montamos também redes de correlações para os genes *RXRα*, *CALM1* e *CALR* (Figura 5).

Quando comparando os grupos genotípicos quanto a diferenças de correlação entre os 41 genes, foi observado que os genes diferentemente expressos entre os grupos (*IGF2BP2*, forma de *splicing* B do gene *TCF7L2*, *ATP2A2*, *JAZF1*, *CDKAL1*, *CDK4*, *FKBP1A*, *CAMK1D*, *CDKN2B* e *JUN*) participam de aproximadamente 67% (56 de 83) das diferenças de correlações significantes (Tabela 14). Ilustrando as diferenças significativas em padrões de correlações de expressão entre os grupos, usando um limiar de $\pm 0,4$ para pelo menos uma das correlações comparadas, montamos redes de correlação com os genes diferentemente expressos (Figura 6). Nota-se também que os genes *RXRα*, *CALR*, *CSF3* e *CXCL3* (os quais

possuem várias diferenças de coeficientes de correlação com as formas de *splicing* do gene *TCF7L2*), juntamente com o gene *CALM1*, respondem por aproximadamente 55% (31 de 56) das diferenças acima citadas de coeficientes de correlação entre os grupos (7, 7, 5, 5 e 7, respectivamente) (Tabela 14).

Quando observamos as diferenças de coeficientes de correlação entre os grupos utilizando um p-valor menor ou igual a 0,01, observamos que 8 das 17 diferenças estão concentradas nos genes *RXR α* e *CALR* (Tabela 12), e essas diferenças de correlações são com genes diferentemente expressos entre os grupos.

Tabela 12 - Comparações entre os grupos genotípicos de coeficientes de correlação gênica dos genes com maiores números de diferenças de coeficientes de correlação

COEFICIENTES DE CORRELAÇÃO				
GENES	GENES	GRUPO SELVAGEM	GRUPO DE RISCO	P-VALOR
<i>RXRα</i>	<i>AKT1</i>	0.02	0.42	0.048
<i>RXRα</i>	<i>ATP2A2</i>	0.72	0.30	0.005*
<i>RXRα</i>	<i>CDKN2B</i>	0.49	0.08	0.034
<i>RXRα</i>	<i>CSRP1</i>	0.74	0.40	0.017
<i>RXRα</i>	<i>EP300</i>	0.50	0.10	0.037
<i>RXRα</i>	<i>FKBP1A</i>	0.65	0.29	0.026
<i>RXRα</i>	<i>IGF2BP2</i>	0.67	0.17	0.003* #
<i>RXRα</i>	<i>JAZF1</i>	0.66	0.28	0.017
<i>RXRα</i>	<i>JUN</i>	0.52	0.14	0.042
<i>RXRα</i>	<i>PPARγ</i>	0.38	-0.03	0.047
<i>RXRα</i>	<i>TCF7L2 (B)</i>	0.54	0.11	0.023
<i>CALR</i>	<i>ATP2A2</i>	-0.03	0.55	0.002*
<i>CALR</i>	<i>CALM1</i>	0.81	0.92	0.046#
<i>CALR</i>	<i>CDK4</i>	0.02	0.47	0.024
<i>CALR</i>	<i>CDKAL1</i>	0.05	0.58	0.004*
<i>CALR</i>	<i>EP300</i>	0.06	0.51	0.019
<i>CALR</i>	<i>FKBP1A</i>	0.02	0.54	0.006*
<i>CALR</i>	<i>IGF2BP2</i>	-0.02	0.53	0.004* #
<i>CALR</i>	<i>JAZF1</i>	-0.02	0.50	0.009*
<i>CALR</i>	<i>TCF7L2 (B)</i>	-0.10	0.44	0.009*
<i>CALR</i>	<i>TOMM20</i>	0.19	0.56	0.039
<i>CALM1</i>	<i>ATP2A2</i>	0.07	0.55	0.011
<i>CALM1</i>	<i>CDK4</i>	0.13	0.50	0.050
<i>CALM1</i>	<i>CDKAL1</i>	0.23	0.66	0.011
<i>CALM1</i>	<i>EP300</i>	0.25	0.60	0.042
<i>CALM1</i>	<i>FKBP1A</i>	0.17	0.54	0.047
<i>CALM1</i>	<i>IGF2BP2</i>	0.12	0.57	0.013#
<i>CALM1</i>	<i>JAZF1</i>	0.11	0.52	0.032
<i>CALM1</i>	<i>CHTF18</i>	0.67	0.85	0.038
<i>CALM1</i>	<i>TCF7L2 (B)</i>	0.00	0.44	0.029
<i>IGF2BP2</i>	<i>CSF3</i>	-0.44	0.07	0.012
<i>IGF2BP2</i>	<i>CYBA</i>	0.02	0.54	0.006*
<i>IGF2BP2</i>	<i>CXCL3</i>	-0.38	0.11	0.017
<i>IGF2BP2</i>	<i>MAP4K4</i>	0.06	0.45	0.048
<i>IGF2BP2</i>	<i>IL1R1</i>	0.30	0.64	0.034
<i>IGF2BP2</i>	<i>TOMM20</i>	0.02	0.42	0.044
<i>IGF2BP2</i>	<i>TXNIP</i>	0.10	0.57	0.012

NOTA: As comparações apresentadas mostram apenas diferenças significativas entre os coeficientes de correlação ($P\text{-valor} \leq 0,05$). * $P\text{-valor} \leq 0,01$. # correlações que são apontadas apenas uma vez nesta tabela.

Tabela 13 – Demais comparações entre os grupos genotípicos de coeficientes de correlação gênica significativamente diferentes

COEFICIENTES DE CORRELAÇÃO				
GENES	GENES	GRUPO SELVAGEM	GRUPO DE RISCO	P-VALOR
<i>ATP2A2</i>	<i>TOMM20</i>	-0.13	0.29	0.047
<i>ATP2A2</i>	<i>TXNIP</i>	0.03	0.50	0.018
<i>BAD</i>	<i>BMP2</i>	0.40	-0.04	0.031
<i>BAD</i>	<i>CSRP2</i>	0.49	0.02	0.017
<i>BAD</i>	<i>DTNB</i>	0.54	0.17	0.045
<i>BDKRB2</i>	<i>PLN</i>	-0.31	0.10	0.048
<i>BDKRB2</i>	<i>PPARγ</i>	0.45	0.00	0.027
<i>CAMK1D</i>	<i>ERBB4</i>	0.55	0.05	0.008
<i>CAMK1D</i>	<i>PPARγ</i>	0.45	0.05	0.044
<i>CAMK1D</i>	<i>TNFα</i>	0.49	0.11	0.046
<i>CDK4</i>	<i>SPTBN1</i>	0.00	0.40	0.050
<i>CDKAL1</i>	<i>BDKRB2</i>	0.07	0.50	0.026
<i>CDKAL1</i>	<i>CHTF18</i>	0.07	0.51	0.021
<i>CDKAL1</i>	<i>MAP4K4</i>	0.16	0.60	0.013
<i>CSF3</i>	<i>ATP2A2</i>	-0.40	0.02	0.038
<i>CSF3</i>	<i>BAD</i>	0.34	-0.34	0.001
<i>CSF3</i>	<i>CDKN2B</i>	-0.43	-0.01	0.034
<i>CSF3</i>	<i>CXCL3</i>	0.92	0.70	0.001
<i>CSF3</i>	<i>EP300</i>	-0.15	0.30	0.030
<i>CSF3</i>	<i>JAZF1</i>	-0.44	0.01	0.025
<i>CSF3</i>	<i>WFS1</i>	-0.57	0.00	0.002
<i>CXCL3</i>	<i>ATP2A2</i>	-0.35	0.06	0.047
<i>CXCL3</i>	<i>BAD</i>	0.28	-0.18	0.028
<i>CXCL3</i>	<i>BDKRB2</i>	0.50	0.11	0.040
<i>CXCL3</i>	<i>CDK4</i>	-0.34	0.11	0.030
<i>CXCL3</i>	<i>JAZF1</i>	-0.38	0.04	0.044
<i>CXCL3</i>	<i>WFS1</i>	-0.58	-0.01	0.002
<i>CYBA</i>	<i>ATP2A2</i>	0.05	0.48	0.029
<i>CYBA</i>	<i>CDKAL1</i>	-0.01	0.49	0.011
<i>CYBA</i>	<i>DTNB</i>	0.25	0.68	0.008
<i>CYBA</i>	<i>EP300</i>	0.14	0.58	0.018
<i>CYBA</i>	<i>IL1R1</i>	0.18	0.60	0.017
<i>CYBA</i>	<i>JAZF1</i>	0.02	0.45	0.033
<i>CYBA</i>	<i>TOMM20</i>	0.23	0.68	0.005
<i>DTNB</i>	<i>MAP4K4</i>	0.87	0.66	0.010
<i>DTNB</i>	<i>TOMM20</i>	0.13	0.53	0.032
<i>JAZF1</i>	<i>BDKRB2</i>	-0.03	0.41	0.033
<i>JAZF1</i>	<i>IL1R1</i>	0.24	0.61	0.033
<i>TXNIP</i>	<i>CDK4</i>	-0.01	0.48	0.014
<i>TXNIP</i>	<i>FKBP1A</i>	0.04	0.45	0.038

NOTA: As comparações apresentadas mostram apenas diferenças significativas entre os coeficientes de correlação (P-valor $\leq 0,05$). As comparações envolvendo a forma de *splicing* B do gene *TCF7L2* foram colocadas na tabela 11.

Figura 5 - Redes de coexpressão *RXRα*, *CALM1* e *CALR* nas quais as faixas de valores de coeficientes de correlação de Spearman são $\leq -0,4$ e $\geq 0,4$. **A.** Rede do genótipo selvagem. **B.** Rede do genótipo *Δ*. Comparação das redes evidencia padrão de correlação diferente para os genes *RXRα*, *CALR* e *CALM1*, os quais tem muitas interações (linhas, correlações de Spearman acima de 0,4 ou abaixo de -0,4) em apenas uma das redes e poucas interações na outra rede. Linhas sólidas representam diferenças de correlações significativas entre os dois grupos que são visualizadas em apenas uma das redes e linhas tracejadas representam diferenças de correlações significativas que estão em ambas as redes.

Tabela 14 - Comparações entre os grupos genotípicos de coeficientes de correlação gênica com apenas os genes diferentemente expressos

COEFICIENTES DE CORRELAÇÃO				
GENES	GENES	GRUPO SELVAGEM	GRUPO DE RISCO	P-VALOR
<i>IGF2BP2</i>	<i>CALM1</i>	0.12	0.57	0.013
<i>IGF2BP2</i>	<i>CALR</i>	-0.02	0.53	0.004 *
<i>IGF2BP2</i>	<i>CSF3</i>	-0.44	0.07	0.012
<i>IGF2BP2</i>	<i>CXCL3</i>	-0.38	0.11	0.017
<i>IGF2BP2</i>	<i>CYBA</i>	0.02	0.54	0.006 *
<i>IGF2BP2</i>	<i>IL1R1</i>	0.30	0.64	0.034
<i>IGF2BP2</i>	<i>MAP4K4</i>	0.06	0.45	0.048
<i>IGF2BP2</i>	<i>RXRα</i>	0.67	0.17	0.003 *
<i>IGF2BP2</i>	<i>TOMM20</i>	0.02	0.42	0.044
<i>IGF2BP2</i>	<i>TXNIP</i>	0.10	0.57	0.012
<i>TCF7L2 (B)</i>	<i>CALM1</i>	0.00	0.44	0.029
<i>TCF7L2 (B)</i>	<i>CALR</i>	-0.10	0.44	0.009 *
<i>TCF7L2 (B)</i>	<i>CSF3</i>	-0.50	0.01	0.011
<i>TCF7L2 (B)</i>	<i>CXCL3</i>	-0.46	0.04	0.013
<i>TCF7L2 (B)</i>	<i>CYBA</i>	-0.01	0.44	0.026
<i>TCF7L2 (B)</i>	<i>IL1R1</i>	0.17	0.64	0.006 *
<i>TCF7L2 (B)</i>	<i>MAP4K4</i>	-0.09	0.32	0.048
<i>TCF7L2 (B)</i>	<i>RXRα</i>	0.54	0.11	0.023
<i>TCF7L2 (B)</i>	<i>TXNIP</i>	0.03	0.45	0.035
<i>ATP2A2</i>	<i>CALM1</i>	0.07	0.55	0.011
<i>ATP2A2</i>	<i>CALR</i>	-0.03	0.55	0.002 *
<i>ATP2A2</i>	<i>CSF3</i>	-0.40	0.02	0.038
<i>ATP2A2</i>	<i>CXCL3</i>	-0.35	0.06	0.047
<i>ATP2A2</i>	<i>CYBA</i>	0.05	0.48	0.029
<i>ATP2A2</i>	<i>RXRα</i>	0.72	0.30	0.005 *
<i>ATP2A2</i>	<i>TOMM20</i>	-0.13	0.29	0.047
<i>ATP2A2</i>	<i>TXNIP</i>	0.03	0.50	0.018
<i>JAZF1</i>	<i>BDKRB2</i>	-0.03	0.41	0.033
<i>JAZF1</i>	<i>CALM1</i>	0.11	0.52	0.032
<i>JAZF1</i>	<i>CALR</i>	-0.02	0.50	0.009 *
<i>JAZF1</i>	<i>CSF3</i>	-0.44	0.01	0.025
<i>JAZF1</i>	<i>CXCL3</i>	-0.38	0.04	0.044
<i>JAZF1</i>	<i>CYBA</i>	0.02	0.45	0.033
<i>JAZF1</i>	<i>IL1R1</i>	0.24	0.61	0.033
<i>JAZF1</i>	<i>RXRα</i>	0.66	0.28	0.017

Continua

Conclusão

Tabela 14

COEFICIENTES DE CORRELAÇÃO				
GENES	GENES	GRUPO SELVAGEM	GRUPO DE RISCO	P-VALOR
<i>CDKAL1</i>	<i>BDKRB2</i>	0.07	0.50	0.026
<i>CDKAL1</i>	<i>CALM1</i>	0.23	0.66	0.011
<i>CDKAL1</i>	<i>CALR</i>	0.05	0.58	0.004 *
<i>CDKAL1</i>	<i>CHTF18</i>	0.07	0.51	0.021
<i>CDKAL1</i>	<i>CYBA</i>	-0.01	0.49	0.011
<i>CDKAL1</i>	<i>MAP4K4</i>	0.16	0.60	0.013
<i>CDK4</i>	<i>CALM1</i>	0.13	0.50	0.050
<i>CDK4</i>	<i>CALR</i>	0.02	0.47	0.024
<i>CDK4</i>	<i>CXCL3</i>	-0.34	0.11	0.030
<i>CDK4</i>	<i>SPTBN1</i>	0.00	0.40	0.050
<i>CDK4</i>	<i>TXNIP</i>	-0.01	0.48	0.014
<i>FKBP1A</i>	<i>CALM1</i>	0.17	0.54	0.047
<i>FKBP1A</i>	<i>CALR</i>	0.02	0.54	0.006 *
<i>FKBP1A</i>	<i>RXRα</i>	0.65	0.29	0.026
<i>FKBP1A</i>	<i>TXNIP</i>	0.04	0.45	0.038
<i>CAMK1D</i>	<i>ERBB4</i>	0.55	0.05	0.008 *
<i>CAMK1D</i>	<i>PPARγ</i>	0.45	0.05	0.044
<i>CAMK1D</i>	<i>TNFα</i>	0.49	0.11	0.046
<i>CDKN2B</i>	<i>CSF3</i>	-0.43	-0.01	0.034
<i>CDKN2B</i>	<i>RXRα</i>	0.49	0.08	0.034
<i>JUN</i>	<i>RXRα</i>	0.52	0.14	0.042

NOTA: As comparações apresentadas mostram apenas diferenças significativas entre os coeficientes de correlação ($P\text{-valor} \leq 0,05$). * $P\text{-valor} \leq 0,01$.

Figura 6. Redes de coexpressão com os genes diferentemente expressos nas quais as faixas de valores de coeficientes de correlação de Spearman são $\leq -0,4$ e $\geq 0,4$. **A.** Rede do genótipo selvagem. **B.** Rede do genótipo de risco. Comparação das redes de evidência padrão de correlação diferente para os genes diferentemente expressos entre os grupos genotípicos, genes os quais tem muitas interações (linhas, correlações de Spearman acima de 0,4 ou abaixo de -0,4) em apenas uma das redes e poucas interações na outra rede. Linhas sólidas representam diferenças de correlações significativas entre os dois grupos.

5. DISCUSSÃO

Sabe-se que diabetes tipo 2 é um importante fator de risco para aterosclerose, sendo associado com um aumento de incidências de doenças cardiovasculares, e podendo ocasionar um risco aumentado de duas a três vezes de doença arterial coronariana^{21,95,96}. Apesar de hoje ser conhecida a associação entre o polimorfismo rs7903146 do gene *TCF7L2* com o risco de DM2 e de DAC^{21,46}, pouco se sabe sobre o modo molecular de ação do genótipo nessas doenças.

SNPs localizados em genes associados a doenças podem estar de algum modo regulando a expressão dos mesmos (regulação em *cis*). Com relação ao DM2, vários estudos primeiramente atentaram em observar o efeito do polimorfismo rs7903146 em tecidos reguladores do metabolismo glicolítico. *Lyssenko et al*⁴⁸ demonstraram que variantes genéticas no gene *TCF7L2* poderiam estar influenciando a expressão desse gene em ilhotas pancreáticas, porém esse experimento não foi replicado por *Cotsapas et al*⁵⁵, que não obtiveram os mesmos resultados analisando tecidos hepáticos, intestinais e pancreáticos. Estudos recentes^{97,98} descreveram o alelo de risco estudado como sendo responsável por abertura de cromatina e atividade gênica aumentada, o que, deste modo, levaria a uma super-expressão do gene *TCF7L2* em indivíduos que possuem o alelo de risco. Baseados nesses resultados, *Savic et al*⁴⁷, analisaram tecidos relacionados à homeostase glicolítica; e, através de deleção seletiva do intervalo genômico que contém o rs7903146 e outros polimorfismos, demonstraram que essa região é crítica na modulação da expressão do *TCF7L2*. Neste mesmo estudo foi observado que a super-expressão do gene pode estar associada com o aumento do risco de DM2 através de uma diminuição de tolerância a glicose.

Com relação a doenças cardiovasculares, sabe-se que o gene *TCF7L2* codifica o fator de transcrição TCF-4, que se liga dentro do núcleo celular com o produto do gene β -catenina (*CTNNB1*), e, desse modo, participa da transmissão do sinal da via de sinalização regular Wnt². Essa via de

sinalização é responsável por alguns processos como proliferação, migração e apoptose em diversos tipos celulares. Tais processos alterados são características de CMLs durante a aterosclerose, e, caso a via Wnt esteja sinalizando em CMLs durante a aterosclerose, apresentaria-se como uma provável mediadora de processos celulares associados com as doenças cardiovasculares². De fato, foi mostrado que o complexo transcripcional β -catenina e TCF-4 pode possuir um importante papel no remodelamento vascular através da regulação da proliferação de células de músculo liso^{23,87}.

Alguns estudos inicialmente também tentaram estabelecer uma ligação entre a desregulação do equilíbrio glicolítico e a proliferação de células vasculares de músculo liso. Suzuki *et al*⁵ mostraram que o diabetes acelera a proliferação e acúmulo de CMLs em lesões ateroscleróticas, porém, utilizando CMLs de porcos e humanos cultivadas não obtiveram resultados quanto ao efeito de glicose como estímulo da proliferação celular. Já outros estudos apontaram que a proliferação e migração de células de músculo liso estavam aumentados na presença de hiperglicemia^{99,100}. Contudo, os mediadores moleculares que respondem a essa condição e ocasionam a proliferação de CMLs ainda tem de ser determinados. Muendlein *et al*²¹ hipotetizaram que variantes do gene *TCF7L2* poderiam influenciar o desenvolvimento de aterosclerose diretamente regulando a proliferação de células vasculares de músculo liso ou ativando a via NFK- β por intermédio da via de sinalização Wnt^{101,102}, ou indiretamente através do aumento de níveis de glicose sangüínea ocasionado pelo funcionamento inadequado das células β . Neste ultimo caso, o aumento de glicose sangüínea seria responsável pela ativação do gene *NFK β* ^{21,103,104}, o qual, dentro de sua via, foi proposto como regulador da expressão de diversos genes inflamatórios que são associados à aterosclerose^{20,21}. Porém esta última hipótese não explica os achados do nosso grupo de pesquisa, que demonstram uma associação do rs7903146 com uma incidência aumentada de doença arterial coronariana, assim como com eventos cardiovasculares em indivíduos não diabéticos⁴⁶.

Nesse presente estudo nós utilizamos células vasculares de músculo liso humano, as quais, portanto, são muito relevantes em doenças cardiovasculares¹⁰⁵, sendo também, de acordo com nosso conhecimento atual, um tipo de célula ainda não estudado em relação aos efeitos de regulação em *cis* de polimorfismos do gene *TCF7L2*. O principal objetivo do uso desse tipo celular foi entender melhor a associação, citada previamente, encontrada por nosso grupo de pesquisa⁴⁶. E para tanto nós obtivemos 92 amostras, o que vem a ser um número interessante considerando-se a dificuldade em se obter tal tecido. Além do mais, através do processo de cultura celular, acreditamos na possibilidade de ter reduzido alguns efeitos epigenéticos (como idade, sexo e medicações) nas medições de expressão gênica. Essa redução dos efeitos foi assumida e pode ser mais bem visualizada com base nas tabelas 6 e 7, que comparam os indivíduos diabéticos e não diabéticos.

Deste modo, com a finalidade de se entender o efeito do polimorfismo rs7903146 na patologia do DM2 e da DAC, nós utilizamos duas estratégias (comparação de níveis medianos de expressão de genes relacionados às doenças utilizando redes de interação protéica; e comparação de padrões de co-expressão dos mesmos genes) para comparar fenotipicamente células de indivíduos que não possuem e possuem o determinado alelo de risco (grupo selvagem e grupo de risco, respectivamente). De modo geral, os resultados de ambas as estratégias sugerem a existência de diferenças fenotípicas entre os grupos.

A primeira estratégia de comparação fenotípica de células dos dois grupos genotípicos, teve por finalidade encontrar o efeito do polimorfismo rs7903146 na expressão do gene *TCF7L2*, utilizando uma abordagem de eQTL^{29,106}. Em nosso estudo observamos com significância estatística uma maior expressão de cinco grupos de transcritos (formas de *splicing*) do gene *TCF7L2* (A, B, C, D e E) no grupo de risco se comparado com o grupo selvagem (Figura 3). Observamos também que é possível que o genótipo esteja influenciando no balanço da produção dos transcritos do *TCF7L2* e que embora a expressão das formas do ensaio G não esteja alterada entre

os grupos genotípicos (Figura 3), os relacionamentos dessas formas com vários genes diferentemente expressos possam estar alterados (Tabela 11). Em humanos, esse efeito de genótipos do *TCF7L2* na regulação em *cis* já foi mostrado por outros pesquisadores em células pancreáticas, adiposas e do sangue^{84,85,107}, porém, de nosso conhecimento, o presente estudo apresenta a primeira demonstração desses efeitos em CMLs humanas.

Devido ao grande número de transcritos descritos do *TCF7L2*¹⁰⁸ e a grande similaridade dos mesmos, nós, assim como outros pesquisadores^{84,85,109}, encontramos dificuldades em conduzir ensaios transcrito-específicos de RT-PCR em tempo real. Um dos possíveis transcritos do *TCF7L2* que pode ter sido detectado em nossos ensaios A e B foi o HM353847, o qual possui o *exon* 15 e foi descrito como produzindo notáveis características de transformação celular maligna em linhagens celulares do fígado, caracterizadas por altas taxas de proliferação celular, migração e formação de colônia¹⁰⁹. Observamos também que o ensaio B foi o único que selecionou formas de *splicing* que possuem o *exon* 9 longo. A parte da proteína codificada pelo *exon* 9, na ausência de CTNNB1 no núcleo, permanece ligada ao correpressor TLE/GROUCHO¹¹⁰. Assim, uma hipótese é que o correpressor TLE/GROUCHO, em humanos, possa interagir diferentemente com proteínas produzidas por transcritos com o *exon* 9 curto e com o *exon* 9 longo, e, deste modo, não desempenhar de modo satisfatório sua função de repressão em proteínas de transcritos com o *exon* 9 longo. Notou-se também que os ensaios A, B e C abrangem formas de *splicing* que não possuem o *exon* 17 como codificador de proteína, sendo que a parte da proteína codificada por tal *exon* é um sítio de ligação do correpressor C-terminal binding protein 1 (CtBP-1)¹¹⁰. Deste modo, é possível que a CtBP-1 não se ligando a proteínas de formas sem o *exon* 17 codificador estaria deixando de reprimir a sinalização do gene *TCF7L2*. No futuro, mostram-se necessários esforços para desenvolver ensaios que sejam capazes de determinar efeitos dos genótipos do *TCF7L2* nos níveis de transcritos específicos, assim como será necessário o melhor

entendimento da relação de formas da proteína do *TCF7L2* com seus repressores.

Posteriormente, comparamos entre os grupos genotípicos as expressões de outros genes que possuem ou estão perto de SNPs associados ao DM2, de genes relacionados a vias de sinalização de DM2 e de genes relacionados à proliferação celular; a fim de investigar se, hipoteticamente, as expressões desses genes poderiam estar respondendo ao sinal de mudanças nas expressões de formas do *TCF7L2* (Figura 4). Os genes *CDKAL1*, *IGF2BP2*, *JAZF1*, *CDKN2B*, *CAMK1D*, *JUN*, *CDK4*, *ATP2A2*, e *FKBP1A* foram observados como tendo diferenças de níveis medianos de expressão entre os grupos, um número significativamente mais alto do que seria esperado ao acaso. É interessante notar que entre os nove genes com diferença de expressão, apresentam-se os genes *IGF2BP2*, *JAZF1*, *CDKN2B*, *CAMK1D*, e *CDKAL1*, os quais são genes que estão perto ou possuem SNPs associados ao DM2^{33,36,39,40,111}. Por essa relação, esses SNPs têm sido amplamente estudados em tecidos relacionados à homeostase glicolítica e são propostos como sendo variantes que afetam a função de células β pancreáticas⁴². Observa-se também que os nove genes com diferença de expressão entre os grupos já foram relacionados a eventos de proliferação celular em alguns estudos^{36,112-118}. Tais observações sugerem que uma desregulação na homeostase glicolítica e/ou uma desregulação no ciclo celular possam ser mecanismos patogênicos comuns do DM2.

Quando analisamos em conjunto os dados de expressão inseridos em uma plataforma que agrega dados de interação protéica, observamos que 29 dos 41 genes medidos possuíam um nível mediano nominal de expressão maior no grupo de risco quando comparado com o grupo selvagem, deste modo, essa pode ser uma evidência de que mudanças em expressão gênica podem parecer mais obvias quando analisadas em conjunto dentro de um determinado sistema¹¹⁹ e que essas mudanças em conjunto estejam sendo ocasionadas pela mudança de expressão de alguma isoforma do *TCF7L2*.

A segunda estratégia aplicada teve por fim inferir se os genes estudados interagem diferentemente de acordo com o grupo genotípico. Com esse propósito, comparamos correlações de expressão gênica entre os grupos. Deste modo, procuramos por genes com padrões de correlações altamente alterados entre os grupos, comportamento que pode evidenciar genes que possuem um papel crucial na patologia de doenças^{68,120,121}, e nesse caso especificamente, na modulação genética do risco de DM2 e DAC pelo gene *TCF7L2*.

Pela abordagem de comparação de co-expressão encontramos os genes *RXR α* , *CALM1*, *CALR*, e *IGF2BP2*, os quais mostraram ter as estruturas de correlação mais significativamente diferentes entre os grupos selvagem e de risco; e os genes *RXR α* , *CALR*, *CALM1*, *CXCL3* e *CSF3*, os quais, de acordo com o genótipo, mudam suas estruturas de correlação com as formas de *splicing* dos ensaios do *TCF7L2* e/ou com os genes diferentemente expressos entre os grupos.

Durante os últimos anos o gene *RXR α* tem sido extensamente estudado em dois campos relacionados à patogênese de doenças cardiovasculares: proliferação celular e homeostase de glicose¹²²⁻¹²⁴. O papel de retinóides no desenvolvimento de doenças cardiovasculares atraiu a atenção dos pesquisadores pela habilidade de todos os ácidos trans-retinoicos de suprimir o crescimento de células vasculares de músculo liso^{124,125}. Outro estudo revelou a importância do tratamento com ácido retinóico, o qual pode induzir sinais de inibição de crescimento em células de músculo liso pulmonar¹²⁶. Foi também descrito que o mecanismo de fosforilação/degradação/transativação alterado do *RXR α* caracteriza proliferação anormal, embora benigna, de células de músculo liso uterinas¹²⁷.

Com relação à homeostase de glicose, o heterodímero *RXR α /PPAR γ* é um conhecido regulador da homeostase de lipídios e de glicose¹²³, os quais são importantes mecanismos desregulados na patologia do DM2. O gene *PPAR γ* produz uma proteína nuclear que está envolvida na regulação transcricional de uma variedade de enzimas relacionadas ao metabolismo

de lipídios e glicose¹²⁸, além de ser descrito como possuindo SNPs associados ao risco de DM2^{30,36}, e de ser um importante alvo para intervenções de doenças metabólicas^{129,130}. O *PPAR γ* é um conhecido alvo para drogas sensibilizadoras de insulina, como as tiazolidinedionas (TZD), as quais são usadas para tratar DM2. Proteínas co-ativadoras e co-repressoras podem regular a atividade do *PPAR γ* modulando a expressão gênica mediada por esse gene¹³¹⁻¹³³. Li *et al*¹³⁴ demonstraram, através da deleção do *NCoR*, co-repressor do *PPAR γ* , que a ativação do programa transcricional mediado pelo *PPAR γ* em adipócitos é o principal responsável pela sensibilidade a insulina.

CALM1 e *CALR* são genes conhecidos que participam de vias de sinalização de cálcio. *CALM1* codifica CaM, a qual se liga ao Ca^{2+} intracelular e faz a mediação de vários sinais envolvidos em funções celulares, incluindo proliferação, mobilidade, progressão do ciclo celular, e transcrição¹³⁵. Mais além, o aumento de sinais de concentração intracelular de cálcio mediado por vasoconstritores tem sido amplamente descritos em células vasculares de músculo liso como um fator que contribui para hiperreatividade a vasoconstritores¹³⁶⁻¹³⁸, característica que é relacionada com DM2^{136,139}. O aumento do influxo de íons cálcio parece também estar envolvido no mecanismo de aumento de proliferação de CMLs cultivadas¹³⁸.

Interessantemente, outro gene que foi observado como possuindo um padrão diferente de expressão e co-expressão entre os grupos, *IGF2BP2*, além de ser um gene relacionado ao mecanismo de disfunção de células-beta⁴², também foi apontado como possuindo um papel em proliferação celular e em câncer¹⁴⁰. Em CMLs, a proteína IGF-2, a qual tem sua transcrição regulada pelo gene *IGF2BP2*, já foi descrita como estimulando significativamente a proliferação de células cultivadas¹¹⁶.

O gene *CXCL3* faz parte de uma família de proteínas de ligação de heparina que exibem papéis únicos e importantes na regulação da angiogênese. As proteínas desse gene contêm o motivo de ligação ELR e desse modo se mostram como potentes promotores da angiogênese,

mediando a atividade angiogênica através da ligação e ativação do gene CXCR2 no endotélio¹⁴¹.

O gene *CSF3* já foi descrito como participando da via PI3K-AKT, a qual tem grande importância relativa à absorção de glicose celular⁸⁶. Além disso, esse gene pode também ativar a via de sinalização Wnt para manter junções “gap” através do recrutamento de β -catenina e caderina^{142,143}, que são importantes reguladores do comportamento de células de músculo liso vasculares¹⁴⁴. A expressão desse gene já foi também descrita como possuindo uma função mediadora da proliferação de células de músculo liso vascular¹⁴⁵.

Deste modo, nossa análise abre a possibilidade de que os resultados observados em humanos poderiam não apenas serem derivados de mecanismos sistêmicos de homeostase de glicose¹⁴⁶, mas também de um desarranjo na homeostase de cálcio e de uma proliferação celular alterada, em um sistema onde o *TCF7L2* é central.

Nós reconhecemos as limitações das abordagens utilizadas nesse estudo. Tentamos identificar e quantificar a maioria das isoformas de transcritos do gene *TCF7L2*, entretanto, não conseguimos identificar transcritos específicos do gene, os quais podem contribuir de fato para as doenças, e acabamos por identificar apenas agrupamentos de isoformas. Os p-valores para as associações entre os genótipos de risco e as expressões dos genes não foram ajustados para múltiplos testes, de modo que entendemos que nosso poder estatístico obtido para tais análises não foi alto por possuímos um número de amostras, embora relativamente grande se comparado com os números descritos na literatura, insuficiente. Desse modo o objetivo desse trabalho foi de identificar potenciais diferenças importantes biologicamente que pudessem ser testadas em outros conjuntos de amostras e com métodos (não funcionais e funcionais) adicionais.

6. CONCLUSÃO

Acreditamos que esse trabalho¹⁴⁷ contribui significativamente para o entendimento de que alelos de risco de DM2 podem estar influenciando a expressão de transcritos do gene *TCF7L2* em células vasculares de músculo liso; e que diferenças de expressão de transcritos do *TCF7L2* podem possivelmente causar mudanças significativas nas expressões de outros genes e/ou em interações gênicas, deste modo afetando sistemas multidimensionais como redes gênicas responsáveis pela homeostase glicolítica, por vias de sinalização de cálcio e por proliferação celular. Juntas, nossas estratégias podem ajudar a entender diferenças fenotípicas entre grupos e também proporcionam um melhor conhecimento de como genes podem se relacionar de acordo com o grupo genotípico. Além do mais, alguns genes como *RXRα*, *CALM1*, *CALR*, *IGF2BP2*, *CSF3* e *CXCL3* são apontados como importantes candidatos para serem estudados funcionalmente através de modulação farmacológica, sendo também possíveis candidatos-alvo para tratamento de redução de risco de DM2 e de DAC em indivíduos com alto risco de desenvolver essas doenças, especialmente indivíduos que possuem os genótipos de risco do *TCF7L2*.

7. REFERENCIAS BIBLIOGRÁFICAS

1. AvailableFrom:<http://www.who.int/en/> O. 2010.
2. Mill C, George SJ: Wnt signalling in smooth muscle cells and its role in cardiovascular disorders. *Cardiovasc Res* 2012; **95**: 233-240.
3. A Joint Editorial Statement by the American Diabetes Association; the National Heart L, and Blood Institute; the Juvenile Diabetes Foundation International; the National Institute of Diabetes and Digestive and Kidney Diseases; and the American Heart Association Diabetes mellitus: a major risk factor for cardiovascular disease. . *Circulation* 1999; **100**: 1132-1133.
4. Norhammar A, Malmberg K, Diderholm E *et al*: Diabetes mellitus: the major risk factor in unstable coronary artery disease even after consideration of the extent of coronary artery disease and benefits of revascularization. *J Am Coll Cardiol* 2004; **43**: 585-591.
5. Suzuki LA, Poot M, Gerrity RG, Bornfeldt KE: Diabetes accelerates smooth muscle accumulation in lesions of atherosclerosis: lack of direct growth-promoting effects of high glucose levels. *Diabetes* 2001; **50**: 851-860.
6. de Maria CAB, Moreira RFA, Marcílio R: *Bioquímica do Diabetes Melito*: Ed. INTERCIENCIA, 2011.
7. Lyssenko V, Jonsson A, Almgren P *et al*: Clinical risk factors, DNA variants, and the development of type 2 diabetes. *N Engl J Med* 2008; **359**: 2220-2232.
8. Butler AE, Janson J, Bonner-Weir S, Ritzel R, Rizza RA, Butler PC: Beta-cell deficit and increased beta-cell apoptosis in humans with type 2 diabetes. *Diabetes* 2003; **52**: 102-110.
9. Lyssenko V, Almgren P, Anevski D *et al*: Predictors of and longitudinal changes in insulin sensitivity and secretion preceding onset of type 2 diabetes. *Diabetes* 2005; **54**: 166-174.
10. Tirosh A, Shai I, Tekes-Manova D *et al*: Normal fasting plasma glucose levels and type 2 diabetes in young men. *N Engl J Med* 2005; **353**: 1454-1462.
11. Stern MP, Williams K, Haffner SM: Identification of persons at high risk for type 2 diabetes mellitus: do we need the oral glucose tolerance test? *Ann Intern Med* 2002; **136**: 575-581.
12. Wilson PW, Meigs JB, Sullivan L, Fox CS, Nathan DM, D'Agostino RB, Sr.: Prediction of incident diabetes mellitus in middle-aged adults: the Framingham Offspring Study. *Arch Intern Med* 2007; **167**: 1068-1074.
13. Xu H, Barnes GT, Yang Q *et al*: Chronic inflammation in fat plays a crucial role in the development of obesity-related insulin resistance. *J Clin Invest* 2003; **112**: 1821-1830.
14. Dor Y, Brown J, Martinez OI, Melton DA: Adult pancreatic beta-cells are formed by self-duplication rather than stem-cell differentiation. *Nature* 2004; **429**: 41-46.
15. Keller MP, Choi Y, Wang P *et al*: A gene expression network model of type 2 diabetes links cell cycle regulation in islets with diabetes susceptibility. *Genome Res* 2008; **18**: 706-716.
16. Scheen AJ, Paquot N, Lefebvre PJ: [Glucotoxicity and lipotoxicity, two implicated accomplices in the vicious circle of type 2 diabetes]. *Rev Med Liege* 1999; **54**: 535-538.
17. Roden M: Muscle triglycerides and mitochondrial function: possible mechanisms for the development of type 2 diabetes. *Int J Obes (Lond)* 2005; **29 Suppl 2**: S111-115.
18. Kim JK, Wi JK, Youn JH: Metabolic impairment precedes insulin resistance in skeletal muscle during high-fat feeding in rats. *Diabetes* 1996; **45**: 651-658.

19. Hoy AJ, Bruce CR, Cederberg A *et al*: Glucose infusion causes insulin resistance in skeletal muscle of rats without changes in Akt and AS160 phosphorylation. *Am J Physiol Endocrinol Metab* 2007; **293**: E1358-1364.
20. Mazzone T, Chait A, Plutzky J: Cardiovascular disease risk in type 2 diabetes mellitus: insights from mechanistic studies. *Lancet* 2008; **371**: 1800-1809.
21. Muendlein A, Saely CH, Geller-Rhomberg S *et al*: Single nucleotide polymorphisms of TCF7L2 are linked to diabetic coronary atherosclerosis. *PLoS One* 2011; **6**: e17978.
22. Zhu L, Sun G, Zhang H *et al*: PGC-1alpha is a key regulator of glucose-induced proliferation and migration in vascular smooth muscle cells. *PLoS One* 2009; **4**: e4182.
23. Quaschnick H, Slater SC, Beeching CA, Boehm M, Sala-Newby GB, George SJ: Regulation of smooth muscle cell proliferation by beta-catenin/T-cell factor signaling involves modulation of cyclin D1 and p21 expression. *Circ Res* 2006; **99**: 1329-1337.
24. Vaag A, Poulsen P: Twins in metabolic and diabetes research: what do they tell us? *Curr Opin Clin Nutr Metab Care* 2007; **10**: 591-596.
25. Newman B, Selby JV, King MC, Slemenda C, Fabsitz R, Friedman GD: Concordance for type 2 (non-insulin-dependent) diabetes mellitus in male twins. *Diabetologia* 1987; **30**: 763-768.
26. Kaprio J, Tuomilehto J, Koskenvuo M *et al*: Concordance for type 1 (insulin-dependent) and type 2 (non-insulin-dependent) diabetes mellitus in a population-based cohort of twins in Finland. *Diabetologia* 1992; **35**: 1060-1067.
27. Poulsen P, Kyvik KO, Vaag A, Beck-Nielsen H: Heritability of type II (non-insulin-dependent) diabetes mellitus and abnormal glucose tolerance--a population-based twin study. *Diabetologia* 1999; **42**: 139-145.
28. Kobberling J TH: Empirical risk figures for first degree relatives of non-insulin dependent diabetics: *The genetics of diabetes mellitus*. London: Academic Press, 1982, pp 201 - 209.
29. Cookson W, Liang L, Abecasis G, Moffatt M, Lathrop M: Mapping complex disease traits with global gene expression. *Nat Rev Genet* 2009; **10**: 184-194.
30. Altshuler D, Hirschhorn JN, Klannemark M *et al*: The common PPARgamma Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. *Nat Genet* 2000; **26**: 76-80.
31. Gloyn AL, Weedon MN, Owen KR *et al*: Large-scale association studies of variants in genes encoding the pancreatic beta-cell KATP channel subunits Kir6.2 (KCNJ11) and SUR1 (ABCC8) confirm that the KCNJ11 E23K variant is associated with type 2 diabetes. *Diabetes* 2003; **52**: 568-572.
32. Grant SF, Thorleifsson G, Reynisdottir I *et al*: Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat Genet* 2006; **38**: 320-323.
33. Saxena R, Voight BF, Lyssenko V *et al*: Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* 2007; **316**: 1331-1336.
34. Frayling TM, Timpson NJ, Weedon MN *et al*: A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* 2007; **316**: 889-894.
35. Sandhu MS, Weedon MN, Fawcett KA *et al*: Common variants in WFS1 confer risk of type 2 diabetes. *Nat Genet* 2007; **39**: 951-953.
36. Scott LJ, Mohlke KL, Bonnycastle LL *et al*: A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* 2007; **316**: 1341-1345.
37. Sladek R, Rocheleau G, Rung J *et al*: A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* 2007; **445**: 881-885.

38. Steinthorsdottir V, Thorleifsson G, Reynisdottir I *et al*: A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. *Nat Genet* 2007; **39**: 770-775.
39. Zeggini E, Weedon MN, Lindgren CM *et al*: Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science* 2007; **316**: 1336-1341.
40. Zeggini E, Scott LJ, Saxena R *et al*: Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat Genet* 2008; **40**: 638-645.
41. Voight BF, Scott LJ, Steinthorsdottir V *et al*: Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat Genet* 2010; **42**: 579-589.
42. Grarup N, Sparso T, Hansen T: Physiologic characterization of type 2 diabetes-related loci. *Curr Diab Rep* 2010; **10**: 485-497.
43. Waters KM, Stram DO, Hassanein MT *et al*: Consistent association of type 2 diabetes risk variants found in europeans in diverse racial and ethnic groups. *PLoS Genet* 2010; **6**.
44. Cauchi S, El Achhab Y, Choquet H *et al*: TCF7L2 is reproducibly associated with type 2 diabetes in various ethnic groups: a global meta-analysis. *J Mol Med (Berl)* 2007; **85**: 777-782.
45. Marqueline GF, Pereira AC, Sousa AG, Mill JG, Hueb WA, Krieger JE: TCF7L2 variant genotypes and type 2 diabetes risk in Brazil: significant association, but not a significant tool for risk stratification in the general population. *BMC Med Genet* 2008; **9**: 106.
46. Sousa AG, Marqueline GF, Lemos PA *et al*: TCF7L2 polymorphism rs7903146 is associated with coronary artery disease severity and mortality. *PLoS One* 2009; **4**: e7697.
47. Savic D, Ye H, Aneas I, Park SY, Bell GI, Nobrega MA: Alterations in TCF7L2 expression define its role as a key regulator of glucose metabolism. *Genome Res* 2011; **21**: 1417-1425.
48. Lyssenko V, Lupi R, Marchetti P *et al*: Mechanisms by which common variants in the TCF7L2 gene increase risk of type 2 diabetes. *J Clin Invest* 2007; **117**: 2155-2163.
49. Jin T, Liu L: The Wnt signaling pathway effector TCF7L2 and type 2 diabetes mellitus. *Mol Endocrinol* 2008; **22**: 2383-2392.
50. AvailableFrom:http://www.ncbi.nlm.nih.gov/SNP/snp_summary.cgi dSGVNCfBINdSi. 2010 Jan 28. .
51. Redon R, Ishikawa S, Fitch KR *et al*: Global variation in copy number in the human genome. *Nature* 2006; **444**: 444-454.
52. Burton PR, Clayton DG, Cardon LR *et al*: Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. *Nat Genet* 2007; **39**: 1329-1337.
53. Cheung VG, Spielman RS: Genetics of human gene expression: mapping DNA variants that influence gene expression. *Nat Rev Genet* 2009; **10**: 595-604.
54. Whitney AR, Diehn M, Popper SJ *et al*: Individuality and variation in gene expression patterns in human blood. *Proc Natl Acad Sci U S A* 2003; **100**: 1896-1901.
55. Cotsapas C, Prokunina-Olsson L, Welch C *et al*: Expression analysis of loci associated with type 2 diabetes in human tissues. *Diabetologia* 2010; **53**: 2334-2339.
56. Cauchi S, Meyre D, Dina C *et al*: Transcription factor TCF7L2 genetic study in the French population: expression in human beta-cells and adipose tissue and strong association with type 2 diabetes. *Diabetes* 2006; **55**: 2903-2908.

57. Vandesompele J, De Preter K, Pattyn F *et al*: Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 2002; **3**: RESEARCH0034.
58. Engwegen JYMN: Clinical proteomics in colorectal and renal cell cancer. 2008.
59. Wong ML, Medrano JF: Real-time PCR for mRNA quantitation. *Biotechniques* 2005; **39**: 75-85.
60. Cui X, Churchill GA: Statistical tests for differential expression in cDNA microarray experiments. *Genome Biol* 2003; **4**: 210.
61. Schmittgen TD, Lee EJ, Jiang J *et al*: Real-time PCR quantification of precursor and mature microRNA. *Methods* 2008; **44**: 31-38.
62. Emilsson V, Thorleifsson G, Zhang B *et al*: Genetics of gene expression and its effect on disease. *Nature* 2008; **452**: 423-428.
63. Schadt EE, Molony C, Chudin E *et al*: Mapping the genetic architecture of gene expression in human liver. *PLoS Biol* 2008; **6**: e107.
64. Chen Y, Zhu J, Lum PY *et al*: Variations in DNA elucidate molecular networks that cause disease. *Nature* 2008; **452**: 429-435.
65. Ackermann M, Strimmer K: A general modular framework for gene set enrichment analysis. *BMC Bioinformatics* 2009; **10**: 47.
66. Carter SL, Brechbuhler CM, Griffin M, Bond AT: Gene co-expression network topology provides a framework for molecular characterization of cellular state. *Bioinformatics* 2004; **20**: 2242-2250.
67. Hudson NJ, Reverter A, Dalrymple BP: A differential wiring analysis of expression data correctly identifies the gene containing the causal mutation. *PLoS Comput Biol* 2009; **5**: e1000382.
68. de la Fuente A: From 'differential expression' to 'differential networking' - identification of dysfunctional regulatory networks in diseases. *Trends Genet* 2010; **26**: 326-333.
69. Subramanian A, Tamayo P, Mootha VK *et al*: Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 2005; **102**: 15545-15550.
70. Dinu I, Potter JD, Mueller T *et al*: Improving gene set analysis of microarray data by SAM-GS. *BMC Bioinformatics* 2007; **8**: 242.
71. Ideker T, Sharan R: Protein networks in disease. *Genome Res* 2008; **18**: 644-652.
72. Nayak RR, Kearns M, Spielman RS, Cheung VG: Coexpression network based on natural variation in human gene expression reveals gene interactions and functions. *Genome Res* 2009; **19**: 1953-1962.
73. Southworth LK, Owen AB, Kim SK: Aging mice show a decreasing correlation of gene expression within genetic modules. *PLoS Genet* 2009; **5**: e1000776.
74. Aten JE, Fuller TF, Lusi AJ, Horvath S: Using genetic markers to orient the edges in quantitative trait networks: the NEO software. *BMC Syst Biol* 2008; **2**: 34.
75. Ghazalpour A, Doss S, Zhang B *et al*: Integrating genetic and network analysis to characterize genes related to mouse weight. *PLoS Genet* 2006; **2**: e130.
76. Miller SA, Dykes DD, Polesky HF: A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 1988; **16**: 1215.

77. Fujita PA, Rhead B, Zweig AS *et al*: The UCSC Genome Browser database: update 2011. *Nucleic Acids Res* 2011; **39**: D876-882.
78. Rozen S, Skaletsky H: Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* 2000; **132**: 365-386.
79. Santos PC, Soares RA, Krieger JE, Guerra-Shinohara EM, Pereira AC: Genotyping of the hemochromatosis HFE p.H63D and p.C282Y mutations by high-resolution melting with the Rotor-Gene 6000(R) instrument. *Clin Chem Lab Med* 2011; **49**: 1633-1636.
80. Campos LC, Miyakawa AA, Barauna VG *et al*: Induction of CRP3/MLP expression during vein arterIALIZATION is dependent on stretch rather than shear stress. *Cardiovasc Res* 2009; **83**: 140-147.
81. Voight BF, Scott LJ, Steinthorsdottir V *et al*: Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat Genet*; **42**: 579-589.
82. Pang DX, Smith AJ, Humphries SE: Functional analysis of TCF7L2 genetic variants associated with type 2 diabetes. *Nutr Metab Cardiovasc Dis* 2012.
83. Tsedensodnom O, Koga H, Rosenberg SA *et al*: Identification of T-cell factor-4 isoforms that contribute to the malignant phenotype of hepatocellular carcinoma cells. *Exp Cell Res* 2011; **317**: 920-931.
84. Mondal AK, Das SK, Baldini G *et al*: Genotype and tissue-specific effects on alternative splicing of the transcription factor 7-like 2 gene in humans. *J Clin Endocrinol Metab* 2010; **95**: 1450-1457.
85. Prokunina-Olsson L, Welch C, Hansson O *et al*: Tissue-specific alternative splicing of TCF7L2. *Hum Mol Genet* 2009; **18**: 3795-3804.
86. Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M: KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res* 1999; **27**: 29-34.
87. Wang X, Xiao Y, Mou Y, Zhao Y, Blankesteyn WM, Hall JL: A role for the beta-catenin/T-cell factor signaling cascade in vascular remodeling. *Circ Res* 2002; **90**: 340-347.
88. Brown KR, Jurisica I: Online predicted human interaction database. *Bioinformatics* 2005; **21**: 2076-2082.
89. Schmittgen TD, Livak KJ: Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc* 2008; **3**: 1101-1108.
90. Livak KJ, Schmittgen TD: Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 2001; **25**: 402-408.
91. Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T: Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* 2010; **27**: 431-432.
92. Cline MS, Smoot M, Cerami E *et al*: Integration of biological networks and gene expression data using Cytoscape. *Nat Protoc* 2007; **2**: 2366-2382.
93. Assenov Y, Ramirez F, Schelhorn SE, Lengauer T, Albrecht M: Computing topological parameters of biological networks. *Bioinformatics* 2008; **24**: 282-284.
94. Myers LaS, M. J.: Spearman Correlation Coefficients, Differences between: *Encyclopedia of Statistical Sciences.*, 2006.
95. Kannel WB, McGee DL: Diabetes and cardiovascular disease. The Framingham study. *JAMA* 1979; **241**: 2035-2038.
96. Mulnier HE, Seaman HE, Raleigh VS *et al*: Risk of myocardial infarction in men and women with type 2 diabetes in the UK: a cohort study using the General Practice Research Database. *Diabetologia* 2008; **51**: 1639-1645.

97. Gaulton KJ, Nammo T, Pasquali L *et al*: A map of open chromatin in human pancreatic islets. *Nat Genet* 2010; **42**: 255-259.
98. Stitzel ML, Sethupathy P, Pearson DS *et al*: Global epigenomic analysis of primary human pancreatic islets provides insights into type 2 diabetes susceptibility loci. *Cell Metab* 2010; **12**: 443-455.
99. Madi HA, Riches K, Warburton P, O'Regan DJ, Turner NA, Porter KE: Inherent differences in morphology, proliferation, and migration in saphenous vein smooth muscle cells cultured from nondiabetic and Type 2 diabetic patients. *Am J Physiol Cell Physiol* 2009; **297**: C1307-1317.
100. Orasanu G, Plutzky J: The pathologic continuum of diabetic vascular disease. *J Am Coll Cardiol* 2009; **53**: S35-42.
101. Lamberti C, Lin KM, Yamamoto Y *et al*: Regulation of beta-catenin function by the IkkappaB kinases. *J Biol Chem* 2001; **276**: 42276-42286.
102. Saegusa M, Hashimura M, Kuwata T, Hamano M, Okayasu I: Crosstalk between NF-kappaB/p65 and beta-catenin/TCF4/p300 signalling pathways through alterations in GSK-3beta expression during trans-differentiation of endometrial carcinoma cells. *J Pathol* 2007; **213**: 35-45.
103. Piga R, Naito Y, Kokura S, Handa O, Yoshikawa T: Short-term high glucose exposure induces monocyte-endothelial cells adhesion and transmigration by increasing VCAM-1 and MCP-1 expression in human aortic endothelial cells. *Atherosclerosis* 2007; **193**: 328-334.
104. Yan SD, Schmidt AM, Anderson GM *et al*: Enhanced cellular oxidant stress by the interaction of advanced glycation end products with their receptors/binding proteins. *J Biol Chem* 1994; **269**: 9889-9897.
105. Askari B, Renard CB, Bornfeldt KE: Regulation of smooth muscle cell accumulation in diabetes-accelerated atherosclerosis. *Histol Histopathol* 2002; **17**: 1317-1328.
106. Gilad Y, Rifkin SA, Pritchard JK: Revealing the architecture of gene regulation: the promise of eQTL studies. *Trends Genet* 2008; **24**: 408-415.
107. Pang DX, Smith AJ, Humphries SE: Functional analysis of TCF7L2 genetic variants associated with type 2 diabetes. *Nutr Metab Cardiovasc Dis*.
108. Fujita PA, Rhead B, Zweig AS *et al*: The UCSC Genome Browser database: update 2011. *Nucleic Acids Res*; **39**: D876-882.
109. Tsedensodnom O, Koga H, Rosenberg SA *et al*: Identification of T-cell factor-4 isoforms that contribute to the malignant phenotype of hepatocellular carcinoma cells. *Exp Cell Res*; **317**: 920-931.
110. Hansson O, Zhou Y, Renstrom E, Osmark P: Molecular function of TCF7L2: Consequences of TCF7L2 splicing for molecular function and risk for type 2 diabetes. *Curr Diab Rep* 2010; **10**: 444-451.
111. Pascoe L, Tura A, Patel SK *et al*: Common variants of the novel type 2 diabetes genes CDKAL1 and HHEX/IDE are associated with decreased pancreatic beta-cell function. *Diabetes* 2007; **56**: 3101-3104.
112. Ridderstrale M, Groop L: Genetic dissection of type 2 diabetes. *Mol Cell Endocrinol* 2009; **297**: 10-17.
113. Bergamaschi A, Kim YH, Kwei KA *et al*: CAMK1D amplification implicated in epithelial-mesenchymal transition in basal-like breast cancer. *Mol Oncol* 2008; **2**: 327-339.
114. Sabapathy K, Hochedlinger K, Nam SY, Bauer A, Karin M, Wagner EF: Distinct roles for JNK1 and JNK2 in regulating JNK activity and c-Jun-dependent cell proliferation. *Mol Cell* 2004; **15**: 713-725.

115. Kwapiszewska G, Wilhelm J, Wolff S *et al*: Expression profiling of laser-microdissected intrapulmonary arteries in hypoxia-induced pulmonary hypertension. *Respir Res* 2005; **6**: 109.
116. Hinek A, Bodnaruk TD, Bunda S, Wang Y, Liu K: Neuraminidase-1, a subunit of the cell surface elastin receptor, desialylates and functionally inactivates adjacent receptors interacting with the mitogenic growth factors PDGF-BB and IGF-2. *Am J Pathol* 2008; **173**: 1042-1056.
117. Rane SG, Dubus P, Mettus RV *et al*: Loss of Cdk4 expression causes insulin-deficient diabetes and Cdk4 activation results in beta-islet cell hyperplasia. *Nat Genet* 1999; **22**: 44-52.
118. Li H, Wang J, Mor G, Sklar J: A neoplastic gene fusion mimics trans-splicing of RNAs in normal human cells. *Science* 2008; **321**: 1357-1361.
119. Newgard CB, Attie AD: Getting biological about the genetics of diabetes. *Nat Med* 2010; **16**: 388-391.
120. Reverter A, Ingham A, Lehnert SA *et al*: Simultaneous identification of differential gene expression and connectivity in inflammation, adipogenesis and cancer. *Bioinformatics* 2006; **22**: 2396-2404.
121. Leonardson AS, Zhu J, Chen Y *et al*: The effect of food intake on gene expression in human peripheral blood. *Hum Mol Genet* 2010; **19**: 159-169.
122. Hegele RA, Cao H: Single nucleotide polymorphisms of RXRA encoding retinoid X receptor alpha. *J Hum Genet* 2001; **46**: 423-425.
123. Gampe RT, Jr., Montana VG, Lambert MH *et al*: Asymmetry in the PPARgamma/RXRalpha crystal structure reveals the molecular basis of heterodimerization among nuclear receptors. *Mol Cell* 2000; **5**: 545-555.
124. Neuville P, Bochaton-Piallat ML, Gabbiani G: Retinoids and arterial smooth muscle cells. *Arterioscler Thromb Vasc Biol* 2000; **20**: 1882-1888.
125. Miano JM, Berk BC: Retinoids: versatile biological response modifiers of vascular smooth muscle phenotype. *Circ Res* 2000; **87**: 355-362.
126. Preston IR, Tang G, Tilan JU, Hill NS, Suzuki YJ: Retinoids and pulmonary hypertension. *Circulation* 2005; **111**: 782-790.
127. Lattuada D, Vigano P, Mangioni S *et al*: Accumulation of retinoid X receptor-alpha in uterine leiomyomas is associated with a delayed ligand-dependent proteasome-mediated degradation and an alteration of its transcriptional activity. *Mol Endocrinol* 2007; **21**: 602-612.
128. Silha JV, Gui Y, Murphy LJ: Impaired glucose homeostasis in insulin-like growth factor-binding protein-3-transgenic mice. *Am J Physiol Endocrinol Metab* 2002; **283**: E937-945.
129. Kliewer SA, Xu HE, Lambert MH, Willson TM: Peroxisome proliferator-activated receptors: from genes to physiology. *Recent Prog Horm Res* 2001; **56**: 239-263.
130. Forman BM, Tontonoz P, Chen J, Brun RP, Spiegelman BM, Evans RM: 15-Deoxy-delta 12, 14-prostaglandin J2 is a ligand for the adipocyte determination factor PPAR gamma. *Cell* 1995; **83**: 803-812.
131. Feige JN, Auwerx J: Transcriptional coregulators in the control of energy homeostasis. *Trends Cell Biol* 2007; **17**: 292-301.
132. Fowler AM, Alarid ET: Dynamic control of nuclear receptor transcription. *Sci STKE* 2004; **2004**: pe51.
133. Hermanson O, Glass CK, Rosenfeld MG: Nuclear receptor coregulators: multiple modes of modification. *Trends Endocrinol Metab* 2002; **13**: 55-60.

134. Li P, Fan W, Xu J *et al*: Adipocyte NCoR knockout decreases PPARgamma phosphorylation and enhances PPARgamma activity and insulin sensitivity. *Cell* 2011; **147**: 815-826.
135. Mototani H, Iida A, Nakamura Y, Ikegawa S: Identification of sequence polymorphisms in CALM2 and analysis of association with hip osteoarthritis in a Japanese population. *J Bone Miner Metab* 2010; **28**: 547-553.
136. Fleischhacker E, Esenabhalu VE, Spitaler M *et al*: Human diabetes is associated with hyperreactivity of vascular smooth muscle cells due to altered subcellular Ca²⁺ distribution. *Diabetes* 1999; **48**: 1323-1330.
137. Guo Z, Su W, Allen S *et al*: COX-2 up-regulation and vascular smooth muscle contractile hyperreactivity in spontaneous diabetic db/db mice. *Cardiovasc Res* 2005; **67**: 723-735.
138. Yoo HJ, Kozaki K, Akishita M *et al*: Augmented Ca²⁺ influx is involved in the mechanism of enhanced proliferation of cultured vascular smooth muscle cells from spontaneously diabetic Goto-Kakizaki rats. *Atherosclerosis* 1997; **131**: 167-175.
139. Okon EB, Chung AW, Rauniyar P *et al*: Compromised arterial function in human type 2 diabetic patients. *Diabetes* 2005; **54**: 2415-2423.
140. Yisraeli JK: VICKZ proteins: a multi-talented family of regulatory RNA-binding proteins. *Biol Cell* 2005; **97**: 87-96.
141. Strieter RM, Burdick MD, Gomperts BN, Belperio JA, Keane MP: CXC chemokines in angiogenesis. *Cytokine Growth Factor Rev* 2005; **16**: 593-609.
142. Liu HM, Luo T, Zhou X *et al*: Disassociation between left ventricular mechanical and electrical properties in ischemic rat heart after G-CSF treatment. *Cardiovasc Drugs Ther* 2011; **25**: 203-214.
143. Kuwabara M, Kakinuma Y, Katare RG *et al*: Granulocyte colony-stimulating factor activates Wnt signal to sustain gap junction function through recruitment of beta-catenin and cadherin. *FEBS Lett* 2007; **581**: 4821-4830.
144. Lyon C, Mill C, Tsaousi A, Williams H, George S: Regulation of VSMC behavior by the cadherin-catenin complex. *Front Biosci* 2011; **16**: 644-657.
145. Chen X, Kelemen SE, Autieri MV: AIF-1 expression modulates proliferation of human vascular smooth muscle cells by autocrine expression of G-CSF. *Arterioscler Thromb Vasc Biol* 2004; **24**: 1217-1222.
146. Yi F, Brubaker PL, Jin T: TCF-4 mediates cell type-specific regulation of proglucagon gene expression by beta-catenin and glycogen synthase kinase-3beta. *J Biol Chem* 2005; **280**: 1457-1464.
147. Vaquero AR, Ferreira NE, Omae SV *et al*: Using gene-network landscape to dissect genotype effects of TCF7L2 genetic variant on diabetes and cardiovascular risk. *Physiol Genomics* 2012; **44**: 903-914.