International Research Journal of Biochemistry and Bioinformatics (ISSN-2250-9941) Vol. 3(6) pp. 115-124, August, 2013
DOI: http:/dx.doi.org/10.14303/irjbb.2013.017
Available online http://www.interesjournals.org/IRJBB

Full Length Research Paper

Basic bioinformatics tools for genotyping studies. A practical exercise with type 2 diabetes mellitus

Alessandro Brito dos Santos[1], Jéssica Nayara Góes de Araújo[1], Rafael Caetano Bastos[1], André Ducati Luchessi[1], Adriana Augusto de Rezende[1], Mario Hiroyuki Hirata[2], Rosario Dominguez Crespo Hirata[2] and Vivian Nogueira Silbiger[1]*

[1]Faculty of Pharmacy, University Federal of Rio Grande do Norte, Brazil.
[2]School of Pharmaceutical Sciences, University of Sao Paulo, Sao Paulo, Brazil.
*Corresponding Authors E-mail: viviansilbiger@hotmail.com

ABSTRACT

The knowledge of bioinformatics tools is essential to investigate the molecular basis of complex diseases, such as type 2 diabetes mellitus (T2DM). The aims of this study were select single nucleotide polymorphisms (SNPs) of candidate genes and elaborate genotyping strategies of T2DM using bioinformatics tools. The genes were selected searching in PubMed/NCBI for articles published from 2007 to 2011 involving T2DM studies. Interactions among the selected genes were studied using the software Ingenuity Pathway Analysis (IPA). SNPs with frequencies higher than 10% were chose according data available at dbSNP/NCBI. Genotyping strategies were based on TaqMan® real-time PCR, PCR-RFLP and sequencing methods, in this order of priority. Primers, probes or enzymes for genotyping strategies were determined using websites and free software. Among the 56 selected genes by PubMed/NCBI research, 32 and 7 were direct and indirectly associated with T2DM, respectively. Five genes were selected to elaborate the genotyping strategies, ADCY5, AGER, IRS2, KCNQ1 and PREX1, for each gene were possible to select 7, 10, 9, 9 and 11 SNPs, respectively. It was possible to select primers and probes for TaqMan® real time-PCR in 65.2% (30/46) of the SNPs selected. Seven of the sixteen SNPs remaining (43.8 %) were viable to genotype by PCR-RFLP, the second strategy. The others nine SNPs were test by third option and just one of them could not be genotype by pyrosequencing, following for the last strategy by Sanger conventional sequencing. In conclusion, this study shows that bioinformatics tools allow the development of a set of genotype strategies for application in future gene candidate association studies for T2DM.

Keywords: bioinformatics, genotyping, database, diabetes mellitus and cardiovascular disease.

INTRODUCTION

Bioinformatics is a highly interdisciplinary field that aims to analyze biological data through technologies and methods of mathematics, statistics, computer science, physics, biology and medicine (Romano *et al.,* 2011). During the past decade, bioinformatics methods and tools have been developed to identify and analyze various components of the cells, such as gene and protein function, interactions and metabolic and regulatory pathways (Zhang *et al.,* 2010). The bioinformatics field has contributed with the exponential growth of all biological databases in recent years, showing that more than one possible abstraction. Bioinformatics became an essential tool for data analysis in molecular biology and structural genomics, such as the development of

methods and strategies for DNA genotyping (Sethi and Theodos, 2009).

Computational investigations of gene functions and their involvement in the pathophysiology of several diseases is promoted greatly by the accumulation of data in the database, of which interaction data have been exploited to identify disease-causing genes (Guo et al., 2011). The identification of genes involved with complex disease has long been a challenge in the study of human genetics (Oti and Brunner, 2007) and the bioinformatics tools are essential in this task.

Single nucleotide polymorphisms (SNPs) are the most abundant variants in many genomes,
and are very important in many fields of genomics. SNPs located in coding and regulatory regions of specific genes have been proven to be associated with increased risk for several diseases, such as type 2 diabetes mellitus (T2DM) (Himelfarb *et al.,* 2011; Cerda *et al*., 2010). Therefore SNPs have becoming useful tools for diagnosis and prognosis in clinical practice (Rodrigues *et al*., 2013).

Many high-throughput SNP genotyping methods have been developed including SNP microarrays, Matrix Assisted Laser Desorption/Ionization-Time of Flight (MALDI-TOF), PCR sequencing, and others. However, most of them require expensive equipments that are unsuitable for small-scale genotyping (Chang *et al.,* 2010).

TaqMan® real-time PCR, for instance, has been applied to SNP genotyping in many candidate gene association studies (Yap *et al.,* 2012; Lima-Neto LG *et al.,* 2012) but some SNPs can't be genotyped using TaqMan® probes (Birdsell *et al.,* 2012; Yang *et al.,* 2012). As an alternative, PCR-Restriction Fragment Length Polymorphism (PCR-RFLP) analysis is able to provide cost-effective SNP genotyping and mutation detection.

Independently of each genotype methods or objective of the study, after human genome sequence it is essentially that all student of science know some basic and free bioinformatics tools to start a genotyping investigation of complex disease, as T2DM.

Considering that the T2DM it is a multi-factorial disease with a strong genetic component (Ismail-Beigi, 2012) and the patients have a four times greater risk of death from cardiovascular events than the general population (Souza *et al.,* 2011), the aims of this study were create a list of potential candidate genes and SNPs, with their primers, probes or restriction enzyme for different strategies of genotype using basic and free bioinformatics tools.

MATERIAL AND METHODS

Genes Selection

Candidate genes for T2DM were selected using the database PubMed of the National Center for Biotechnology Information (NCBI) website to detect a group of relevant studies published as full articles from 2007 to 2011 using some keywords, such as "type 2 diabetes mellitus AND genotyping", "type 2 diabetes mellitus AND SNP", "type 2 diabetes mellitus AND gene" and "type 2 diabetes mellitus AND polymorphism". The list of relevant genes was analyzed by Ingenuity Pathway Analysis (IPA) software (Ingenuity Systems, California, USA) in order to organize the genes into networks of interacting genes, and to identify modules of functionally related genes that correspond to pathways. IPA was an important tool to identify genes that have direct or indirect association with T2DM.

SNPs Selection

SNPs were selected from the dbSNP database build 132. The selection criteria stipulated in accordance with order of priority were: (1) SNPs present in the coding regions with minor allele frequency (MAF) >5%; (2) SNPs in non-coding and promoter regions with MAF >10% included in the HapMap Project; (3) SNPs present in the portions missense and frame shift; (4) SNPs with higher frequency of minor allele. A total less than twelve SNPs were select for each gene.

Genotyping strategies and selection of Primers and Probes

The genotyping strategies were proposed in this study according to the following this order of option: (1) TaqMan® real-time PCR, (2) PCR-RFLP, (3) pyrosequencing by PyroMark®, and (4) conventional sequencing by Sanger method.

The TaqMan® real-time PCR allows the allelic discrimination with high specificity besides it has high trough put and is less time consuming is not expensive, considering that 96 samples or more can be processed in almost 1h 30min. The primers and probes were selected using Primer Express® 3.0 software and the company's website that producer of the software Applied Biosystems.

The PCR-RFLP was the second strategy of choice because it has less through put than TaqMan® system besides the time consuming, usually 12 samples can be processed per assay, with duration of 24 hours each cycle. The primers and restriction enzymes were selected using Primer-BLAST and NEBcutter v2.0 (Vincze *et al.,* 2003), respectively. The PCR products of RFLP present in its sequence a recognition site of restriction enzyme, called constitutive site that allows monitoring of enzyme activity during the reaction of restriction (Lima-Neto *et al.,* 2009).

Table 1. List of genes related with T2DM in different study populations.

Genes	Population Study	References	Genes	Population Study	References
ACAD10	American	Bian L *et al.,* 2010	*IDE*	Japanese	Chiefari E *et al.,* 2010
ACE	Hungarian	Nagy G *et al.,* 2009	*IGF2BP2*	German	Heni M *et al.,* 2010
ADAMTS9	Swedish	Parikh *et al.,* 2009	*IRS1*	Swedish	Parikh *et al.,* 2009
ADCY5	Danish	Andersson EA *et al.,* 2010	*IRS2*	Italian	Cardellini M *et al.,* 2011
ADIPOQ	Brazilian	Vendramini MF *et al.,* 2010	*KCNJ11*	Chinese	Wang F *et al.,* 2009
AGER	Netherlander	Engelen L *et al.,* 2010	*KCNQ1*	Chinese	Tsai FJ *et al.,* 2010
APOE	Brazilian	Terra N *et al.,* 2011	*KIF11*	Japanese	Solaas K *et al.,* 2010
AP3S1	Chinese	Zhou JB *et al.,* 2010	*LOC387761*	Indian	Rong R *et al.,* 2009
BAZ1B	Indian	Chidambaram M *et al.,* 2010	*MAP4K4*	Swedish	Bouzakri *et al.,* 2009
CAMK1D	American	Ezzidi I *et al.,* 2010	*MBL2*	American	MullerYL *et al.,* 2010
CAPN10	Tunisian	Boesgaard TW *et al.,* 2009	*MTNR1B*	Chinese	Liu C *et al.,* 2010
CDKAL1	Japanese	Miyaki K *et al.,* 2010	*NOS1AP*	Chinese	Qin W *et al.,* 2010
CDKN2A	German	Heni M *et al.,* 2010	*NOS2*	Brazilian	Bagarolli RA *et al.,* 2010
CDKN2B	German	Heni M *et al.,* 2010	*NR1H2*	Norwegian	Furukawa *et al.,* 2008
CYP11B2	Saudi Arabia	Chmaisse HN *et al.,* 2009	*PPARG*	Finnish	Vangipurapu J *et al.,* 2010
C2CD4B	American	Shu XO *et al.,* 2010	*PON1*	Indian	Bhaskar S *et al.,* 2011
ELMO2	American	Bento JL *et al.,* 2008	*PREX1*	American	Lewis JP *et al.,* 2010
EMID2	American	Morris DL *et al.,* 2009	*PTPN1*	Spanish	Valverde AM. 2011
EXT2	American	Rong R *et al.,* 2009	*RETN*	Japanese	Asano H *et al.,* 2010
FABP5	American	Bu L *et al.,* 2011	*RXR-α*	Chinese	Lu Y *et al.,* 2011
FOXO1	German	Müssig K *et al.,* 2009	*SERPINA12*	German	Kempf *et al.,* 2010
FTO	British	Timpson NJ *et al.,* 2009	*SPRY2*	American	Bento JL *et al.,* 2008
GCKR	Chinese	Lin Y *et al.,* 2010	*SULF2*	America	Shu XO *et al.,* 2010
GIPR	German	Sauber J *et al.,* 2010	*TAS1R2*	Canadian	Eny KM *et al.,* 2010
G6PC2	Chinese	Hug C *et al.,* 2010	*TLR4*	Brazilian	Bagarolli RA *et al.,* 2010
HHEX	Chinese	Hug C *et al.,* 2010	*TCF7L2*	Brazilian	Franklin CS *et al.,* 2010
HIF-1α	Hungarian	Furukawa *et al.,* 2008	*WFS1*	German	Heni M *et al.,* 2010
HMGA1	Italian	Nagy G *et al.,* 2009	*ZNF334*	American	Bento JL *et al.,* 2008

The pyrosequencing by PyroMark® of the QIAGEN company was the third strategy of choice, this it is more expensive technology than the others, but is faster, usually 190 – 300 samples can be processed per assay with duration of 45 minutes each. The primers and probe were selected using the PyroMark® Assay Design 2.0.
Finally, the last technology choice at this study was the conventional sequence method of Sanger, which it still more expensive that the all methods presented at this study. The primers were select using Primer-BLAST.
The specificity of all primers and probes were tested by Primer-BLAST tools available at NCBI website.

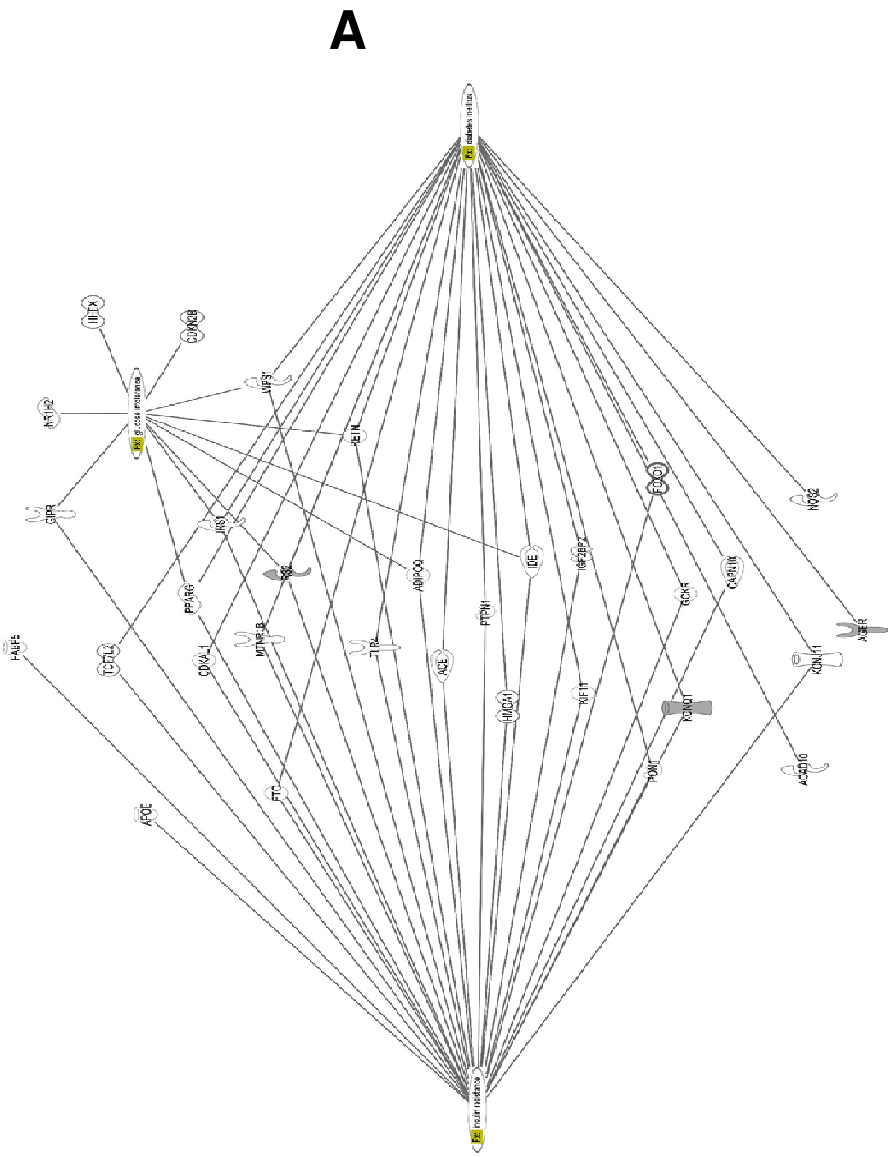
RESULTS

Selected genes

Fifty-six candidate genes for T2DM were selected based on the information available in PubMed public database from 2007 to 2011 (Table 1). All of these genes were analyzed by IPA software to verify putative relationships with T2DM. Thirty-two genes showed direct association with "T2DM", "resistance of insulin" or "glucose intolerance" (Figure 1.A) and other seven (*AP3S1*, *CDKN2A*, *HIF1A*, *MAP4K4*, *SERPINA12*, *SPRY2*) were indirectly associated with genes directly linked with some nodes of this network (Figure 1.B). Seventeen genes were not direct or indirect associated with T2DM, but three of them (*ADCY5*, *CYP11B2*, *PREX1*) were linked to cardiovascular disease by IPA (Figure 1.C). Among these genes *ADCY5*, *PREX1* (in gray at Figure 3.C) and *AGER, IRS2, KCNQ1* (in gray at Figure 1.A) were selected as representative to the next steps of this study.

Selected SNPs and genotyping strategies

Based in the established criteria, 46 SNPs were select in *ADCY5* (7), *AGER* (10), *IRS2* (9), *KCNQ1* (9), and



PREX1 (11). Figure 2 shows gene structure and location of the selected SNPs. The genotyping strategies, oligonucleotide sequences and/or restriction enzymes are shown in Tables 2 to 6 for *ADCY5*, *AGER*, *IRS2*, *KCNQ1* and *PREX1*, respectively. All oligonucleotide were specific to their genes by Primer-BLAST verification.

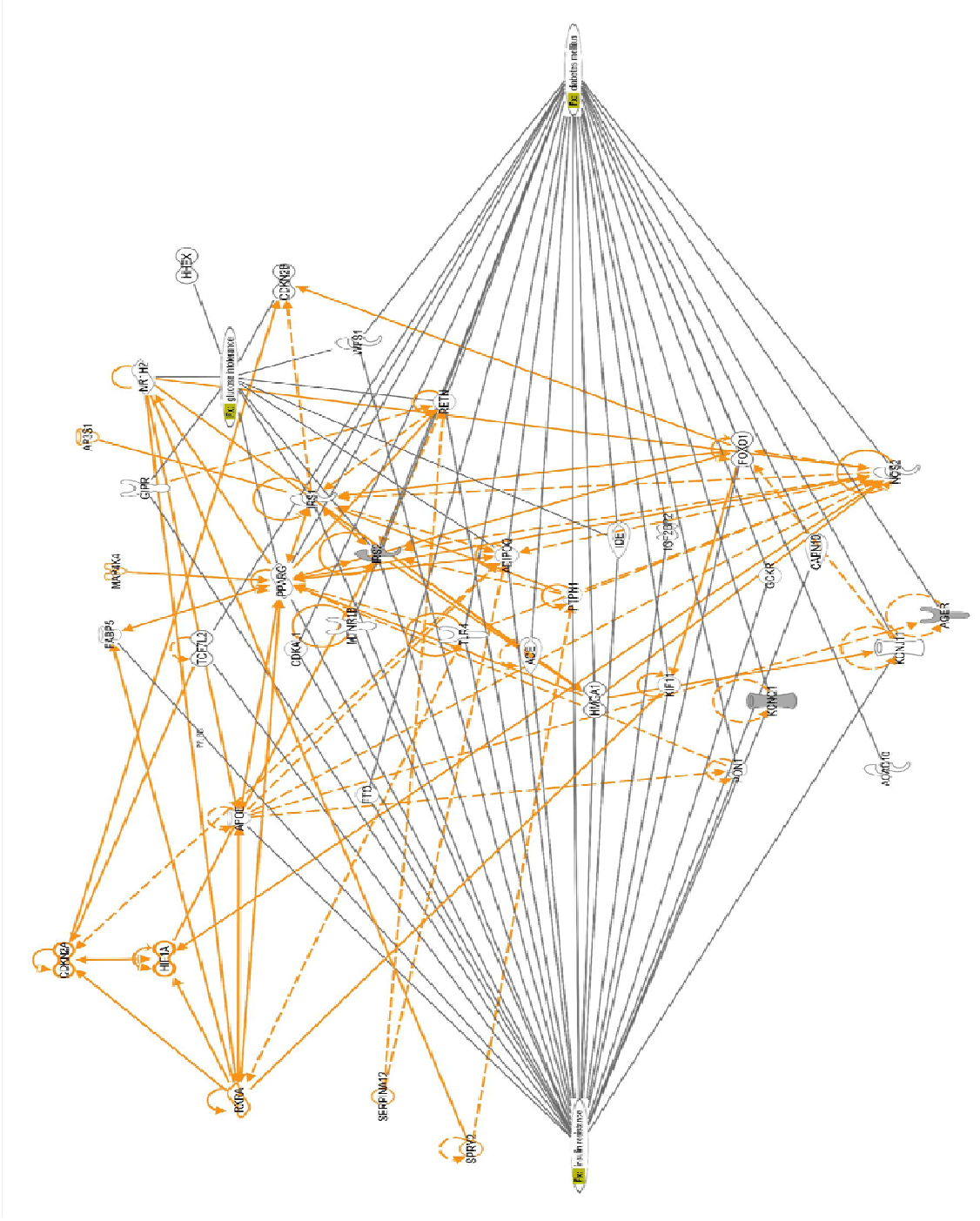
The allelic discrimination by TaqMan® system was possible for 65.2% (30/46) of the SNPs selected 19 of them by Primer Express® 3.0 software and 11 by Applied Biosystems database (Assay ID).

Seven of the sixteen SNPs remaining (43.8 %) were viable to genotype by PCR-RFLP, the second strategy. The others nine SNPs were test by third option and just one of them could not be genotype by pyrosequencing, following for the last strategy by Sanger conventional sequencing.

DISCUSSION

The T2DM was because several studies show that T2DM patients present 2-6 times higher incidence of coronary artery disease and cerebrovascular accident being regarded as having an equivalent risk to a patient without diabetes with pre-existing disease of the heart (Gomes *et al.,* 2009). T2DM is a multi-factorial disease, both in the formation of atherosclerotic plaque, a major cause of mortality, as being the leading cause of blindness, amputations of legs and chronic kidney disease (Ismail-Beigi, 2012), and being genes and SNPs significantly associated with this pathology, the genotypic identification of the component is presented as the main factor for the understanding, treatment and prevention of T2DM.

B



According to Dupuis and collaborators (2010) variations in the *ADCY5* in European populations is strongly associated with T2DM. This suggests that the variants within this gene may exert it's by β-cell dysfunction pancreatic and insulin secretion (Freathy *et al.,* 2010). In this study show three SNPs in *ADCY5* that cause a missense mutation, the rs1123376754 and rs111942308 with high MAF (50%), and the rs61734561 with 5.6%.

The presence of polymorphisms in the regulatory regions in the *AGER* provide alterations that evidence the

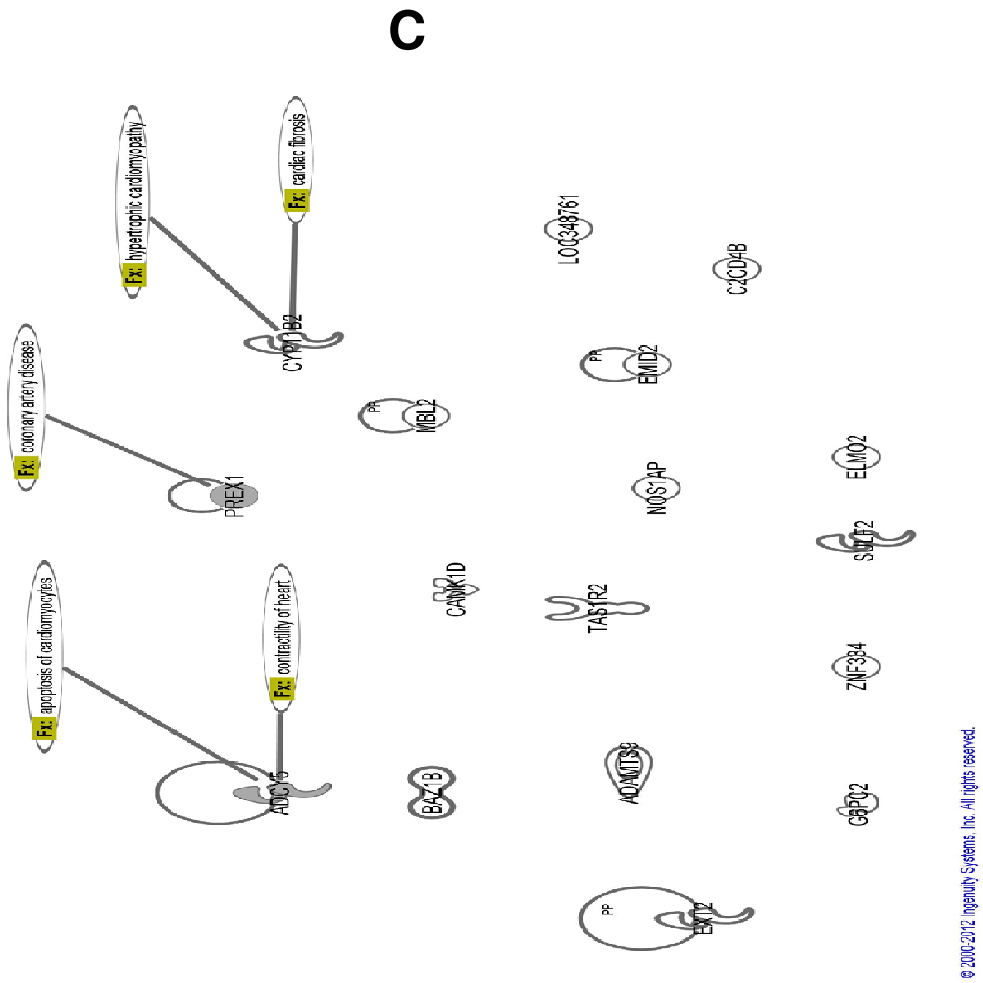


Figure 1. IPA analysis with all genes selected by literature review. A. genes associated with the nodes "T2DM", "insulin resistance" and "glucose intolerance". B. genes not linked with the nodes but associated between each other gene, the orange lines represent direct (solid line) and indirect (dashed line). C. genes associated with nodes involved with cardiovascular disease.

association between SNPs in the promoter region and quantitative traits related to glucose homeostasis (Boutant *et al.,* 2012). Two SNPs observed in promoter region (rs1800624, rs113760854) show MAF > 25%.

Liu *et al.,* demonstrated that mice knockout for *IRS2* have marked hyperglycemia due to various abnormalities in insulin action in peripheral tissues and the failure of the secretory activity of pancreatic b-cells followed by a significant reduction in pancreatic beta cells mass (Liu *et al.,* 2011).

According to this study, 8 out of 9 selected SNPs of the *IRS2* were in the coding region. Two of these SNPs (rs113301332; rs116906692) have missense mutations with MAF > 35%. The SNPs rs11838537 is located in the promoter region

The increased risk for development of T2DM is directly linked to genetic variations in the *KCNQ1*, resulting in deficient secretion of insulin by the pancreas (Müssig *et al.,* 2009a,b).

In this study, all SNPs selected to *KCNQ1*. One (rs10644111) is a frameshift mutation located in the coding region, resulting in the insertion of a triplete (ACC) in the gene sequence (Masood and Kayani, 2011).

PREX1 polymorphisms are strongly associated with the onset of T2DM in Europe with American ancestry, mainly polymorphisms associated with the region of chromosome 20q12-1.13 (Lewis *et al.,* 2010). Two SNPs (rs1885290, rs6019445) are present in the promoter region located at position 47445579 and 47445915, respectively, of chromosome 20. All nine SNPs in the region coding cause a missense mutation.

In this study, we could design TaqMan®-based genotyping strategies for 30 out of 46 selected SNPs .The remain (16) SNPs had no putative efficient assays based in TaqMan strategy because the amplicons sizes (up to 80 bp) were considered to be small, the regions flanked by primers and probes may be non-specific sequences with similarity to other genes or even similar

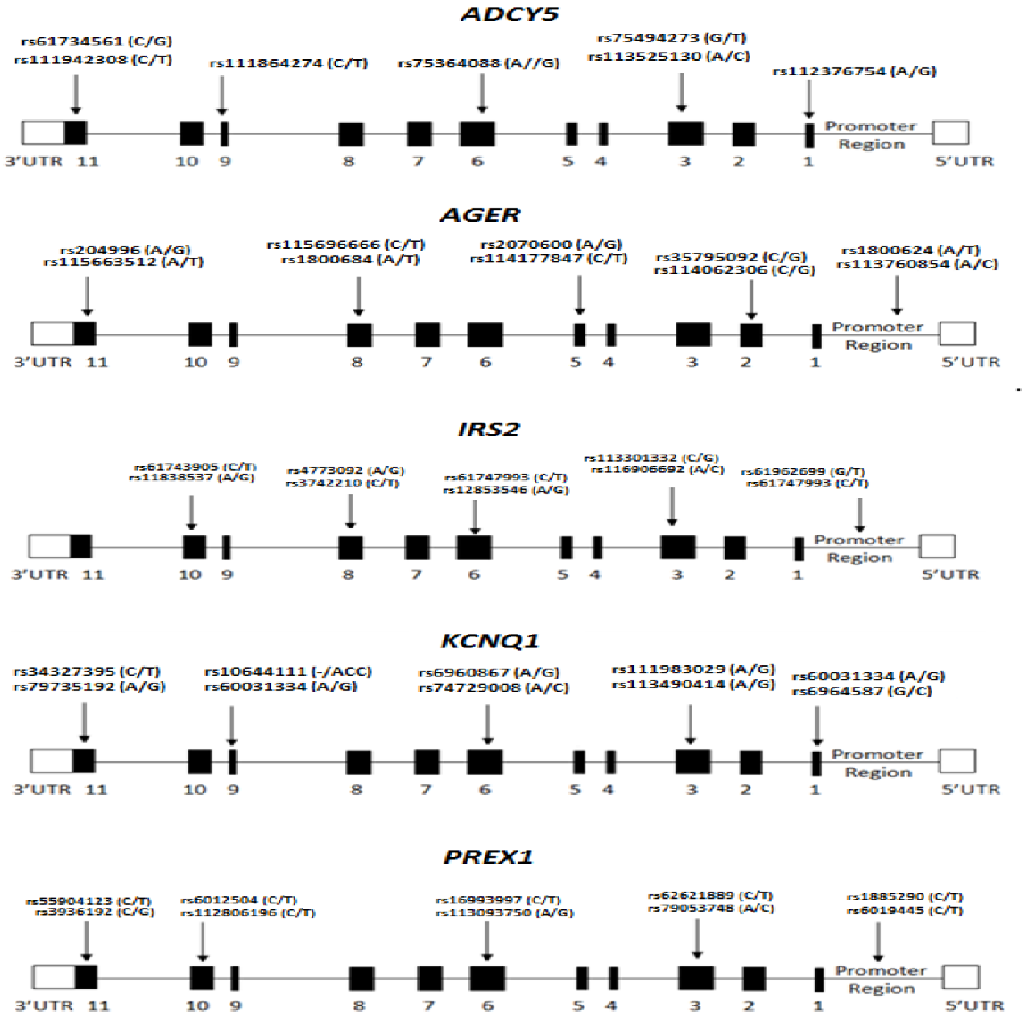


Figure 2. Gene structure and location of the selected SNPs.

to non-human genes. Another difficulty was the occurrence of a mismatch between primers and probe design with a DNA template therefore this mismatch decreases the extension efficiency of Taq polymerase by 15% to 50% per cycle resulting in lower PCR efficiency (Birdsell *et al.,* 2012).

Seven out of 16 SNPs without an effective genotyping strategy by TaqMan® real-time PCR, had an assay designed by PCR-RFLP, following our proposed .Besides the polymorphic site, the proposed strategies considered the presence of a second cleavage site for evaluation of the endonuclease activity in every assay to assure the correct genotyping. Nine SNPs could not be detected by PCR-RFLP due to some limitations such as the lack of endonuclease site, a single endonuclease site and multiple enzyme sites.

The third option for genotype strategic was by pyrosequencing. According to Shi and collaborators (2011), pyrosequencing allows the use of accurately and quickly detects polymorphisms in the genes in time saving. The PyroMark® also allow a high throughput for the detection of minor variations of less than 1% (Stürmer and Reinheime, 2012), indicating that this method can be sufficiently sensitive for use in association studies involving complex diseases where a small frequency difference between cases and controls allele is expected (Shi *et al.,* 2011). However, the high cost to the genotyping was significant for using this method. Eight out of 9 SNPs, which could not be detected by the previous strategies, were tested to be detected by pyrosequencing. We could not design a pyrosequencing strategy for one SNP because the oligonucleotide

sequences selected by software PyroMark® were not specific for the DNA template by Prime-BLAST verification, this last SNP could be genotype by the conventional sequence method of Sanger, our last option.

In conclusion, those methodologies through various bioinformatics tools have a high yield in both the selection of candidate genes and SNPs and the selection of primers, probes and enzymes for numerous diseases. And this exercise using the T2DM shows that bioinformatics tools allow the development of a set of genotype strategic for application in future association-studies with complex disease.

ACKNOWLEDGMENTS

Alessandro Brito dos Santos and Elyson Diego S. S. Osório are recipients of fellowships from REUNI/PIBIQ-CNPq, Brazil. Mario H. Hirata and Rosario DC Hirata are recipients of fellowships from CNPq, Brazil.

REFERENCES

Andersson EA, Pilgaard K, Pisinger C, Harder MN, Grarup N, Faerch K, Poulsen P, Witte DR, Jørgensen T, Vaag A, Hansen T, Pedersen O (2010). Type 2 diabetes risk alleles near ADCY5, CDKAL1 and HHEX-IDE are associated with reduced birthweight. Diabetologia. 53(9): 1908-16.

Asano H, Izawa H, Nagata K, Nakatochi M, Kobayashi M, Hirashiki A, Shintani S, Nishizawa T, Tanimura D, Naruse K, Matsubara T, Murohara T, Yokota M (2010). Plasma resistin concentration determined by common variants in the resistin gene and associated with metabolic traits in an aged Japanese population. Diabetologia. 53(2): 234-46.

Bagarolli RA, Saad MJ, Saad ST (2010). Toll-like receptor 4 and inducible nitric oxide synthase gene polymorphisms are associated with Type 2 diabetes. J Diabetes Complications. 24(3): 192-8.

Bento JL, Palmer ND, Zhong M, Roh B, Lewis JP, Wing MR, Pandya H, Freedman BI, Langefeld CD, Rich SS, Bowden DW, Mychaleckyj JC (2008). Heterogeneity in gene loci associated with type 2 diabetes on human chromosome 20q13.1. Genomics. 92(4): 226-34.

Bhaskar S, Ganesan M, Chandak GR, Mani R, Idris MM, Khaja N, Gulla S, Kumar U, Movva S, Vattam KK, Eppa K, Hasan Q, Pulakurthy UR (2011). Association of PON1 and APOA5 Gene Polymorphisms in a Cohort of Indian Patients Having Coronary Artery Disease With and Without Type 2 Diabetes. Genet Test Mol Biomarkers. 15(7-8): 507-12.

Birdsell DN, Pearson T, Price EP, Hornstra HM, Nera RD, Stone N, Gruendike J, Kaufman EL, Pettus AH, Hurbon AN, Buchhagen JL, Harms NJ, Chanturia G, Gyuranecz M, Wagner DM, Keim PS (2012). Melt Analysis of Mismatch Amplification Mutation Assays (Melt-MAMA): A Functional Study of a Cost-Effective SNP Genotyping Assay in Bacterial Models. PLoS One. 7(3): e32866.

Bian L, Hanson RL, Muller YL, Ma L; MAGIC Investigators, Kobes S, Knowler WC, Bogardus C, Baier LJ (2010). Variants in ACAD10 are associated with type 2 diabetes, insulin resistance and lipid oxidation in Pima Indians. Diabetologia. 53(7): 1349-53.

Boesgaard TW, Gjesing AP, Grarup N, Rutanen J, Jansson PA, Hribal ML, Sesti G, Fritsche A, Stefan N, Staiger H, Häring H, Smith U, Laakso M, Pedersen O, Hansen T (2009). Variant near ADAMTS9 known to associate with type 2 diabetes is related to insulin resistance in offspring of type 2 diabetes patients--EUGENE2 study. PLoS One. 4(9): e7236.

Boutant M, Ramos OH, Lecoeur C, Vaillant E, Philippe J, Zhang P, Perilhou A, Valcarcel B, Sebert S, Jarvelin MR, Balkau B, Scott D, Froguel P, Vaxillaire M, Vasseur-Cognet M (2012). Glucose-Dependent Regulation of NR2F2 Promoter and Influence of SNP-rs3743462 on Whole Body Insulin Sensitivity. PLoS One. 7(5): e35810.

Bouzakri K, Ribaux P, Halban PA (2009). Silencing mitogen-activated protein 4 kinase 4 (MAP4K4) protects beta cells from tumor necrosis factor-alpha-induced decrease of IRS-2 and inhibition of glucose-stimulated insulin secretion. J Biol Chem. 284(41): 27892-8.

Bu L, Salto LM, De Leon KJ, De Leon M (2011). Polymorphisms in fatty acid binding protein 5 show association with type 2 diabetes. Diabetes Res Clin Pract. 92(1): 82-91.

Cerda A, Genvigir FD, Arazi SS, Hirata MH, Dorea EL, Bernik MM, Bertolami MC, Faludi AA, Hirata RD (2010). Influence of SCARB1 polymorphisms on serum lipids of hypercholesterolemic individuals treated with atorvastatin. Clin Chim Acta. 411(9-10): 631-7.

Chang HW, Cheng YH, Chuang LY, Yang CH (2010). SNP-RFLPing 2: an updated and integrated PCR-RFLP tool for SNP genotyping. BMC Bioinformatics. 11: 173.

Chidambaram M, Radha V, Mohan V (2010). Replication of recently described type 2 diabetes gene variants in a South Indian population. Metabolism. 59(12): 1760-6.

Chiefari E, Iiritano S, Paonessa F, Le Pera I, Arcidiacono B, Filocamo M, Foti D, Liebhaber SA, Brunetti A (2010). Pseudogene-mediated posttranscriptional silencing of HMGA1 can result in insulin resistance and type 2 diabetes. Nat Commun. 1: 40.

Chmaisse HN, Jammal M, Fakhoury H, Fakhoury R (2009). A study on the association between angiotensin-I converting enzyme I/D dimorphism and type-2 diabetes mellitus Saudi J Kidney. Dis Transpl. 20(6): 1038-46.

Dupuis J, Langenberg C, Prokopenko I, Saxena R, Soranzo N, Jackson AU, Wheeler E, Glazer NL, Bouatia-Naji N, Gloyn AL *et al.* (2010). New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. Nat Genet. 42(2): 105-16.

Engelen L, Ferreira I, Gaens KH, Henry RM, Dekker JM, Nijpels G, Heine RJ, 't Hart LM, van Greevenbroek MM, van der Kallen CJ, Blaak EE, Feskens EJ, Ten Cate H, Stehouwer CD, Schalkwijk CG (2010). The association between the -374T/A polymorphism of the receptor for advanced glycation endproducts gene and blood pressure and arterial stiffness is modified by glucose metabolism status: the Hoorn and CoDAM studies. J Hypertens. 28(2): 285-93.

Eny KM, Wolever TM, Corey PN, El-Sohemy A (2010). Genetic variation in TAS1R2 (Ile191Val) is associated with consumption of sugars in overweight and obese individuals in 2 distinct populations. Am J Clin Nutr. 92(6): 1501-10.

Ezzidi I, Turki A, Messaoudi S, Chaieb M, Kacem M, Al-Khateeb GM, Mahjoub T, Almawi WY, Mtiraoui N (2010). Common polymorphisms of calpain-10 and the risk of Type 2 Diabetes in a Tunisian Arab population: a case-control study. BMC Med Genet. 11: 75.

Franklin CS, Aulchenko YS, Huffman JE, Vitart V, Hayward C, Polašek O, Knott S, Zgaga L, Zemunik T, Rudan I, Campbell H, Wright AF, Wild SH, Wilson JF (2010). The TCF7L2 diabetes risk variant is associated with HbA_1(C) levels: a genome-wide association meta-analysis. Ann Hum Genet. 74(6): 471-8.

Freathy RM, Mook-Kanamori DO, Sovio U, Prokopenko I, Timpson NJ, Berry DJ, Warrington NM, Widen E, Hottenga JJ, Kaakinen M *et al.* (2010). Variants in ADCY5 and near CCNL1 are associated with fetal growth and birth weight. Nat Genet. 42(5): 430-5.

Furukawa Y, Shimada T, Furuta H, Matsuno S, Kusuyama A, Doi A, Nishi M, Sasaki H, Sanke T, Nanjo K (2008). Polymorphisms in the IDE-KIF11-HHEX gene locus are reproducibly associated with type

2 diabetes in a Japanese population. J Clin Endocrinol Metab. 93(1): 310-.
Gomes MB, Giannella-Neto D, Faria M, Tambascia M, Fonseca RM, Rea R, Macedo G, Modesto-Filho J, Schmid H, Bittencourt AV, Cavalcanti S, Rassi N, Pedrosa H, Dib SA (2009). Estimating cardiovascular risk in patients with type 2 diabetes: a national multicenter study in Brazil. Diabetol Metab Syndr. 1(1): 22.
Guo X, Gao L, Wei C, Yang X, Zhao Y, Dong A (2011). A Computational Method Based on the Integration of Heterogeneous Networks for Predicting Disease-Gene Associations. PLoS One. 6(9): e24171.
Heni M, Ketterer C, Thamer C, Herzberg-Schäfer SA, Guthoff M, Stefan N, Machicao F, Staiger H, Fritsche A, Häring HU (2010). Glycemia determines the effect of type 2 diabetes risk genes on insulin secretion. Diabetes. 59(12): 3247-52.
Himelfarb ST, Silva FA, Arazi SS, Farjado CM, Garofalo A, Bertolami MC, Bertolami A, Faludi A, Sampaio MF, Rezende AA, Hirata RD, Hirata MH (2011). Tumor necrosis factor-α and interleukin-6 expression in leukocytes and their association with polymorphisms and bone markers in diabetic individuals treated with pioglitazone. Drug Metabol Drug Interact. 26(1): 37-40.
Ismail-Beigi F (2012). Pathogenesis and glycemic management of type 2 diabetes mellitus: a physiological approach. Arch Iran Med. 15(4): 239-46.
Kempf K, Rose B, Illig T, Rathmann W, Strassburger K, Thorand B, Meisinger C, Wichmann HE, Herder C, Vollmert C (2010). Vaspin (SERPINA12) genotypes and risk of type 2 diabetes: Results from the MONICA/KORA studies. Exp Clin Endocrinol Diabetes. 118(3): 184-9.
Lewis JP, Palmer ND, Ellington JB, Divers J, Ng MC, Lu L, Langefeld CD, Freedman BI, Bowden DW (2010). Analysis of candidate genes on chromosome 20q12-13.1 reveals evidence for BMI mediated association of PREX1 with type 2 diabetes in European Americans. Genomics. 96(4): 211-9.
Lima-Neto LG, Hirata RD, Luchessi AD, Silbiger VN, Pastorelli CP, Sampaio MF, Armaganijan D, Rezende AA, Doi SQ, Hirata MH (2009). Detection of the TLR4 1196C>T polymorphism by mismatched-polymerase chain reaction using plasmid DNA as internal control in restriction fragment length polymorphism assays. Genet Test Mol Biomarkers. 13(3): 343-7.
Lima-Neto LG, Hirata RD, Luchessi AD, Silbiger VN, Stephano MA, Sampaio MF, Armaganijan D, Hirata MH (2012). CD14 and IL6 polymorphisms are associated with a pro-atherogenic profile in young adults with acute myocardial infarction. J Thromb Thrombolysis. doi: 10.1007/s11239-012-0841-4.
Lin Y, Li P, Cai L, Zhang B, Tang X, Zhang X, Li Y, Xian Y, Yang Y, Wang L, Lu F, Liu X, Rao S, Chen M, Ma S, Shi Y, Bao M, Wu J, Yang Y, Yang J, Yang Z (2010). Association study of genetic variants in eight genes/loci with type 2 diabetes in a Han Chinese population. BMC Med Genet. 11: 97.
Liu C, Wu Y, Li H, Qi Q, Langenberg C, Loos RJ, Lin X (2010). MTNR1B rs10830963 is associated with fasting plasma glucose, HbA1C and impaired beta-cell function in Chinese Hans from Shanghai. BMC Med Genet. 11: 59.
Liu Z, Kim W, Chen Z, Shin YK, Carlson OD, Fiori JL, Xin L, Napora JK, Short R, Odetunde JO, Lao Q, Egan JM (2011). Insulin and glucagon regulate pancreatic α-cell proliferation. PLoS One. 6(1): e16096.
Lu Y, Ye X, Cao Y, Li Q, Yu X, Cheng J, Gao Y, Ma J, Du W, Zhou L (2011). Genetic Variants in Peroxisome Proliferator-Activated Receptor-γ and Retinoid X Receptor-α Gene and Type 2 Diabetes Risk: A Case-Control Study of a Chinese Han Population. Diabetes Technol Ther. 13(2): 157-64.
Masood N, Kayani MA (2011). Mutational analysis of xenobiotic metabolizing genes (CYP1A1 and GSTP1) in sporadic head and neck cancer patients. Genet Mol Biol. 34(4): 533-8.
Miyaki K, Oo T, Song Y, Lwin H, Tomita Y, Hoshino H, Suzuki N, Muramatsu M (2010). Association of a cyclin-dependent kinase 5 regulatory subunit-associated protein 1-like 1 (CDKAL1) polymorphism with elevated hemoglobin A_1(c) levels and the prevalence of metabolic syndrome in Japanese men: interaction with dietary energy intake. Am J Epidemiol. 172(9): 985-91.
Morris DL, Cho KW, Zhou Y, Rui L (2009). SH2B1 enhances insulin sensitivity by both stimulating the insulin receptor and inhibiting tyrosine dephosphorylation of insulin receptor substrate proteins. Diabetes. 58(9): 2039-47.
Muller YL, Hanson RL, Bian L, Mack J, Shi X, Pakyz R, Shuldiner AR, Knowler WC, Bogardus C, Baier LJ (2010). Functional variants in MBL2 are associated with type 2 diabetes and pre-diabetes traits in Pima Indians and the old order Amish. Diabetes. 59(8): 2080-5.
Müssig K, Staiger H, Machicao F, Kirchhoff K, Guthoff M, Schäfer SA, Kantartzis K, Silbernagel G, Stefan N, Holst JJ, Gallwitz B, Häring HU, Fritsche A (2009b). Association of type 2 diabetes candidate polymorphisms in KCNQ1 with incretin and insulin secretion. Diabetes. 58(7): 1715-20.
Müssig K, Staiger H, Machicao F, Stancáková A, Kuusisto J, Laakso M, Thamer C, Machann J, Schick F, Claussen CD, Stefan N, Fritsche A, Häring HU (2009a). Association of common genetic variation in the FOXO1 gene with beta-cell dysfunction, impaired glucose tolerance, and type 2 diabetes. J Clin Endocrinol Metab. 94(4): 1353-60.
Nagy G, Kovacs-Nagy R, Kereszturi E, Somogyi A, Szekely A, Nemeth N, Hosszufalusi N, Panczel P, Ronai Z, Sasvari-Szekely M (2009). Association of hypoxia inducible factor-1 alpha gene polymorphism with both type 1 and type 2 diabetes in a Caucasian (Hungarian) sample. BMC Med Genet. 10: 79.
Oti M, Brunner HG (2007). The modular nature of genetic diseases. Clin Genet. 71(1): 1-11.
Parikh H, Lyssenko V, Groop LC (2009). Prioritizing genes for follow-up from genome wide association studies using information on gene expression in tissues relevant for type 2 diabetes mellitus. BMC Med Genomics. 2: 72.
Qin W, Zhang R, Hu C, Wang CR, Lu JY, Yu WH, Bao YQ, Xiang KS; International Type 2 Diabetes 1q Consortium, Jia WP (2010). A variation in NOS1AP gene is associated with repaglinide efficacy on insulin resistance in type 2 diabetes of Chinese. Acta Pharmacol Sin. 31(4): 450-4.
Rodrigues AC, Sobrino B, Genvigir FD, Willrich MA, Arazi SS, Dorea EL, Bernik MM, Bertolami M, Faludi AA, Brion MJ, Carracedo A, Hirata MH, Hirata RD (2013). Genetic variants in genes related to lipid metabolism and atherosclerosis, dyslipidemia and atorvastatin response. Clin Chim Acta. 417: 8-11.
Romano P, Giugno R, Pulvirenti A (2011). Tools and collaborative environments for bioinformatics research. Brief Bioinform. 12(6): 549–561.
Rong R, Hanson RL, Ortiz D, Wiedrich C, Kobes S, Knowler WC, Bogardus C, Baier LJ (2009). Association analysis of variation in/near FTO, CDKAL1, SLC30A8, HHEX, EXT2, IGF2BP2, LOC387761, and CDKN2B with type 2 diabetes and related quantitative traits in Pima Indians. Diabetes. 58(2): 478-88.
Sanger F, Nicklen S, Coulson AR (1977). DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A. 74(12): 5463-7.
Sauber J, Grothe J, Behm M, Scherag A, Grallert H, Illig T, Hinney A, Hebebrand J, Wiegand S, Grüters A, Krude H, Biebermann H (2010). Association of variants in gastric inhibitory polypeptide receptor gene with impaired glucose homeostasis in obese children and adolescents from Berlin. Eur J Endocrinol. 163(2): 259-64.
Sethi P, Theodos K (2009). Translational bioinformatics and healthcare informatics: computational and ethical challenges. Perspect Health Inf Manag. 6: 1h
Shi H, Yu RL, Ma JF, Ren XY (2011). Development of a single-tube PCR-pyrosequencing method for simultaneous and rapid detection

of the genetic polymorphism of warfarin metabolizing enzymes. Yi Chuan. 33(11): 1283-90.
Shu XO, Long J, Cai Q, Qi L, Xiang YB, Cho YS, Tai ES, Li X, Lin X, Chow WH *et al.* (2010). Identification of new genetic risk variants for type 2 diabetes. PLoS Genet. 6(9): e1001127.
Solaas K, Legry V, Retterstol K, Berg PR, Holven KB, Ferrières J, Amouyel P, Lien S, Romeo J, Valtueña J *et al.* (2010). Suggestive evidence of associations between liver X receptor β polymorphisms with type 2 diabetes mellitus and obesity in three cohort studies: HUNT2 (Norway), MONICA (France) and HELENA (Europe). BMC Med Genet. 11: 144.
Souza AG, Lopes NH, Hueb WA, Krieger JE, Pereira AC (2011). Genetic Variants of Diabetes Risk and Incident Cardiovascular Events in Chronic Coronary Artery Disease. PLoS One. 6(1): e16341.
Stürmer M, Reinheimer C (2012). Description of two commercially available assays for genotyping of HIV-1. Intervirology. 55(2): 134-7.
Terra N, Moriguchi Y, Bittencourt L, Trois RS, Piccoli JE, Cruz IB (2011). Apolipoprotein E polymorphism in elderly Japanese-Brazilian immigrants does not explain the reduced cardiovascular risk factor incidence. Genet Mol Res. 10(3): 1975-85.
Timpson NJ, Lindgren CM, Weedon MN, Randall J, Ouwehand WH, Strachan DP, Rayner NW, Walker M, Hitman GA, Doney AS *et al.* (2009). Adiposity-related heterogeneity in patterns of type 2 diabetes susceptibility observed in genome-wide association data. Diabetes. 58(2): 505-10.
Tsai FJ, Yang CF, Chen CC, Chuang LM, Lu CH, Chang CT, Wang TY, Chen RH, Shiu CF, Liu YM, Chang CC *et al.* (2010) A genome-wide association study identifies susceptibility variants for type 2 diabetes in Han Chinese. PLoS Genet. 6(2):e1000847.
Vangipurapu J, Stančáková A, Pihlajamäki J, Kuulasmaa TM, Kuulasmaa T, Paananen J, Kuusisto J, Ferrannini E, Laakso M (2010). Association of indices of liver and adipocyte insulin resistance with 19 confirmed susceptibility loci for type 2 diabetes in 6,733 non-diabetic Finnish men. Diabetologia. 54(3): 563-71.
Vendramini MF, Pereira AC, Ferreira SR, Kasamatsu TS, Moisés RS (2010). Association of genetic variants in the adiponectin encoding gene (ADIPOQ) with type 2 diabetes in Japanese Brazilians. J Diabetes Complications. 24(2): 115-20.
Vincze T, Posfai J, Roberts RJ (2003). NEBcutter: a program to cleave DNA with restriction enzymes. Nucleic Acids Res. 31: 3688-3691.
Wang F, Han XY, Ren Q, Zhang XY, Han LC, Luo YY, Zhou XH, Ji LN (2009). Effect of genetic variants in KCNJ11, ABCC8, PPARG and HNF4A loci on the susceptibility of type 2 diabetes in Chinese Han population. Chin Med J (Engl). 122(20): 2477-82.
Yang CH, Cheng YH, Yang CH, Chuang LY (2012). Mutagenic Primer Design for Mismatch PCR-RFLP SNP Genotyping using a Genetic Algorithm. IEEE/ACM Trans Comput Biol Bioinform. 9(3): 837-45.
Yap RW, Shidoji Y, Hon WM, Masaki M (2012). Association and interaction between dietary pattern and VEGF receptor-2 (VEGFR2) gene polymorphisms on blood lipids in Chinese Malaysian and Japanese adults. Asia Pac J Clin Nutr. 21(2): 302-11.
Zhang S, Ma G, Xiang J, Cheng A, Wang M, Zhu D, Jia R, Luo Q, Chen Z, Chen X (2010). Expressing gK gene of duck enteritis virus guided by bioinformatics and its applied prospect in diagnosis. Virol J. 7: 168.

Internet Resources

1. Literature database
http://www.ncbi.nlm.nih.gov/pubmed
2. dbSNP database
http://www.ncbi.nlm.nih.gov/SNP/
3. Primer-BLAST
http://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi?LINK_LOC=BlastHome
4. Applied Biosystems database
http://www.appliedbiosystems.com.br
5. Ingenuity® Systems
http://www.ingenuity.com/

How to cite this article: Santos BA, Araújo JN, Bastos RC, Luchessi AD, Rezende AA, Hirata MH, Hirata RDC and Silbiger VN (2013). Basic bioinformatics tools for genotyping studies. A practical exercise with type 2 diabetes mellitus. Int. Res. J. Biochem. Bioinform. 3(6):115-124