UNIVERSITÀ DEGLI STUDI DI MILANO
Scuola di Dottorato in Scienze Biochimiche, Nutrizionali e Metaboliche
Corso di Dottorato di Ricerca in Scienze Biochimiche XXVIII Ciclo

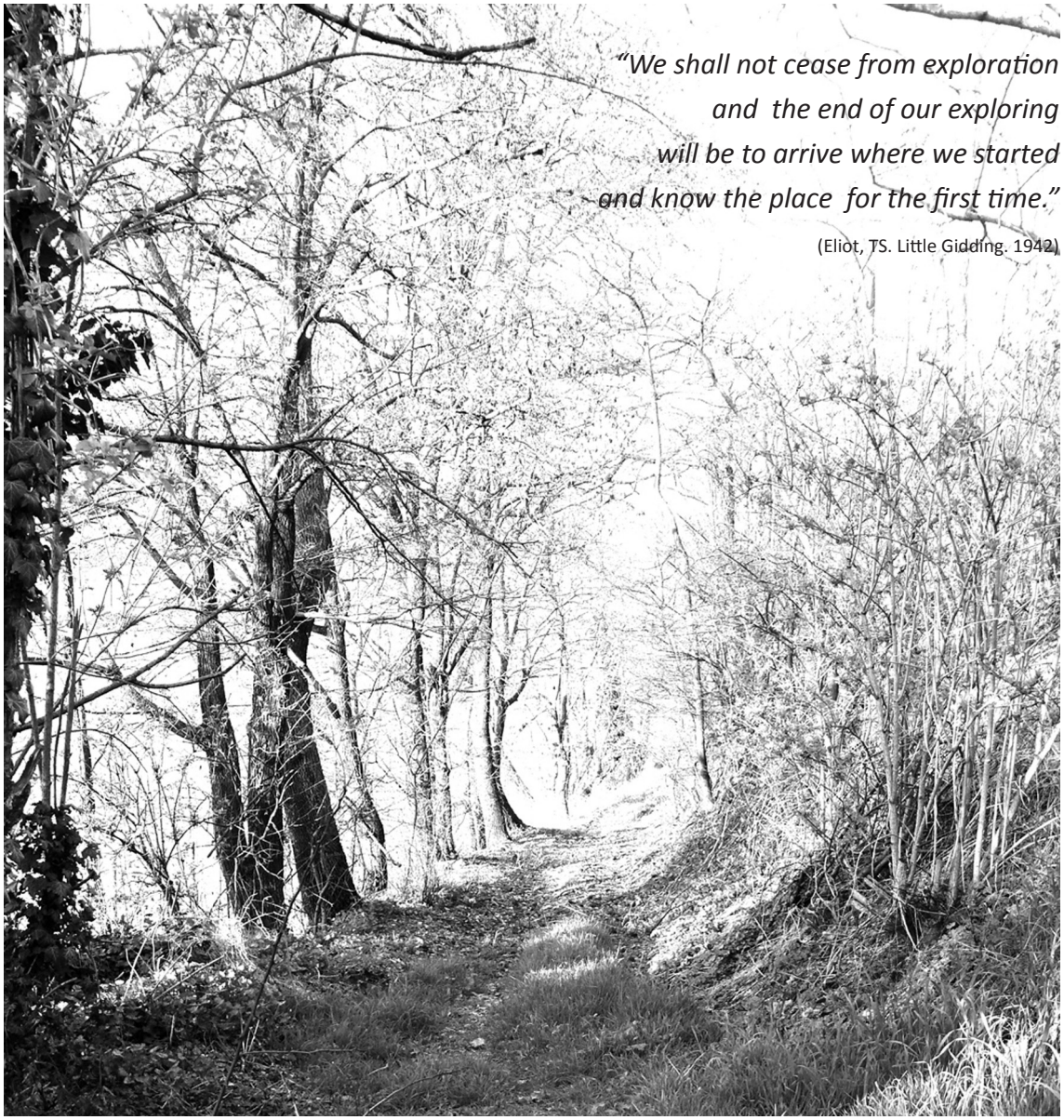
Tesi di Dottorato di Ricerca

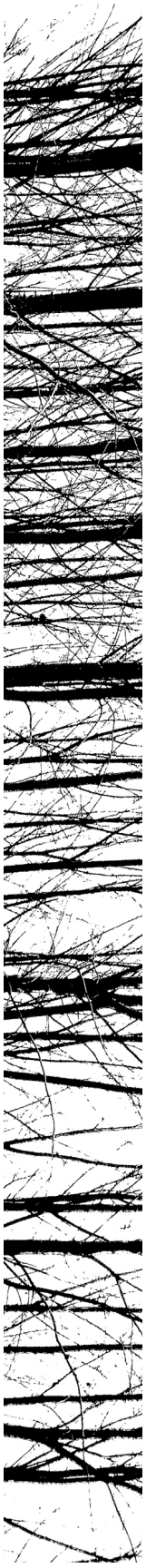
DIABETES MELLITUS: A COMPLEX METABOLIC DISORDER EXPLORING THE DISEASE THROUGH FRUCTOSAMINE 3-KINASE GENE ANALYSIS

Dott.ssa Francesca AVEMARIA
Matricola: R10187

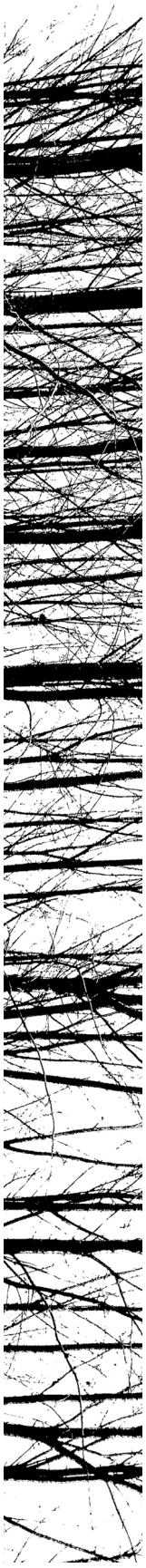
Tutor: Prof. Andrea MOSCA
Correlatore: Dott.ssa Paola CARRERA
Direttore: Prof. Sandro Sonnino
Coordinatore del Dottorato: Prof. Francesco BONOMI

Anno Accademico 2014-2015


"We shall not cease from exploration
and the end of our exploring
will be to arrive where we started
and know the place for the first time."
(Eliot, TS. Little Gidding. 1942)



INDEX



1. ABSTRACT

Diabetes Mellitus is the most prevalent metabolic disorder characterized by chronic hyperglycemia due to primary defects in insulin secretion and/or insulin function. In the last few decades the notion of diabetes has widened, ascertain that the present subdivision into type 1 (T1DM) and type 2 (T2DM) diabetes is a gross oversimplification. Both forms of diabetes seem to result from a complex interplay between genes and environment. Advance technologies have revolutionized the search for genetic influences on complex traits. Genomewide association studies (GWAS) and GWAS meta-analyses have been the most efficient way to identify new T2DM genes. However, despite these advances, the overall effect attributed to these loci is low and their contribution is of little clinical usefulness compared to evaluation of classical risk factors such as body mass index (BMI), age and family history. Therefore, evaluation of glycemic control remain the primary target for diabetes treatments. However, understanding and utilization of single gene effects on specific traits that conglomerate into a complex phenotype is currently the best way to understand the genetic basis at functional level and could results in a possible advantage for disease prevention and management.

In this light, to explore the complexity of this disorder and to give some hints in the pathogenesis of diabetes, we have followed a candidate gene approach by studying the Fructosamine 3-kinase (FN3K) gene, whose product is implicated in non-enzymatic glycation of proteins, in an Italian cohort of diabetic individuals.
Glycation has long been considered irreversible. Thus, the identification of an enzyme, FN3K, able to reverse this process by decomposing fructosamine 3-phosphate to 3-deoxyglucosone, inorganic phosphate and an amine, opened the perspective that fructosamines could be physiologically removed by proteins, suggesting a protective role in the development of diabetic complications and other pathologies characterized by high fructosamines/AGEs levels.

First aim of the present study was to accomplish the analysis of the FN3K gene in a well clinically characterized group of Italian individuals with diabetes (35 T1DM and 35 T2DM) belonging to ADAG study and 33 healthy subjects, by analyzing its promoter region. Then, the FN3K gene (promoter region and all six exons with corresponding intron/exon boundaries) were analyzed in additional 80 T2DM subjects, followed since long time in diabetic clinic.

The molecular screening revealed the presence of 15 different genetic variants. Four of them represented new mutations: the c.2 T>A (p.M1?) in the translation starting codon; the c.465 G>A (p.P115=) located in a consensus sequence for the splicing site; the c.559 C>T (p.R187*) leading to the formation of a truncated protein; the missense mutation c.716 A>G (p.Y239C) in exon 6. Presence of these variants were excluded from control group using DHPLC analysis. Other 11 variants identified were polymorphisms; of them 3 were new: the IVS2-27 A>G in intron 2; the c.421 C>T and c.429delATCGGAG in the promoter region. The remaining 8 were polymorphisms already described: c. -232 A>T and c.-385 A>G in the promoter region; c.187 A>C (p.R63=) in exon 2; c.900 C>G (p.S300=) and c.906 C>T (p.G302=) in exon 6; polymorphisms IVS+26 G>A, IVS+31 A>T and IVS4-9-11delTTG were present in non coding regions of the FN3K gene (intron 2, intron 2 and intron 4, respectively).

An RNA expression study was performed on new variants (c.2 T>A; c.465 G>A; c.559 C>T and c.716 A>G) in order to confirm the hypothesis on pathogenicity of these variants found at genomic level. Furthermore, mutations c.559 C>T (p.R187*) and c.716 A>G (p.Y239C) were analyzed in a familial context. The RNA analysis didn't confirm the result predicted by *in silico* modelling, suggesting that these variants might exert a pathological function with other mechanisms or interacting with other polymorphisms present in non coding regions of the FN3K gene where they may produce a subtle difference in regulation expression.

A second aim was to find a correlation between genotypes and some clinical parameters typical of diabetes, to offer a better comprehension of glycemic control and its predictive role in the middle and log period. A genotype composed by the combination of 3 polymorphisms (c.-385 A>G; c.-232 A>T; c.900 C>G) mostly associated with a variation in FN3K enzymatic activity and with HbA1c values was considered. On the bases of glycated hemoglobin values of each patients, the identified genotypes were compared on HbA1c mean values for the different groups. The analysis didn't find out a difference in HbA1c values among different groups. However, two interesting observations came out: the genotype containing the favorite alleles for the 3 polymorphisms (GG at c.-385 A>G; TT at c.-232 A>T; CC at c.900 C>G) seemed to be related to a low concentration of HbA1c; patients with different complications (macro- or microvascular) displayed different genotypes.

The role of FN3K variants in development or progression of diabetes remain unclear. The lack of an association between polymorphisms identified in the FN3K gene and diabetes explains that probably this enzyme cannot by itself account for the entire deglycation story.

It alone cannot be responsible for the susceptibility to this pathology or for the development of its complications.

Moreover, recent studies provided experimental evidences for epigenetic mechanisms as plausible means by which environmental factors integrate their effects with genetic variants to mediate T2DM risk. Current research provides insights for the importance of nutrition in terms of health and disease prevention. Studies on the "epigenetic diet" have revealed that the consumption of some foods like soy, curry spices, red grapes, as well as blueberries has beneficial effects on the prevention of diseases. One way of exerting its impact is through modulation of non-coding RNA levels, microRNAs (miRNAs), short (≈22 nucleotides) non-coding regulatory elements functioning as translational repressors. miRNAs have an important role in many cellular processes and are proposed as promising pharmaceutical targets in various fields, such as cancer and metabolic diseases. Recent findings have shown that altered circulating miRNA profiles were linked to pathological conditions, thus raising the possibility of their use as promising non-invasive biomarkers for the detection, classification and prognosis of diseases.

We had then the opportunity to collaborate with the University of Hull (UK), involved in understanding the effect of soy phytoestrogen in the management of diabetes. The potential health benefits of soy are widely publicized. These legumes contain complex carbohydrates,vegetable protein, soluble fibers, oligosaccharides, minerals, and phytoestrogens, particularly the isoflavones genistein and daidzein.

Two hundred men with T2DM and compensated hypogonadism were randomized and administered either 30 g of soy protein with 66 mg of isoflavones per day, or 30 g soy protein alone without isoflavones for 12 weeks. Given the estrogenic effect of phytoestrogens, the hypothesis to test was if there were any effects of soy with and without isoflavones on testosterone, this would be exaggerated in men with low testosterone levels.

The primary outcome of this study was the change in testosterone levels. The secondary outcomes were changes in glycemic control and cardiovascular risk markers including insulin resistance, lipid profile, highly sensitive CRP (hsCRP) and endothelial function. miRNA expression profile in peripheral plasma samples of 10 selected subjects under active treatment with soy and isoflavons, was performed to evaluate a possible change in expression level after treatment. Fifty-seven circulating miRNAs differed

for about 2 fold between pre- and post-treatment, including 7 up-regulated and 50 suppressed miRNAs. Among them miR-34a-5p, miR-144-3p and miR-19b-3p differed significantly between pre- and post- treatment.

In this study, after soy supplementation changes in testosterone levels, in glycemic control and cardiovascular risk markers were observed. Moreover, the number of expressed miRNA seemed to increase after treatment. A consistent and truly significant effect of soy on the miRNAs plasma profile could not be demonstrated, since a validation study is needed to confirm these findings and proceed with further evaluation. However, these preliminary results are encouraging for future investigations on soy isoflavones potential and for miRNA utilization as biomarkers.

In conclusion, our findings provide new insights in FN3K genetics and give further attempts on the hypothesis that phenotype might be influenced by allelic heterogeneity and/or mutations at multiple modifier genes. It would be interesting to analyze the FN3K gene in subjects belonging to the soy and testosterone study to evaluate a possible concomitant effect of FN3K genetic variants on the progression of the disease.

2. INTRODUCTION

2.1 Diabetes Mellitus

Diabetes is a group of metabolic diseases characterized by hyperglycemia resulting from defects in insulin secretion, insulin action, or both. The chronic hyperglycemia of diabetes is associated with long-term damage, dysfunction, and failure of different organs, especially the eyes, kidneys, nerves, heart, and blood vessels [1].

Diabetes mellitus has reached pandemic proportion, increasing in incidence by 50% over the past 10 years [2]. The World Health Organization (WHO) states that 347 million people worldwide were suffering from diabetes in 2008, which equates to 9.5% of the adult population. The incidence of diabetes is increasing at an alarming rate with estimations suggesting that this number will almost double by 2030 [2]. Diabetes mellitus occurs throughout the world but is more common in developed countries, due to rapid rise in obesity and life style changes. The greatest increase in prevalence in the near future, however, is expected to occur in Asia, the Middle East [3] and Africa, where it is likely that there will be a 50% increase in diabetes by 2030 [4].

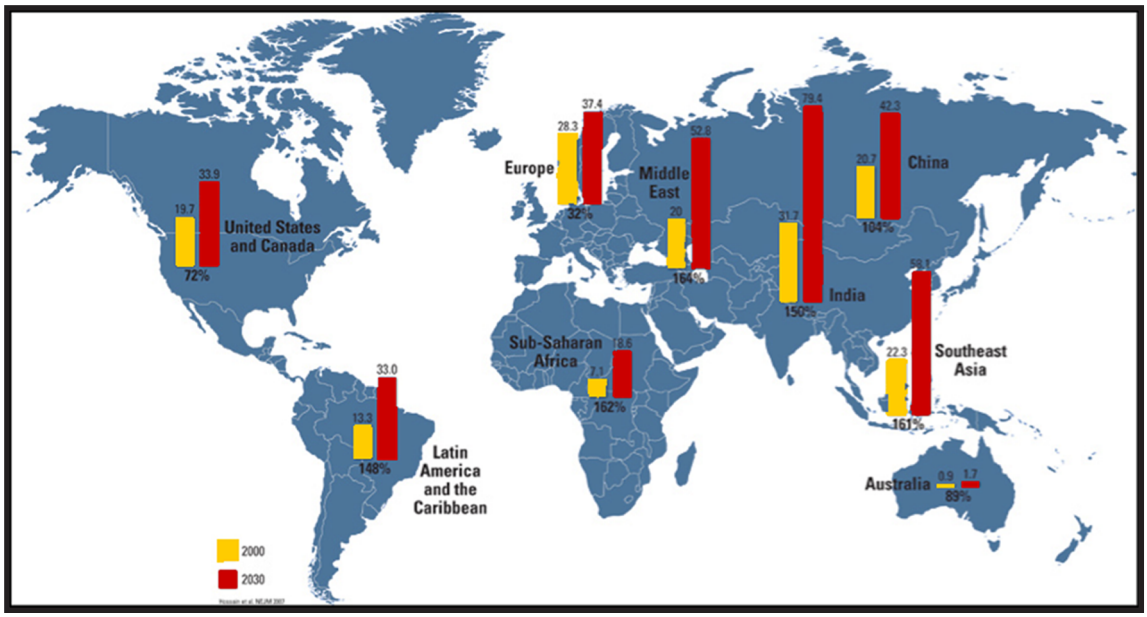


Figure 1. Prevalence of diabetes in 2010 and future estimation for 2030. Adapted from ref. [5].

Traditionally diabetes has been subdivided into two broad etiopathogenic categories: type 1 diabetes (T1DM), also known as insulin-dependent diabetes or juvenile-onset diabetes, which comprises only 5-10% of those with diabetes [1][6]; and type 2 diabetes (T2DM), the much more prevalent category, with 90-95% of diabetic cases worldwide [7]. In T1DM the lack of insulin secretion due to auto-immune mediated destruction of β cells causes hyperglycemia [1]. Autoimmune destruction of pancreas β-cells has multiple genetic predispositions and effect of environmental factors is still being understood. However, markers of immune destruction like islet cell auto-antibodies, auto-antibodies to insulin, GAD65, and tyrosine phosphatases IA-2 and IA-2b and rise in circulating levels of C-peptide help in reasonably accurate diagnosis of T1DM and determine their clinical course [1][6]. On the other hand, in T2DM insulin resistance and inadequate insulin secretory response result in raised circulating glucose levels [1].T2DM is characterized by resistance to insulin action in tissues like liver, skeletal muscles and adipose tissue, which leads to other features such as dyslipidemia and central obesity. Although the specific etiologies are not known, autoimmune destruction of β-cells does not occur [8].Impaired fasting glucose (IFG) or impaired glucose tolerance (IGT) can provide indications of derangements in glucose metabolism. Patients spend a long asymptomatic period characterized by hyperglycemia, which is undetectable but sufficient to cause pathological changes [1].

However, this subdivision is a gross oversimplification, and poorly describes the true range of such disease.T1DM and T2DM are polygenic in nature and caused by interactions between genetic and environmental factors. Monogenic diabetes, on the other hand represents rare, heterogeneous group of disorders due to genetic defects in single genes causing pancreatic β cell dysfunction and marked hyperglycemia [8]. Around 2–5% of global cases correspond to monogenic diabetes. Maturity onset diabetes of the young (MODY) and neonatal diabetes mellitus (NDM) represent two different classes of monogenic diabetes where hyperglycemia is either due to defects in insulin secretion, decrease in β cell mass or both [9].

In addition of these categories, diabetes may also manifest during pregnancy (gestational diabetes, GDM) and under other conditions including drug or chemical toxicity, genetic disorders, endocrinopathies, insulin receptor disorders and in association with pancreatic exocrine disease (Fig. 2) [10].

I. **Type 1 diabetes (b-cell destruction, usually leading to absolute insulin deficiency)**

- A. Immune mediated
- B. Idiopathic

II. **Type 2 diabetes (may range from predominantly insulin resistance with relative insulin deficiency to a predominantly secretory defect with insulin resistance)**

III. **Other specific types**

A. **Genetic defects of b-cell function**
1. MODY 3 (Chromosome 12, HNF-1a)
2. MODY 1 (Chromosome 20, HNF-4a)
3. MODY 2 (Chromosome 7, glucokinase)
4. Other very rare forms of MODY (e.g., MODY 4: Chromosome 13, insulin promoter factor-1; MODY 6: Chromosome 2, NeuroD1; MODY 7: Chromosome 9, carboxyl ester lipase)
5. Transient neonatal diabetes (most commonly ZAC/HYAMI imprinting defect on 6q24)
6. Permanent neonatal diabetes (most commonly KCNJ11 gene encoding Kir6.2 subunit of β-cell KATP channel)
7. Mitochondrial DNA
8. Others

B. **Genetic defects in insulin action**
1. Type A insulin resistance
2. Leprechaunism
3. Rabson-Mendenhall syndrome
4. Lipoatrophic diabetes
5. Others

C. **Diseases of the exocrine pancreas**
1. Pancreatitis
2. Trauma/pancreatectomy
3. Neoplasia
4. Cystic fibrosis
5. Hemochromatosis
6. Fibrocalculous pancreatopathy
7. Others

D. **Endocrinopathies**
1. Acromegaly
2. Cushing's syndrome
3. Glucagonoma
4. Pheochromocytoma
5. Hyperthyroidism
6. Somatostatinoma
7. Aldosteronoma
8. Others

E. **Drug or chemical induced**
1. Vacor
2. Pentamidine
3. Nicotinic acid
4. Glucocorticoids
5. Thyroid hormone
6. Diazoxide
7. b-Adrenergic agonists
8. Thiazides
9. Dilantin
10. g-Interferon
11. Others

F. **Infections**
1. Congenital rubella
2. Cytomegalovirus
3. Others

G. **Uncommon forms of immune-mediated diabetes**
1. "Stiff-man" syndrome
2. Anti-insulin receptor antibodies
3. Others

H. **Other genetic syndromes sometimes associated with diabetes**
1. Down syndrome
2. Klinefelter syndrome
3. Turner syndrome
4. Wolfram syndrome
5. Friedreich ataxia
6. Huntington chorea
7. Laurence-Moon-Biedl syndrome
8. Myotonic dystrophy
9. Porphyria
10. Prader-Willi syndrome
11. Others

IV. **Gestational diabetes mellitus**

Figure 2. Etiologic classification of diabetes mellitus. Adapted from ref. [10].

2.2 Genetics of T2DM

Several lines of evidence support the principle of inherited genetic susceptibility as an important risk factor for T2DM. There is a strong familial aggregation, with a risk of developing the disease of 40% for those who have an affected parent (higher if the mother rather than the father) and of 70% if both parents are diabetics [11]. Moreover, in identical monozygotic twins the concordance rate approaches 100%, much higher than that seen in dizygotic twins or among siblings [12]. Genetic predisposition in T2DM is also supported by the observation that differences in disease prevalence rates exist among populations, even after migration of entire ethnic groups to another country, thus independent from the environmental influence [13].

Enormous efforts have been done to identify the genetic cause of the disease. In the late 1990s and early 2000s family linkage approach, following the hypothesis that complex diseases such as T2DM may harbor major genetic mutations that are severe enough to cause disease under specific exposures, have been utilized to evaluate for co-segregation of genetic markers in multiplex families with T2DM. Although much effort invested, the only major success was the identification of the transcription factor 7-like 2 gene (TCF7L2), while most of the other putative candidates remained to be validated [14].

In more recent years, advances in technology for SNP genotyping, exploitations of recent genetic knowledge gained from the Human Genome Project and development of robust statistical methods have allowed genome-wide association studies (GWAS) to emerge as the method of choice for detecting common genetic variants associated with complex diseases. This methodology has helped to identify dozens of new associations between T2DM and genes with known or unknown functions and some of previous results have been confirmed [15]. To date, about 80 distinct genetic loci have been identified as predisposing to risks of T2DM in European populations through case-control studies and these are summarized in Table 1. Moreover, multiple other GWAS efforts have also been invested in evaluating for related quantitative traits, such as blood glucose and insulin levels as well as T2DM in non-European populations [16][17].

Despite this dramatic success, there is a substantial gap between the discovery of many T2DM-associated SNPs and the understanding of how these SNPs physiologically impact T2DM pathogenesis. So far, 80% of all T2DM-associated SNPs are intergenic or intronic, and the geneticists have named the closest gene on the SNP chromosome as the T2DM-susceptibility gene without having strong clues about the molecular link between the gene and the variant. Furthermore, the individual effect of these SNPs

on T2DM risk is typically modest. Each risk allele usually increases less than 15% the risk of developing T2DM (Odds Ratio <1.15). As a consequence, we can see that all common T2DM-associated SNPs have captured less than 15% of T2DM familial heritability [18].

However, in case in which was possible to ascertain the precise pathogenic mechanism, it has been observed that most of the genes identified are involved in pancreatic β-cell mass and/or function, with a consequent implication in insulin defects [19][20].

The advent of GWAS era and the identification of multiple risk loci have illuminated the pathophysiology of T2DM and have of course confirmed the polygenic nature of the disease. Furthermore, the identification of these loci has also provided an opportunity to translate genetic information into clinical practice. For example, these may have potential role in disease risk prediction or in identifying subjects at risk of developing disease at an early-stage or in clinical management of individuals. Affected individuals could take advantages from the so called legacy effect, i.e., early tight diabetes control resulting in a substantial micro- and cardiovascular benefit [21].

Location	Reported Gene(s)	SNP	Risk Allele Frequency	OR (95% CI)	Location	Reported Gene(s)	SNP	Risk Allele Frequency	OR (95% CI)
10q25.2	TCF7L2	rs7903146	0.30	1.40 (1.35–1.46)	15q24.3	HMG20A	rs7178572	0.70	1.08 (1.04–1.13)
9q21.31	TLE4	rs17791513	0.93	1.21 (1.13–1.31)	11p15.1	KCNJ11	rs5215	0.38	1.08 (1.04–1.12)
6p22.3	CDKAL1	rs7756992	0.26	1.20 (1.16–1.25)	1q32.3	PROX1	rs2075423	0.66	1.08 (1.04–1.12)
9p21.3	CDKN2A, CDKN2B	rs10811661	0.82	1.18 (1.13–1.24)	8q22.1	TP53INP1	rs7845219	0.53	1.08 (1.04–1.12)
8q24.11	SLC30A8	rs3802177	0.70	1.16 (1.11–1.22)	18q21.32	MC4R	rs12970134	0.27	1.08 (1.03–1.12)
12q14.3	HMGA2	rs2261181	0.09	1.16 (1.10–1.23)	2p25.3	TMEM18	rs10190052	0.88	1.07 (1.04–1.10)
3p25.2	PPARG	rs1801282	0.88	1.16 (1.10–1.23)	10q22.1	C10orf35	rs2812533	0.83	1.07 (1.04–1.09)
10q23.33	HHEX, IDE	rs1111875	0.58	1.15 (1.11–1.19)	3q27.3	LPP	rs6808574	0.60	1.07 (1.04–1.09)
2p21	THADA	rs10203174	0.90	1.15 (1.08–1.21)	6p21.33	POU5F1, TCF19	rs3132524	0.74	1.07 (1.04–1.09)
16q12.2	FTO	rs9936385	0.39	1.13 (1.09–1.18)	9q21.32	TLE1	rs2796441	0.63	1.07 (1.03–1.12)
3q27.2	IGF2BP2	rs4402960	0.31	1.13 (1.09–1.17)	20q13.12	HNF4A	rs4812829	0.16	1.07 (1.01–1.12)
11q13.4	ARAP1, CENTD2	rs1552224	0.83	1.13 (1.08–1.19)	5q11.2	ARL15	rs702634	0.71	1.06 (1.04–1.09)
6p21.2	KCNK16	rs1535500	0.59	1.13 (1.08–1.19)	8q24.21	TMEM75	rs1561927	0.23	1.06 (1.04–1.09)
7p21.2	DGKB	rs17168486	0.19	1.13 (1.07–1.19)	12q24.31	MPHOSPH9	rs1727313	0.24	1.06 (1.04–1.08)
5q13.3	ZBED3	rs6878122	0.25	1.13 (1.07–1.18)	13q12.13	RNF6	rs10507349	0.74	1.06 (1.04–1.08)
17q12	HNF1B	rs4430796	0.53	1.13 (1.07–1.09)	6p21.1	VEGFA	rs9472138	0.24	1.06 (1.04–1.08)
7p15.1	JAZF1	rs849135	0.52	1.12 (1.08–1.17)	6p24.3	SSR1, RREB1	rs9502570	0.30	1.06 (1.04–1.08)
12q24.31	HNF1A	rs12427353	0.77	1.12 (1.07–1.18)	10q23.31	PTEN	rs10788575	0.16	1.06 (1.03–1.08)
11q14.3	MTNR1B	rs10830963	0.27	1.11 (1.06–1.16)	15q22.2	C2CD4A	rs7163757	0.56	1.06 (1.02–1.11)
7q32.3	KLF14	rs13233731	0.49	1.10 (1.06–1.13)	19q13.32	GIPR	rs8108269	0.30	1.06 (1.02–1.11)
1p12	NOTCH2	rs10923931	0.11	1.10 (1.04–1.17)	10p13	CDC123	rs11257655	0.23	1.06 (1.01–1.11)
1p32.3	FAF1	rs17106184	0.92	1.10 (1.07–1.14)	7p14.3	CRHR2	rs2284219	0.32	1.05 (1.03–1.08)
8p11.21	ANK1	rs516946	0.77	1.10 (1.06–1.15)	10q26.13	PLEKHA1	rs10510110	0.41	1.05 (1.03–1.07)
13q31.1	SPRY2	rs1359790	0.73	1.10 (1.05–1.14)	1q41	LYPLAL1	rs2820446	0.71	1.05 (1.03–1.07)
3p24.3	UBE2E2	rs7612463	0.87	1.10 (1.04–1.16)	5q31.1	PCBD2	rs319598	0.53	1.05 (1.03–1.07)
10q22.3	ZMIZ1	rs12571751	0.51	1.09 (1.06–1.13)	6q22.32	C6orf173	rs4273712	0.25	1.05 (1.03–1.07)
4p16.1	WFS1	rs4458523	0.59	1.09 (1.06–1.13)	7p21.2	ETV1	rs7795991	0.54	1.05 (1.03–1.07)
3q21.1	ADCY5	rs11717195	0.78	1.09 (1.05–1.14)	9p24.2	GLIS3	rs7041847	0.50	1.05 (1.01–1.09)
12q21.1	TSPAN8	rs7955901	0.47	1.09 (1.05–1.13)	6q23.3	IL20RA	rs6937795	0.42	1.04 (1.02–1.06)
15q25.1	ZFAND6	rs11634397	0.64	1.09 (1.05–1.13)	15q26.1	AP3S2	rs2028299	0.29	1.04 (1.00–1.09)
2q36.3	IRS1	rs2943640	0.63	1.09 (1.05–1.13)	2q24.3	GRB14	rs3923113	0.61	1.04 (1.00–1.09)
11p15.4	KCNQ1	rs163184	0.50	1.09 (1.04–1.13)	3p14.1	PSMD6	rs831571	0.81	1.03 (0.99–1.08)
12p11.22	KLHDC5	rs10842994	0.80	1.09 (1.04–1.13)	3q27.3	ST64GAL1	rs16861329	0.85	1.03 (0.96–1.10)
15q26.1	PRC1	rs12899811	0.30	1.09 (1.04–1.13)	10q22.1	VPS26A	rs1802295	0.33	1.02 (0.98–1.06)
2p16.1	BCL11A	rs243088	0.46	1.09 (1.04–1.13)	15q14	RASGRP1	rs7403531	0.22	1.02 (0.98–1.06)
4q31.3	TMEM154	rs6813195	0.72	1.08 (1.06–1.10)	19q13.11	PEPD	rs3786897	0.57	1.02 (0.98–1.06)

Table 1. List of identified common variants associated with T2DM in populations of European ancestry utilizing a case-control design. Adapted from ref. [22].

2.3 Diagnostic criteria for diabetes mellitus

Assigning a type of diabetes to an individual often depends on the circumstances present at the time of diagnosis, with individuals not necessarily fitting clearly into a single category. Clinical presentation and disease progression may vary considerably in both type of diabetes.

The WHO diagnostic criteria for diabetes mellitus include fasting plasma glucose level ≥126 mg/dL (7.0 mmol/L) or 2-hour plasma glucose level ≥200 mg/dL (11.1 mmol/L) after a 75-g oral glucose load [23][24]. More recently, a glycated hemoglobin level of ≥6.5% is recommended by WHO as the cut point for diagnosing diabetes [25].

Moreover, there are some conditions of intermediate hyperglycemia associated with an increased risk to develop diabetes, collectively known as prediabetes [26]. Impaired glucose tolerance is defined as a 2-hour plasma glucose level ≥140 mg/dL (7.8 mmol/L) and less than 200 mg/dL (11.1 mmol/L) after a 75-g oral glucose load. Impaired fasting glucose is defined as fasting glucose level between 110 mg/dL and 125 mg/dL (6.1–6.9 mmol/L). Glycated hemoglobin value between 5.7% and 6.4% is also associated with prediabetes.

Diagnostic criteria for diabetes are summarized in Figure 3.

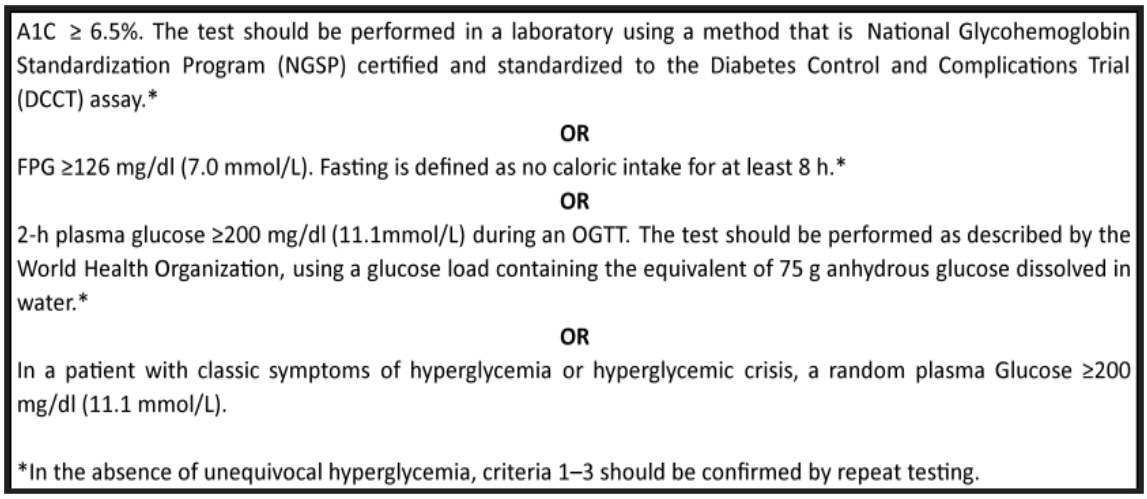


Figure 3. Criteria for the diagnosis of diabetes. Adapted from ref. [1]

2.4 Diabetic complications: the “non-enzymatic glycation hypothesis”

Diabetes is a common condition with multiple complications. Symptoms of marked hyperglycemia include polyuria, polydipsia, weight loss, sometimes with polyphagia, and blurred vision. Impairment of growth and susceptibility to certain infections may also accompany chronic hyperglycemia [27][28]. In human diabetes, hyperglycemia leads to long-term complications such as retinopathy with potential loss of vision; nephropathy leading to renal failure; peripheral neuropathy with risk of foot ulcers, amputations, Charcot joints; and autonomic neuropathy causing gastrointestinal, genitourinary, and cardiovascular symptoms and sexual dysfunction. Patients with diabetes have an increased incidence of atherosclerotic cardiovascular, peripheral arterial and cerebrovascular disease. Moreover, hypertension and abnormalities of lipoprotein metabolism are often found in people with diabetes [27][28].

To date, there has been much work done to elaborate on the etiology, prevention and treatment of diabetes related complications. The DCCT [29] and UKPDS [30] studies have emphasized the role of tight glucose control as being important in reducing diabetes microvascular complications in T1DM (DCCT) and T2DM (UKPDS). On the other hand, the relation between tight glucose control and macrovascular complications is not yet clear [31].

Many hypotheses have been proposed to link the elevated concentration of glucose and the development of related complications [32][33][34]. One of the leading proposals is the ‘‘non-enzymatic glycation hypothesis’’ which postulates that the deleterious effects of chronic hyperglycemia are a result of excessive non-enzymatic modification of proteins and some phospholipids by glucose and its by-products [34][35].

Glycation, also known as Maillard reaction, is a common and spontaneous reaction, occurring in vivo [36]. Protein glycation refers to the binding of glucose or other reducing sugars to proteins. Particularly, glycation with aldoses takes place with the ε-amino group of lysine residues or with the N-terminus of the protein, resulting in the formation of a Schiff base product. This thermodynamically-unstable compound rearranges to form ketoamine derivatives, the Amadori products, that can undergo a series of further rearrangements (dehydration, cyclization, fragmentation and oxidation) to form a wide and heterogeneous group of complex compounds called advanced glycation end products (AGEs) (Fig. 4) [37]. Both fructosamines and AGEs can impair structural and biological properties of proteins in living organisms, with a complex mechanism at least in part independent of hyperglycemia [38][39].

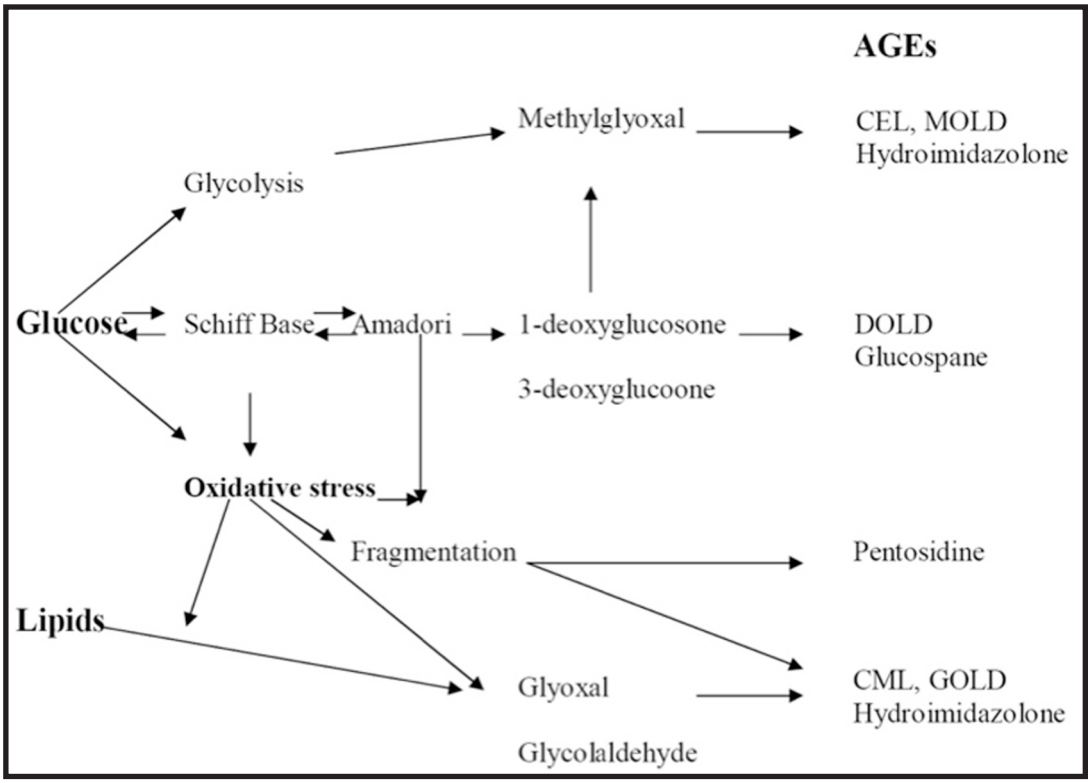


Figure 4. Schematic representation of Maillard reaction and some AGEs formation.
CEL= carboxyethyllysine; MOLD= methylglyoxal lysine dimer; DOLD= 3-deoxyglucosone lysine dimer; CML= carboxymethyllysine; GOLD= glyoxal lysine dimer. Adapted from ref. [40].

2.4.1 Advanced glycation end products and diabetes

AGE formation and accumulation increasingly occurs under diabetic conditions, and even if glycemic control is restored, AGEs can remain in tissues of diabetic subjects for a long time[41]. In people with diabetes, AGE formation is accelerated due to increased concentration of circulating glucose, AGE precursors and oxidative stress. In serum and tissues of patients with T1DM as well as T2DM, increased levels of AGEs have been found. Furthermore, accumulation of AGEs in diabetic tissue was shown to correlate with diabetic complications [42].

The damaging potential of AGEs results from direct alterations on protein structures and functions due to AGEs per se or the cross-linking effect of some AGEs [43]. AGEs are often found in the extracellular matrix and thus modified matrix proteins impair matrix-matrix as well as matrix cell interactions. This may cause cell death, cell differentiation or reduced cell adhesion and migration [44]. Intracellular proteins are also targets of modifications and AGE formation was shown to impair their functions [45]. Besides direct changes in protein structures and functions, AGE-mediated damage occurs via binding of AGEs to the receptor of advanced glycation end products (RAGE) [46]. RAGE belongs to the immunoglobulin superfamily and additionally interacts with a wide range of AGEs which is why RAGE is also classified as a multi-ligand receptor. In recent years the interaction of AGEs with RAGE was studied in vitro pointed out that ligand binding to RAGE activates NAPDH oxidases and thus increases intracellular ROS formation [47][48]. Increased ROS in turn leads to AGE formation, which triggers all described damaging mechanisms mediated by AGEs, but also activates the transcription factor nuclear factor kappa B (NFκB) [49]. Activation of NFκB increases the expression of pro-inflammatory cytokines such as interleukin 6 [50] and monocyte chemoattractant peptide 1 (MCP-1) [51] as well as RAGE itself [52] thus intensifying the inflammatory response.

Glycation of proteins is a major cause of spontaneous damage to the proteome. In recent years, the traditional view of glycation has been confirmed and new insights established. AGEs are particularly insidious since they play an important role in the pathogenesis of diabetic complications that are the principal cause of morbidity and mortality of diabetic patients. It is nevertheless critical to identify logical and directed targets for prevention and therapy. In future research it will be important to identify proteins and hotspot sites within them susceptible to glycation. Probably the most potent strategy will be to develop inducers of enzymatic defense against glycation [53] [54] such as FN3K.

2.5 Fructosamine 3-kinase

Repair is of outstanding importance for life and three different types of enzymes are known to metabolize ketoamines: oxidases, isomerases and kinases [55] [56]. However, only the latter is used to deglycate proteins under physiological conditions.

The discovery of Fructosamine 3-kinase (FN3K), a protein distantly related to aminoglycoside kinases and, even more distantly, to protein kinases [57], able to repair fructosamines, opened new conceptual avenues on Maillard reaction.

The starting point for the discovery of FN3K, was the identification of fructose 3-phosphate in human and animal tissues [58]. This enzyme exhibited a very low affinity for its substrate (Km≥ 30 mM) and a low metabolic capacity, suggesting that it could act on some substrates different from fructose, even if on compounds with a closely related structure [58]. Two independent groups led to the conclusion that "fructose 3-kinase" phosphorylates fructosamines with a Km in the micromolar range, i.e. 4–5 orders of magnitude lower than the Km for fructose (≈50–100 mM).

FN3K appears to represent a part of a cellular defense and/or repair system to control non enzymatic glycation of proteins. Indeed, the enzyme phosphorylates not only fructosamines, but also their C3-epimers psicosamines, as well as ribulosamines and erythrulosamines, even if psicosamines are much poorer substrates than the other ketoamines [59]. In details, FN3K would be able to break down the second intermediate of the non-enzymatic glycation cascade by phosphorylating fructoselysine to a fructoselysine-3-phosphate (FL3P). The latter compound spontaneously decomposes by β-elimination, regenerating an unmodified lysine along with inorganic phosphate and 3-deoxyglucosone, then readily detoxified to inert products such as 3-deoxyfructose or 2-deoxy-3-ketogluconic acid [57][59] (Fig. 5).

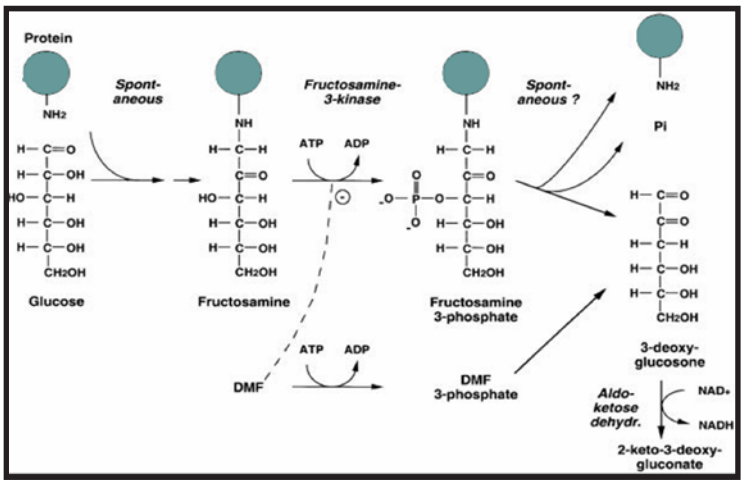


Figure 5. Glycation process and FN3K role in phosphorylation with spontaneous deglycation of the Amadori product. Adapted from ref. [60]

The enzyme was purified from human erythrocytes [59] and its cDNA was cloned [57].

Human FN3K is a monomeric protein of 309 amino acids encoded by a gene (NC_000017.10) located on chromosome 17q25.3. FN3K gene may have arisen by an event of duplication of an ancestral gene, FN3K-Related Protein (FN3K-RP). The gene encoding FN3K-RP is located 8 kb upstream of the FN3K gene, and share a 65% sequence homology with FN3K and an identical genome organization and same orientation on the chromosome (Fig. 6) [61][62].

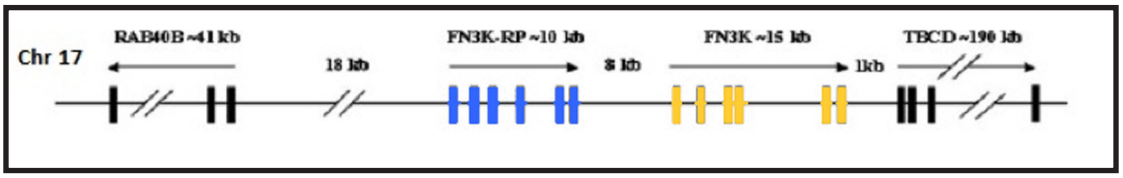


Figure 6. FN3K genomic organization. Adapted from ref. [62].

Both FN3K and FN3K-RP genes have 6 exons with the initial ATG in exon 1 and a STOP codon in exon 6. The intron-exon junctions are in homologous position and both genes are characterized by a large fourth intron. Analysis of the core promoters of these genes showed two similar sequences without TATA or CCAAT boxes and high in GC. Such promoter sequences are consistent with those of housekeeping genes, and suggest they are likely to be expressed in a large variety of tissues [63]. mRNA for both genes was detected, with especially high levels of FN3K, in tissues susceptible to diabetic complications such as the kidney, heart and peripheral and central nervous systems [63]. Furthermore, starvation and diabetes do not change the level of expression of FN3K in different tissues, and no regulation of FN3K expression was observed in human fibroblasts treated with condition mimicking the diabetic state, strengthening the idea of an housekeeping gene (Fig. 7) [63].

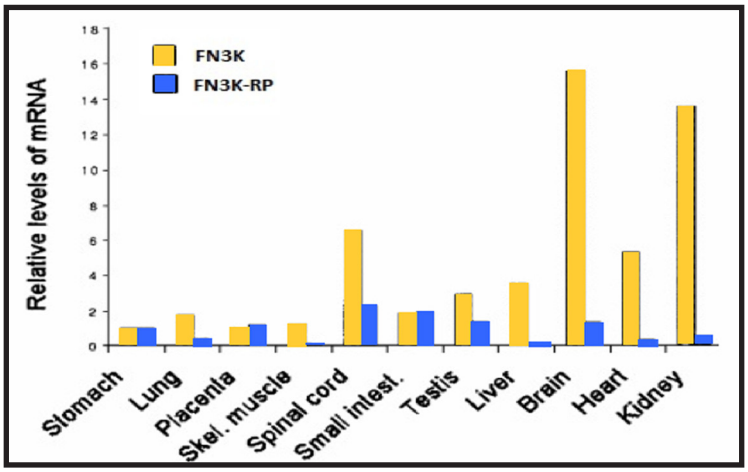


Figure 7. Relative levels of FN3K and FN3K-RP mRNA in human tissues (normalized to β-actin message) as assessed by quantitative RT-PCR. Adapted from ref. [63].

Both FN3K and FN3K-RP phosphorylate psicosamines and ribulosamines, but only the former act on fructosamines [61].

FN3K involvement in deglycation was first provided by finding that its competitive inhibitor, deoxymorpholino-fructose (DMF), increases about 2-fold the rate of accumulation of glycated hemoglobin when erythrocytes are incubated in presence of 200 mM glucose [60]. Definitive evidence for FN3K being responsible for deglycation was provided in animal models: FN3K-/- mice showed a level of hemoglobin-bound fructosamines of about 2.5-fold higher than those observed in FN3K+/+ or FN3K+/- mice [64]. However, FN3K has been recently demonstrated to be able to reduce the glycation of intracellular islet proteins, but does not affect pancreatic β-cell survival and function, even if these are incubated for several weeks in presence of high glucose concentration [65].

The demonstration that FN3K phosphorylates glycated hemoglobin in intact cells causing its partial deglycation was followed by experiments aimed to identify the fructosamines residues removed from hemoglobin in intact erythrocytes, as a result of FN3K action. In vitro studies indicated that several fructosamines bound to lysines are excellent substrates, whereas others are only poorly phosphorylated [66]. Thus, the fructosamines bound to Lys139α, located near the C-terminus of α subunits, and Lys16α, located on a loop of α subunits, are good substrates. On the contrary, fructose bound to Lys61α, whose side chain is partially bound to a heme, is only very slowly phosphorylated. Moreover, the N-terminal glycated valine is a poor substrate, consistent with free fructosevaline being a much poorer substrate than free fructoselysine [67].

The physiological importance of deglycation is unknown at present. Deglycation might prevent the further conversion of fructosamines into advanced glycation products. FN3K would be expected to act on glycated residues that are most accessible and that would therefore be more likely to cause covalent cross-linking of proteins. Another potential role is to prevent the loss of residues that play a critical role in enzyme catalysis or protein function [66]. However, enzymes sometimes undergo conformational changes during catalysis. A glycated lysine might therefore become accessible in an open conformation of an enzyme. Moreover, FN3K could prevent the loss of residues playing a role in protein/protein interactions, that are expected to be rather freely accessible to FN3K [66].

2.5.1 FN3K variability in humans

The pathogenesis of diabetes is incredibly complicated as depicted by the number of pathways implicated. Diligent control of glycemia and blood pressure stabilize the level of morbidity and mortality associated with diabetes [68]. However, good glycemic control alone has been shown to be insufficient to prevent death from cardiovascular event [69] and may increase risk of adverse events [70].

The human erythrocyte FN3K activity, which is stable with time in a single individual, is highly variable from person to person. This variability is apparently not due to differences in blood glucose levels or insulinemia, in addition FN3K activity is not different in erythrocytes from T1DM and normoglycemic subjects [71]. It was reported that the real role of FN3K enzyme in the development of diabetic complications could be demonstrated by comparing the development of complications in patients with low and high FN3K activities on a large cohort of diabetic patients categorized according to known factors predisposing such complication [71]. Different approaches for the measurement of FN3K activity were described. First of all, FN3K has to be partially purified from the erythrocyte lysate in order to determine its activity in erythrocytes [59]. Phosphorylated products are then detected using 31P nuclear magnetic resonance (NMR) spectroscopy. A second strategy is based on the use of the substrate fructose-1,4-(3-aminoethyl)-piperazine and the radiolabeled cosubstrate [γ-33P]-ATP. Quantification of the 33P-labelled phosphorylated product is performed by utilising a scintillation technique after removal of the excess of cosubstrate. In a third approach the radiolabelled substrate 14C-deoxymorpholinofructose (14C-DMF) is used. For the scintillation analysis the phosphorylated product 14C-DMF is isolated using an anion-exchange column. In this case, activity studies are performed using intact erythrocytes or partially purified FN3K [72]. However, these methods are time consuming and inappropriate for routine analysis on large series of subjects, for the use of radioactive substrates and/or NMR equipment. Moreover these methods seem to be deficient in terms of analytical performance for the current clinical use.

A more realistic approach in understanding the FN3K possible role in diabetic complications would be to analyze the presence of different SNPs in the FN3K gene, previously linked to FN3K enzymatic activity [71].

First, Delpierre and collaborators reported an association between FN3K enzymatic activity in red cells and some polymorphisms in the FN3K gene, in a Belgian cohort of 31 T1DM subjects and 26 controls [71]. They found that two SNPs, besides other gene variants, the CC of the c.900C>G (rs1056534) in exon 6 and the GG of the c.-385A>G (rs3859206) in the promoter region, were associated with reduced

enzymatic activity measured in erythrocytes. However, they failed to detect a correlation between FN3K SNPs and HbA1c levels [71]. Then, the group of Mohás analyzed a large cohort of T2DM subjects (859 T2DM and 265 controls) for the presence of the polymorphism c.900C>G (rs1056534) of the FN3K gene [73]. They found that the C allele of rs1056534 was coupled with lower HbA1c concentration and with a later onset of type 2 diabetes. However, no association between this variant and diabetic complications, such as nephropathy, neuropathy or retinopathy, was found in their investigation. In 2014, a group of 314 T2DM subjects was screened for 19 SNPs in six candidate genes encoding for enzymes of metabolic pathways, in order to verify if the genetic variability in such genes could influence the progression of diabetic nephropathy. An association of the polymorphism in exon 6 (rs1056534) of the FN3K gene with the progression of diabetic nephropathy and cardiovascular morbidity and mortality was indeed reported [74]. Recently, Škrha and co-workers, in a cohort of 129 T1DM, 340 T2DM and 126 controls, evaluated the association of FN3K and GLO1 polymorphisms with parameters of endothelial dysfunction and soluble receptor for AGEs (sRAGE) [75]. In 126 subjects (50 T1DM, 52 T2DM and 24 healthy individuals), a significant association of FN3K rs1056534 and rs3848403 SNPs with sRAGE concentration in patients with diabetes was proven. Our research group has also analyzed a Caucasian cohort of 70 diabetic patients, 35 T1DM and 35 T2DM and 33 controls, for the coding part of the FN3K gene, identifying two new mutations and additional variants within the gene. No significant association was found between certain SNPs and diabetic conditions. However, we noted too that the genotype containing c.900 CC alleles (rs1056534) seemed to be related with low concentration of HbA1c [76].

The interpretation of FN3K polymorphisms and their role in influencing the enzymatic activity is hard to evaluate. None of the identified polymorphisms resulted in a modification of the sequence of the encoded protein, indicating that the differences in activity are not due to the production of a hypo- or hyper-active protein. Furthermore, none of the polymorphisms were present in a consensus sequence for a splicing site or for the binding of a transcription factor, suggesting that they are not by themselves responsible for the variations in FN3K activity. A likely explanation is that they are in linkage disequilibrium with polymorphisms in regulatory sequences (enhancer element) [71].

2.6 MicroRNA

The most well-studied sequences in the human genome are those of protein-coding genes. However, the coding exons of these genes account for only 1.5% of the genome, a proportion that increases to 2% if untranslated regions (UTRs) are considered [77]. In recent years, it has become increasingly apparent that the non-protein-coding portion of the genome is of crucial functional importance for normal development and physiology and for disease [78]. The functional relevance of the non-protein-coding genome is particularly evident for a class of small non-coding RNAs (ncRNAs) called microRNAs (miRNAs) [79][80].

They were first described in 1993, in C.elegans as RNA molecules that regulate the developmental timing in the organism [81]; however, after less than two decades, thousands of miRNAs have been identified in plants and animals, including approximately 2,000 miRNAs in humans (miRBase) [82].

miRNAs are short (19–24 nucleotides in length) non-coding RNAs that regulate messenger RNA (mRNA) or protein levels either by promoting mRNA degradation or by attenuating protein translation. Based on computational prediction, it has been estimated that more than 60% of mammalian mRNAs are targeted by at least one miRNA [83].

Biogenesis of miRNAs starts in the nucleus, where long primary molecules (pri-miRNAs) are transcribed by RNA polymerase II [84], or in some cases by RNA polymerase III [85]. The primary transcript is processed by the nuclear RNase III-type enzyme Drosha into a shorter miRNAs precursors (pre-miRNAs) [86]. After the initial processing in the nucleus, pre-miRNAs are then exported to the cytoplasm by the complex of Exportin-5 [87]. A Dicer enzyme cleaves the molecules, which then form RNA-induced silencing complex (RISC). By attaching themselves to complementary sequences of the target RNA,the RISC complexes improve their stability and help in mRNA translation. miRNA inhibits protein synthesis by interacting with partially complementary regions near the 3’-end, which do not undergo translation. Upon binding, the miRNA initiates a pathway that either degrades the transcripts or suppresses their translation [88][89] (Fig. 8).

Mature miRNA regulated gene expression contributes to essential cellular processes such as differentiation, apoptosis and proliferation [91]. Quantitative detection of cellular miRNAs has been suggested to define disease status, as abnormal presence of certain miRNAs correlates with the pathogenesis of diseases such as cancer and diabetes [92].

Importantly, miRNAs can also be detected outside of cells [93][94]. This fraction of miRNAs is regarded a cell-free circulating molecules residing in various extracellular

vesicles such as microparticles, exosomes [95] and apoptotic bodies [96]; or conjugated with RNA binding proteins [97] or lipoprotein complexes [98] (Fig 8).

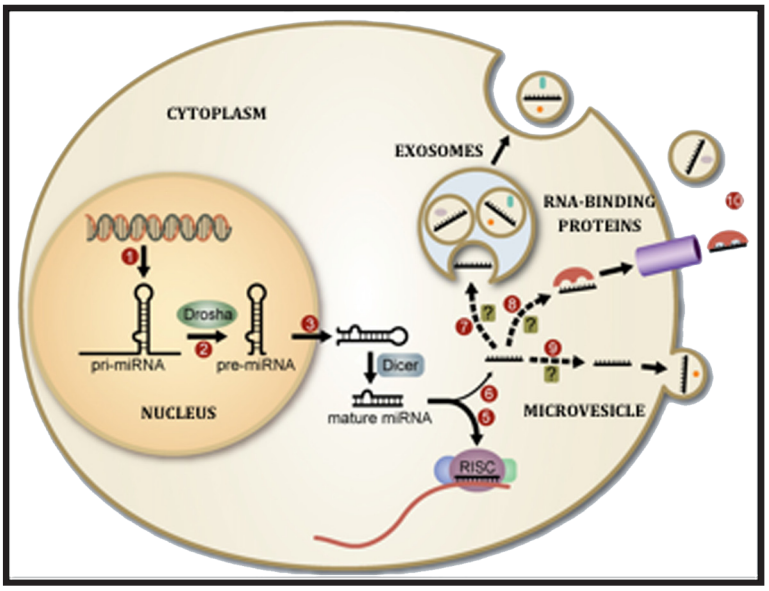


Figure 8. Overview of miRNAs biogenesis. miRNA biogenesis begins with transcription of pri-miRNA transcripts by RNA polymerase II or III (1). In the nucleus, pri-miRNAs are processed by Drosha to produce pre-miRNA hairpins (2), which are then exported into the cytosol (3). Here, pre-miRNA hairpins are processed into 19–24 nucleotide mature miRNA duplexes by Dicer (4). One strand of the mature miRNA duplex is incorporated into the RISC complex where it can regulate expression of target mRNAs (5). The other strand may either be degraded, or possibly prepared for export from the cell (6). Some miRNAs have been found packaged in exosomes derived from multivesicular bodies (7). Others may be exported in the presence of RNA-binding proteins (8). Still others might be exported microvesicles shed during membrane blebbing (9). Once in the extracellular space, these miRNAs could be taken up by other cells, degraded by RNases, or excreted (10). Adapted from ref. [90].

However, the function of these circulating extracellular miRNAs remains poorly understood. One of the most intriguing ideas is that extracellular miRNAs are used as mediators of cell–cell communication [94][99][100].

A growing list of reports indicates that these circulating miRNAs can be detected and quantitatively analyzed in biofluids, including serum, plasma, urine, and saliva [90].

Although the target genes and biological functions of individual miRNAs remain largely unknown, there is a growing body of evidence indicates that miRNAs have diverse functions in both normal and pathological states [101][90]. Therefore, investigation of miRNA regulation and function may add new diagnostic and therapeutic tools for various diseases. There is great expectation for miRNAs as novel biomarkers with diagnostic, prognostic and predictive value. However, we are still at a primordial phase of these investigations and so far no miRNA has been validated to be superior to the traditional biomarkers in use.

2.6.1 Soy and isoflavons

Diet plays an important role in daily life. People nowadays consider nutrition not just as nourishment, but are more conscious about food and define their lifestyle by what they eat. Also, there is growing public interest in achieving and sustaining good health through healthy eating. It is well recognize that changes in lifestyle and the intake of diets low in fat and high in complex carbohydrates from grains, fruits, and vegetables is associated with a lower risk of chronic disease [102].

The potential health benefits of soy are widely publicized. Research over the past two decades has led to renewed interest in health and nutritional aspects of soy-foods that focuses on much more than their protein and fat content. There is particular interest in understanding the possible protective role of soy against coronary heart disease [103][104], certain forms of cancer, such as breast and prostate cancer [105] [106], osteoporosis [107], and alleviation of menopausal hot flushes [108].

Isoflavones are diphenolic compounds present in plants like soybeans, red clover, and kudzu root. Soy is the most common dietary source for isoflavones. The predominant isoflavones in soy are genistein, daidzein, and, in lower concentration, glycitein. Soy isoflavones have structural similarities to endogenous 17β-estradiol, which explains their extrogenic activity and identifies them as phytoestrogens (Fig. 9).

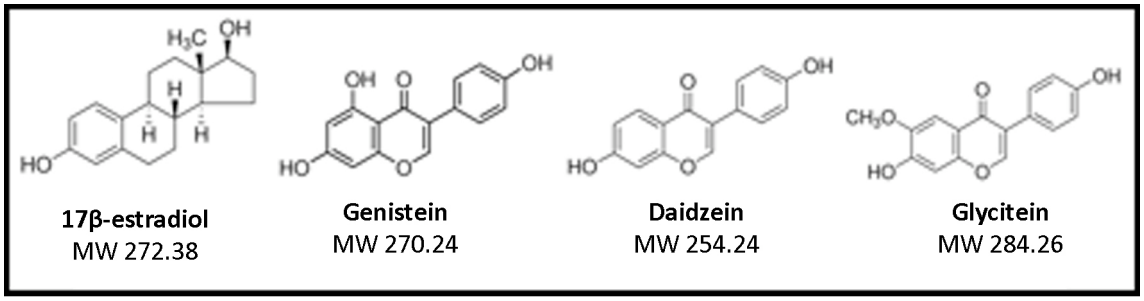


Figure 9. Molecular structures of 17β-estradiol and isoflavones genistein, daidzein and glycitein. MW= Molecular weight.

Because of this structural and functional similarity to estradiol, soy isoflavones are able to selectively bind to estrogen receptors (ER) with distinct affinities, modulating the recruitment of co-repressors and co-activators and thus affecting ER-signaling [109]. Independent of ER binding, isoflavones and especially genistein, hold the potential to exert physiological effects since they affect signal transduction pathways by inhibiting the activity of many enzymes (e.g., tyrosine protein kinase, mitogen activated kinase, and DNA topoisomerase) and regulating cellular factors that control the growth and differentiation of cells [110][111].

2.6.2 Plasma profile of microRNAs after soy supplementation in patients with T2DM and subclinical hypogonadism–a randomized double-blind parallel study: Rationale of the study

Isoflavones exposure has been shown to lowers circulating testosterone levels and sperm concentration in rodents [112][113]. In humans only one report on inverse association between soy food intake and sperm concentration has been reported so far [114], and a case report describing a 60 years old man who developed gynecomastia allegedly in response to the consumption of soymilk [115]. However, in contrast to the effects in rodents, a recently published meta-analysis of the clinical research found that there were no effects of soyfoods or isoflavone supplements on total or free testosterone or dihydrotestosterone levels [116]. Similarly, the clinical evidence shows that soy does not raise estrogen levels [117]. Finally, testosterone deficiency is common in men with type 2 diabetes [118], and lower testosterone levels seems to be due to a combination of factors, including obesity and insulin resistance [119][120].

Given the estrogenic effect of phytoestrogens, a randomized double-blind, parallel study was undertaken in men with T2DM and compensated hypogonadism. The hypothesis to test in this study was if there were any effects of soy with and without isoflavones on testosterone levels, this would be exaggerated in men with low testosterone levels. The primary outcome of this study was the change in testosterone levels. The secondary outcomes were changes in glycemic control and cardiovascular risk markers including insulin resistance, lipid profile, highly sensitive CRP (hsCRP) and endothelial function.

Then, miRNA expression profile was assessed in peripheral plasma samples of 10 subjects in order to evaluate a possible change in expression level after treatment with soy and isoflavones.

3. AIM OF THE STUDY

Genetic studies have greatly contributed to our understanding of diabetes mechanisms by identifying new genes and pathways, previously not linked to diabetes based on existing hypothetical models. However, the impact of these findings in the clinic with respect to identification of predictive and prognostic markers as well as new therapeutic targets, is still low and unsatisfying.

Aim of the present study was investigating on genetic variation contributing to glycaemic trait variability, representing a fascinating line of research because it is centered to the pathophysiological processes leading to diabetes through the relationships to other metabolic traits. In particular, we focused our study on the molecular characterization of the FN3K gene, whose product is implicated in non-enzymatic glycation reaction, in a population of Italian individuals with diabetes. First, the analysis of the FN3K gene in a well clinically characterized cohort of 35 T1DM and 35 T2DMand 33 healthy subjects, was completed by analyzing its promoter region. Then, the FN3K gene (promoter region and all six exons with corresponding intron/exon boundaries) were analyzed in additional 80 T2DM subjects, followed since long time. The main purposes was to identify and investigate on genetic variants within the FN3K gene and correlate genotypes with some clinical parameters typically associated with diabetes in order to offer a better comprehension of the glycemic control and of its predictive role in the middle and long period of the disease.

A second and innovative aim of this work refers to new evidences for an epigenetic role in influencing the development of diabetes. Particularly, environmental exposure to nutrients may change gene expression and alter disease susceptibility through epigenetic modifications. For this purposes miRNA plasma profile of 10 subjects with T2DM and compensated hypogonadism administrated with soy and isoflavons was investigated to evaluate a possible effect of soy in influencing miRNA expression level.

4. MATERIALS AND METHODS

4.1 FN3K ANALISYS

4.1.1 Patients

Diabetic subjects analyzed for the FN3K gene were totally 150. All the individuals were recruited at the outpatient Diabetic Center of the Department of Medical and Surgical Sciences at the University of Padua. Patients were recruited in different period: the first, cohort 1 was made up of 70 patients, 35 T1DM and 35 T2DM; while the second, cohort 2, consisted of 80 T2DM subjects.

Diabetic subjects belonging from a subgroup of patients enrolled in the ADAG (A1c-Derived Average Glucose) study [121]. The presence of T1DM and T2DM was assessed according to WHO criteria [24]. The diabetic patients had to have stable glycemic control, as shown by two HbA1c values in the 6 months prior to enrolment, with a variation of no more than 1%. Furthermore, they had to be able to perform self-monitoring of blood glucose and continuous glucose monitoring.

Patients affected by any condition that could cause changes in glycemic concentrations, as disease requiring steroid therapy or plans for pregnancy during the study, as well as patients affected by any condition that could influence HbA1c concentrations (anemia, high erythrocyte turnover, blood loss, transfusions, hemoglobinopathies, liver and kidney disease, infections) were excluded from the study. All diabetic patients were between 18 and 70 years of age.

The control cohort consisted of 33 healthy blood donors also recruited from Padua Hospital. Control subjects had no history of diabetes, fasting glucose values ≤100 mg/dL and HbA1c levels comprises between 4 and 6%. Blood samples were collected in EDTA tubes and stored at –80°C until use.

Written informed consent was obtained from all participants, also with regards to genetic testing. The study was approved by the local Ethic Committee and was conducted in accordance with the Declaration of Helsinki [122].

4.1.2 DNA extraction

DNA was extracted from whole blood of subjects using the QIAamp® DNABlood Mini Kit (QIAGEN); the procedure comprised 4 steps and was automated on the QIAcube (QIAGEN) (Fig. 10), an instrument that uses advanced technology to process QIAGEN spin columns, enabling seamless integration of automated, low-throughput sample preparation. Up to 12 samples could be processed per run.

Figure 10. QIAcube instrument. The outside on the left and the inside on the right.

The principle of the QIAcube extractor is based on the lysis and elimination of red blood cells and on the subsequent extraction of genomic DNA from leukocytes through cells lysis by using an anionic detergent. Contaminating RNA is removed by enzymatic treatment, while other contaminants, such as proteins, are eliminated by separation on pre-packed columns.
Sample preparation using QIAcube followed the same 4 steps as the manual spin procedure (Fig. 11). Each whole blood sample was thawed and 200 µL were used for DNA extraction in order to obtain about 3-12 µg of DNA. Briefly, optimized buffers lysed samples, stabilized nucleic acids and enhanced selective DNA adsorption to the QIAamp membrane. Alcohol was added and lysates were loaded onto the spin columns. DNA was adsorbed onto the silica membrane during a brief centrifugation. Salt and pH conditions in the lysate ensured that proteins and other contaminants, which could inhibit PCR and downstream enzymatic reactions, were not retained on the membrane. DNA bound to the membrane was washed in two centrifugation steps by using two different wash buffers; wash conditions ensured the complete removal of any residual contaminants without affecting DNA binding. Purified DNA was then eluted from the spin column in a concentrated form in a low salt buffer (100 µL final volume).

The concentration of DNA extracted was checked by spectrophotometry; DNA purity was assessed through λ260/λ280 ratio (expected ratio ca. 1.75).
DNA was stored at -20 °C until its use.

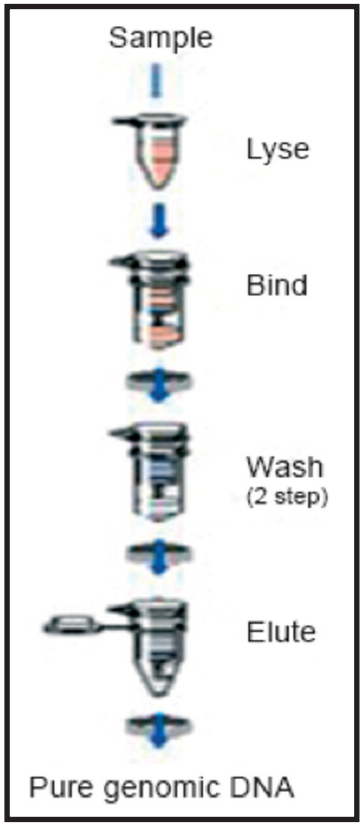


Figure 11. Schematic representation of the spin procedure for DNA extraction by using the QIAamp® DNA Blood Mini Kit

4.1.3 DNA Amplification

Extracted DNA was amplified by means of Polymerase Chain Reaction (PCR). Primers were designed to amplify the promoter region and the 6 exons and the corresponding intron/exon boundaries of FN3K gene [71], in order to exclude pseudogene amplification located on chromosome 22 [61].

Specificprotocols were developed in our laboratory in order to amplify the largest number of exons with the same protocol. Because of its length, exon 6 was divided into two parts and amplified with two different PCR reactions.

Table 2 shows the length of PCR products and some characteristics of the primers (sequence, melting temperature) used in each PCR amplification.

Exon	Primer		PCR product (bp)	Tm (°C)
Promoter	forward	TGGAATCAGCAGCCCCAC	276	67
	revense	GGCTTCGCGCTTCTTGCG		72
1	forward	AGCGCCTCAGCCGACAGG	244	63
	revense	AGGCCCCGCAGAGACCCGC		68
2	forward	TGGTAGCTTCTGTCAAGGGCCCA	314	66
	revense	CCTTTCTCTGTCTCTCTCTGTTTGT		64
3	forward	TCTCAAAATGGAACAAACTGGGTGG	217	64
	revense	GGTAGACAATGTGGCTGGATCAAG		65
4	forward	GACCACCTATATTCTAGCATGCGT	251	64
	revense	AAAGAGTGGGTGACCAGCATTAGG		65
5	forward	AGGCAGCACTTGGCTCCGAATTGC	227	69
	revense	GGATCATGAGAGGACCAGC		59
6-(5')	forward	GAACGAGGTGATGAGACCG	274	59
	revense	GGATCTTCCGGTGGTAGGC		62
6-(3')	forward	CGAGTTTGAACTGGCAATCG	256	58
	revense	ACGGGGAGACGGAGACGGGGACGG		75

Table 2. PCR products lengths, primers sequences and melting temperature (Tm).

Each PCR reaction was performed in a total volume of 25µL (PCR mixture).

PCR mix for amplification of the promoter region:

- 16µL H2O
- 2.5 µL PCR Buffer II (100 mM Tris-HCl, pH 8.3, 500 mM KCl)
- 0.6 µL 25 mM MgCl2
- 2 µL 2.5 mM dNTPs
- 1 µL 12 mM primer forward
- 1 µL 12 mM primer reverse
- 0.4 µL FastStartTaq® DNA Polymerase
- 1,5 µL genomic DNA

PCR mix for amplification of exon 1:

- 14.4 µL H2O
- 2.5 µL PCR Buffer II (100 mM Tris-HCl, pH 8.3, 500 mM KCl)
- 0.8 µL 25 mM MgCl2
- 2 µL 2.5 mMdNTPs
- 2 µL 10 mM primer forward
- 2 µL 10 mM primer reverse
- 0.3 µL AmpliTaq® Gold DNA Polymerase
- 1 µL genomic DNA

PCR mix for amplification of exons 2, 3, 4, 5, 6-(5'):

- 13.7 µL H2O
- 2.5 µL PCR Buffer II (100 mM Tris-HCl, pH 8.3, 500 mM KCl)
- 1.5 µL 25 mM MgCl2
- 2 µL 2.5 mMdNTPs
- 2 µL 10 mM primer forward
- 2 µL 10 mM primer reverse
- 0.3 µL AmpliTaq® Gold DNA Polymerase
- 1 µL genomic DNA

PCR mix for amplification of exon 6-(3'):

- 14.3 µL H2O
- 2.5 µL PCR Buffer II (100 mM Tris-HCl, pH 8.3, 500 mM KCl)
- 0.9 µL 25 mM MgCl2
- 2 µL 2.5 mMdNTPs
- 2 µL 10 mM primer forward
- 2 µL 10 mM primer reverse
- 0.3 µL AmpliTaq® Gold DNA Polymerase
- 1 µL genomic DNA

Two different programs on a thermal cycler (GeneAmp® PCR System 9700, Applied Biosystems) were used for PCR amplifications.
For the promoter region:
• 1 cycle at 96 °C for 10 minutes, to ensure the complete denaturation of template DNA and the activation of the polymerase;
• 30 cycles, each one composed of: 95 °C for 40 seconds (DNA denaturation),62.2°C for 60 seconds (primers annealing), 72 °C for 60 seconds (DNA extension);
• 1 cycle at 72 °C for 10 minutes to fill-in the protruding ends of reaction products.
For all exons:
• 1 cycle at 96 °C for 10 minutes, to ensure the complete denaturation of template DNA and the activation of the polymerase;
• 35 cycles, each one composed of: 95 °C for 30 seconds (DNA denaturation), 56°C for 60 seconds (primers annealing), 72 °C for 60 seconds (DNA extension);
• 1 cycle at 72 °C for 5 minutes to fill-in the protruding ends of reaction products.

For each amplification protocol, a negative control was used: it consisted of a PCR mixture aliquot with no DNA added. Negative control had to give no PCR products and was indicative of the presence or absence of contamination in the reaction. Once the amplification was completed, agarose gels were prepared in order to test the quality of the PCR products.
Agarose gel could be prepared at different percentages to allow the separation of PCR products on the basis of their lengths, expressed as base pairs. For the confirmation of our PCR products, a 2% agarose gel was sufficient. In order to prepare a 100 mL solution, 2 g of agarose were weighed and added to 100 mL of TAE buffer (10 mM Tris/HCl, pH 8.0, 1 mM EDTA); the solution was then brought to the boil until it was limpid. After cooling, 8µL of GelRed™ 10000X (Biotium) were added; this red fluorescent nucleic acid dye can be excited with a common300 nm UV transilluminator. Three microliters of amplified DNA were mixed with 2 µL of the gel-loading buffer (GelPilot DNA Loading Dye 5X, QIAGEN), previously diluted 1:2, and the samples were loaded into the gel wells. For sizing the PCR products (217 to 314 bp), a DNAladder (GeneRuler™ 100 bp DNA Ladder, Genenco) was always used; it contained the following 11 discrete fragments: 1031, 900, 800, 700, 600, 500, 400, 300, 200, 100and 80 bp.

The amplification protocol for FN3Kgene was performed using Biomek NXP (Beckman Coulter). BiomekNXP puts every aspect of liquid handling, including pipetting, dilution,

dispensing and integration, into a single, automated system that is as powerful and flexible as it is efficient and economical (Fig. 12).

Figure 12. Biomek NXP (Beckman Coulter).

Biomek NXP offers proven pipetting performance for low-volume reaction setup and assay miniaturization, as well as accurate and repeatable results extending into the submicroliter range. Since many functions are combined into a single step, setup is fast and effortless.

Automation of reaction setup can also eliminate both human error and contamination associated with manual processing, while assuring high quality results. The reliability of this instrument, repeatable pipetting, tube to plate capability and an efficient automation combine to provide an ideal automation solution for the amplification protocol. Moreover its low-volume capability helps miniaturize reaction setups and save on costly reagents.

Biomek NXP combined with robust Biomek Software is used to manage and program the instrument (Fig. 13). This software provides the ability to customize the interface setting for every PCR step, as type of tip to use, source and destination of reagents and it is able to configure methods with drag-and-drop ease. The users have only to prepare the worktop with tips, reagents, tubes, primers, DNA and plate of destination in the correct position.

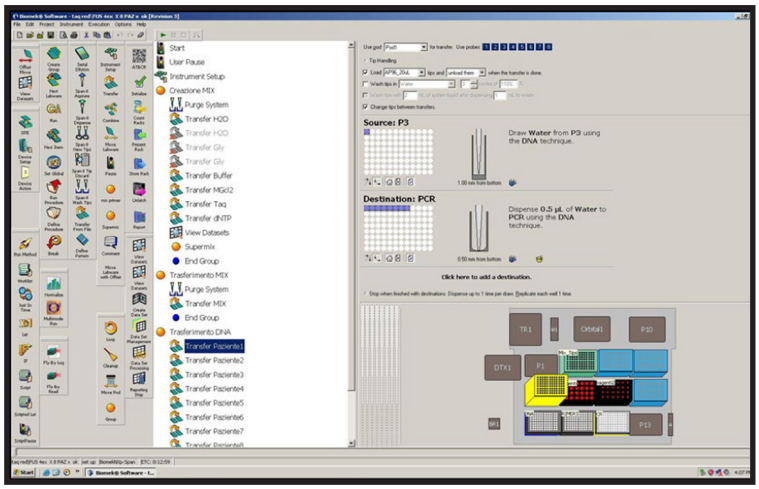


Figure 13. Software interface of Biomek NXP.

4.1.4 Purification of PCR products

Amplicons have to be purified to eliminate excess oligos, nucleotides, salts, enzymes and other contaminants that may inhibit subsequent reactions. The purification was performed using AMPure® PCR Purification Kit (Agencourt® Bioscience Corporation) (Fig. 14) and automating the process through Biomek 3000 (Beckman Coulter).
The Agencourt AMPure method utilizes Solid Phase Reversible Immobilization (SPRI) magnetic bead-based technology, which requires no centrifugation or filtration. The Agencourt AMPure purification process binds PCR amplicons to para-magnetic particles and draws them out of solution, allowing contaminants such as primers, primer dimers, salts, and dNTPs to be easily rinsed away providing purified PCR product ideal for downstream genomics applications. The para-magnetic particles a reconstituted for 40% of magnetite and have uniform size (1 µm); their negative charge binds PCR products 100 bp and larger. Placed in a magnetic field, the amplicons remain bound to the magnetic beads in the plate wells, while the washes are performed and the contaminants removed. At the end the purified DNA is eluted with dd H2O and transferred to another plate.

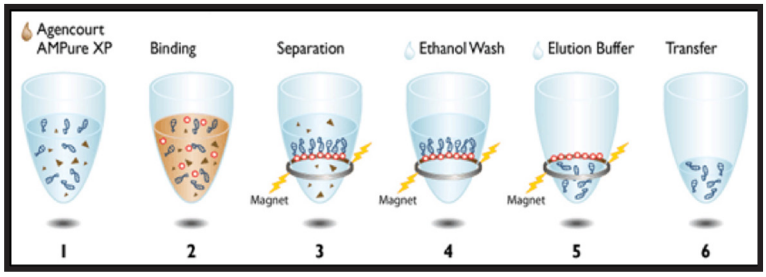


Figure 14. Schematic representation of PCR purification.

The AMPure procedure is performed in 8 steps:

1. MixAMPure reagent with the PCR reaction mixture. Add 1.8 MlAMPure per1.0 µL of PCR product;

2. Bind PCR products to para-magnetic beads and incubate for 3-5 minutes;

3. Move the plate on the magnetic support, incubate for 5-10 minutes to separate beads from solution;

4. Aspirate the cleared solution from the reaction plate and discard;

5. Dispense 200 µL of 85% ethanol to each well of the reaction plate and incubate for 30 seconds at room temperature; aspirate out the ethanol and discard; repeat for a total of two washes;

6. Place the reaction plate on bench top to air-dry; the plate should be left for10-20 minutes to allow complete evaporation of residual ethanol that may inhibit the subsequent reactions;

7. Add 40 µL of dd H2O to each well of the reaction plate and mix in order to elute purified PCR product from the beads;
8. Transfer the purified products to a new plate.

This process has been automated and optimized by using the Biomek 3000 (Beckman Coulter) (Fig. 15). This instrument is an automated laboratory work station fully compatible with Agencourt AMPure.
As for the Biomek NXP, Biomek 3000 automates many processes. This instrument has a mechanic alarm that moves on 3 axes (X, Y, Z) mono and multichannel pipettes (P20, MP20, P200, MP200, P1000) besides the gripper tool that picks up and dislocates the plates. On the worktop there is space to accommodate the magnetic support as well as containers for reagents, plates, tips and waste. User-friendly software facilitates tailoring each protocol to specific requirements, or building new protocols from scratch. The user must have the material correctly placed on the worktop, as indicated on the program, and start it.

Figure 15. Biomek 3000 (Beckman Coulter).

Purified DNA fragments were quantified in order to settle the suitable quantity of DNA to use in the subsequent sequencing reaction. DNA yield was determined by 2% agarose gel analysis. The amount of loaded DNA sample was estimated by visual comparison of the band intensity with that of the standards (GeneRuler™ 100 bp DNA Ladder, Genenco); DNA quantity for a correct sequencing reaction is 10 ng per 100 bp. Purified PCR products were stored at 4 °C until their use.

4.1.5 Sequencing analysis

Sequencing a DNA fragment amplified by PCR allows its nucleotide sequence to be determined, individualizing the possible presence of a variation. The most commonly used approach is the dideoxy chain termination method (Sanger method) [123]. The key principle of the Sanger method is the use of dideoxynucleotide triphosphates (ddNTPs) as DNA chain terminators: they are analogues of the normal dNTPs, but differ in that they lack a hydroxyl group at the 3' carbonposition as well as at the 2' carbon. Lacking this group, any ddNTP that is incorporated into a growing DNA chain cannot participate in phosphodiester bonding at its 3' carbon atom, thereby causing abrupt termination of chain synthesis. Competition for incorporation into the growing DNA chain between a ddNTP and its normal dNTP analogue results in a population of fragments of different lengths. The technique requires a single-stranded DNA template, a DNA primer and a sequencing kit containing DNA polymerase, dNTPs and the four ddNTPs labeled with fluorescent dyes, each of which with different wave lengths of fluorescence and emission.

The newly synthesized and labeled DNA fragments are heat denatured and separated by size by capillary electrophoresis (CE). During CE, the DNA fragments are electrokinetically injected into capillaries filled with a high resolution polymer.

High voltage is applied so that the negatively charged DNA fragments move through the polymer toward the positive electrode. CE can resolve DNA molecules that differ in molecular weight by only one nucleotide. Shortly before reaching the positive electrode, the fluorescently labeled DNA fragments move throught he path of a laser beam that causes the dyes on the fragments to fluoresce. An optical detection device detects the fluorescence; specific software converts the fluorescence signal to digital data and records them (Fig. 16). Since each dye emits light at a different wavelength when excited by the laser, all four colors, and therefore, all four bases, can be detected and distinguished in one capillary injection (Fig. 17).

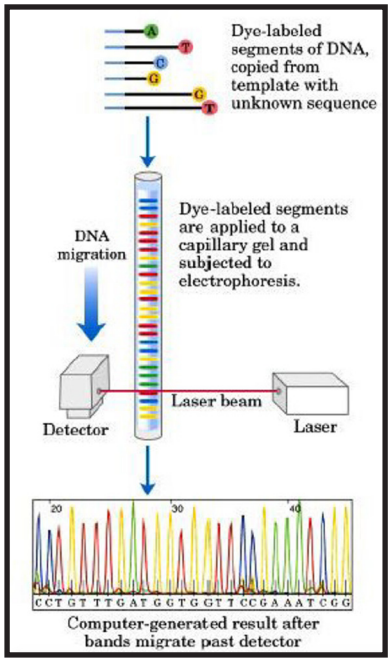


Figure 16. Schematic representation of the automated DNA sequencing with capillary electrophoresis (www.appliedbiosystem.com).

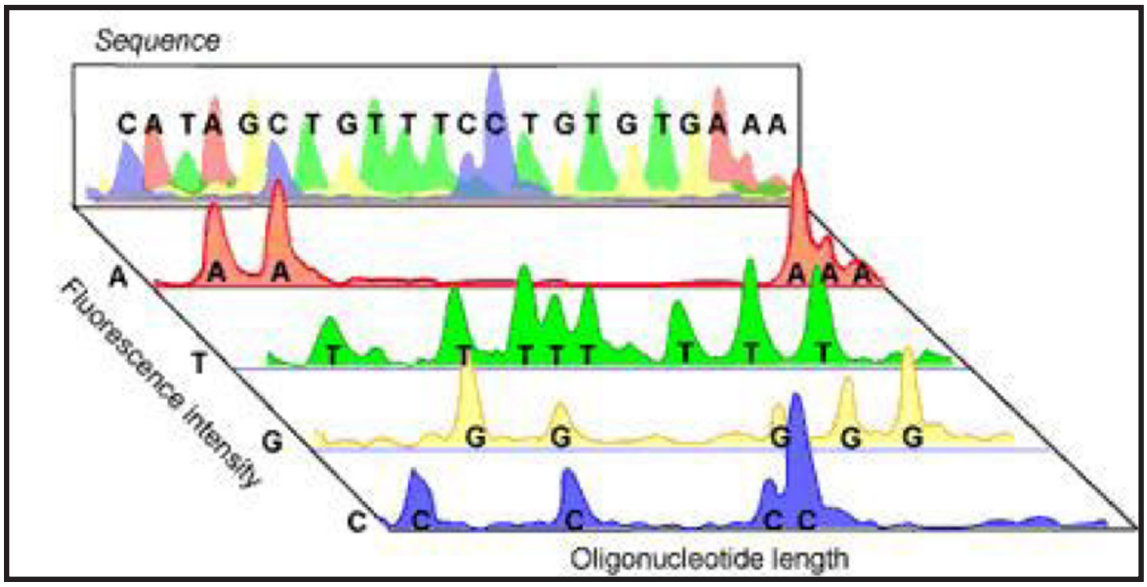


Figure 17. Electropherogram (www.appliedbiosystem.com).

4.1.6 Sequencing reaction protocol

The sequencing reaction was performed by using the BigDye Term v1.1 Cycle Seq Kit (Applied Biosystems).

Sequencing protocol involved the use of:

- 1 µL of 10 µM primer forward or reverse;
- 1 µL of DNA sequencing Mix;
- 1µL of BigDye® Terminator v1.1 5X Sequencing Buffer;
- DNA quantity established following the confirmation of PCR products on the agarose gel (usually 1 µL of amplified DNA was used);
- H2O to a final volume of 10 µL.

Each amplicon of the FN3K gene was sequenced on both strands; therefore two sequencing reactions, one with primer forward and the other with primer reverse, were performed. The reaction plate with the sequencing mix was automated using Biomek 3000 (Beckman Coulter).

For the sequencing reaction the following program on the thermalcycler was performed:

• 25 cycles, each one composed of: 96 °C for 10 seconds (DNA denaturation), 50 °C for 5 seconds (primer annealing), 60 °C for 4 minutes (DNA extension).

4.1.7 Purification of sequencing reaction products

It was necessary to purify the sequencing reaction products in order to remove unincorporated dye terminators and all the other reagents and by-products. Removal of unincorporated dye-terminators is required prior to capillary electrophoresis, because they can cause a “background noise” which makes the electropherograms unreadable. The CleanSEQ Sequencing Reaction Clean-Up has proven an efficient and highly effective dye-terminator removal system with the ABI 3730 sequencer. This system utilizes Agencourt’s patented SPRI™ para-magnetic bead technology as in the AMPure purification process previously described.

Also in this case, the process is automated through Biomek 3000 (Beckman Coulter). The protocol can be performed directly in the thermal cycler plate. CleanSEQ contains magnetic particles in an optimized binding buffer to selectively capture sequencing extension products, whereas unincorporated dyes, nucleotides, salts and contaminants are removed using a simple washing procedure.

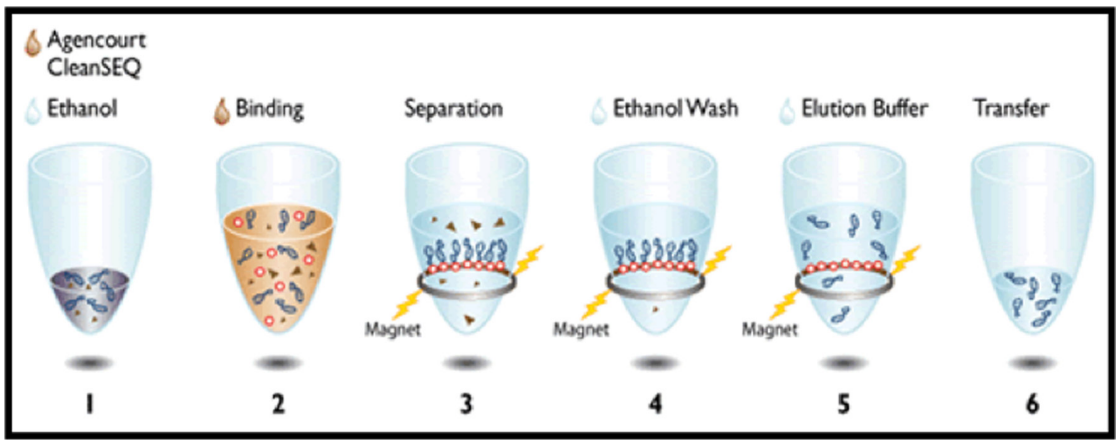


Figure 18. Schematic representation of CleanSeq purification.

The CleanSEQ procedure is performed in 7 steps:

1. Gently shake the CleanSEQ bottle to re-suspend any magnetic particles that may have settled. Add 10 µL of CleanSEQ to the reaction plate;
2. Add 42 µL of 85% ethanol to the reaction plate; pipette the mix 7 times;
3. Place the reaction plate onto magnetic support for 3 minutes to separate beads from solutions;
4. Aspirate the cleared solution from the reaction plate and discard;
5. Dispense 100 µL of 85% ethanol and incubate at room temperature for at least 30 seconds, then aspirate out the ethanol and discard; repeat for a total of 2 washes;
6. Let the reaction plate air-dry for 10 minutes at room temperature;
7. Add 40 µL of dd H2O and incubate the plate for 5 minutes at room temperature and then transfer samples into a new clean reaction plate.

4.1.8 Sequencing

Samples preparation

For DNA sequencing a 48-capillary 3730 DNA Analyzer (Applied Biosystems) instrument was used. Once purified, samples containing the sequencing reactions were prepared on a special plate for multi-channel electrophoresis runs. This plate was formed by 96 wells; purified sequences were added in alternate rows since the 48 capillaries enter simultaneously in different wells. Two microliters of purified products together with 8 µL of deionized formamide were placed in each well. The plate was covered with a rubber mat and centrifuged to bring the samples down to the bottom of the wells. Samples were denatured at 95 °C for 2 minutes on a thermal cycler and then incubated at -20 °C for at least 5 minutes.

Samples loading

Before loading the plate in the instrument, the worksheet showing the order of samples was entered in the computer. The first step was to appoint the plate with the "bar code" present on its side, together with the date of the transaction. The worksheet is a table in which rows represent the name of wells and columns the name of the sample, the protocol type and the reading type to be used. The plate was then inserted in the instrument: the 48 samples placed in alternating columns were analyzed first.

4.1.9 Software for sequence analysis

After the electrophoretic run and the acquisition of data by the analyzer, each DNA fragment sequence could be viewed on the computer screen: the four bases, labeled with different fluorophores, were detected and represented as peaks of four colors (green for adenine, blue for cytosine, black for guanine and red for thymine). Raw data analysis was performed by using the "Sequencing Analysis" software (Fig. 19), designed to base-call, trim, display, edit, and print DNA sequencing data generated from the genetic analyzer; this step allowed to check if the procedure was successful or if the electrophoretic run had to be repeated. Moreover, this software provided multiple metrics (sample score, read length, signal-to-noise values) that help pinpoint the cause of data failure.

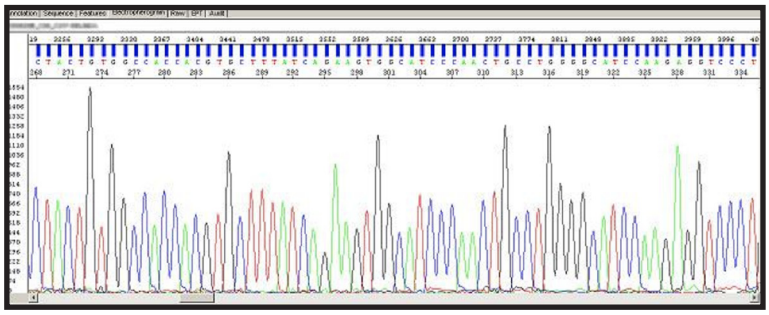


Figure 19. "Sequencing Analysis" interface: the software was used to observe DNA sequences.

By using the SeqScape®software (Fig. 20), all the variations (deletions, insertions, mismatches, heterozygous bases, and heterozygous insertion/deletions) between the reference genomic sequence and each subject sample sequence were detected. Software examines the quality of each forward and reverse trace and forms an accurate consensus call. Variations between the consensus and the reference sequence are reported as mutations in the Mutations report. All analysis in SeqScape®software occurred in a project; before creating a project, a project template containing a reference data group, analysis defaults and display settings was created. Once a project was created, the project template was used recurrently. As for the "Sequencing Analysis" software, the sequence was represented by a succession of peaks characterized by four different colors, one for each nucleotide. A peak overlap (identified by letter N) showed the presence of two different nucleotides in the same position and, therefore, the presence of a variant in the heterozygous state. If the alteration was present in the homozygous state, a different peak in comparison with the reference sequence was found: however, the change was signaled by the software. The existence of an insertion or a deletion in the sample was characterized by the presence of a complicated pattern of overlapping peaks from the site of the alteration because of a different positioning between the two strands. The presence of any genetic variant was always verified also on the complementary strand (3' —> 5').

Figure 20. SeqScape® interface: the software was used to identified polymorphisms/mutations by comparing the DNA fragment of interest with the reference sequence.

4.2 Denaturing High Performance Liquid Chromatography (DHPLC) analysis

Denaturing high performance liquid chromatography (DHPLC) is a variant of heteroduplex analysis [124] and has emerged as one of the most versatile technologies for the analysis of genetic variations [125]. DHPLC compares two chromosomes as a mixture of denatured and reannealed PCR amplicons. When the PCR amplicon does not contain a mutation, the sense strand and the anti-sense strand of the PCR amplicons are completely complementary to each other. In this case, the molecule is referred to as a homoduplex. When a patient is heterozygous for a mutation, the mutant sense and antisense strands will not only form a homoduplex, but also heteroduplices with their corresponding wild type sense and anti-sense strands upon re-naturation of the denatured PCR product. The physicochemical difference between a homoduplex and heteroduplices can be detected by HPLC using a reversed-phase chromatography column with an affinity to double-stranded DNA. The differential retention of homo- and heteroduplex DNA on the column after partial denaturation indicates the presence of a mutation (Fig.21).

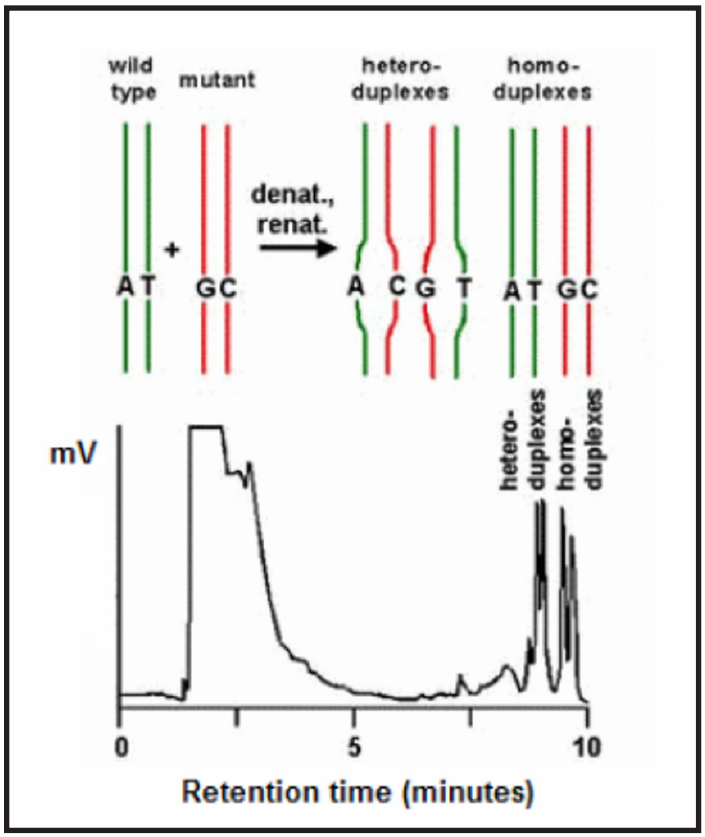


Figure 21. Schematic representation of separation of heteroduplexes from homoduplexes under partially DHPLC conditions for mutation detection.

The temperature of the column determines sensitivity, and the optimal temperature is predicted by a mathematical formula based on the amplicon sequence; the output of this formula can be calculated using a computer program [126]. The stationary phase present in the column is electrically neutral or hydrophobic; DNA fragments that are negatively charged, cannot adsorb to the column's matrix by themselves. Therefore, a "bridging" molecule is needed to help the adsorption of DNA fragments to the stationary phase. Such a molecule is triethylammonium acetate (TEAA), that binds on the one hand to DNA via its positively charged ammonium group, while it also interacts with the hydrophobic column material via its hydrophobic ethyl groups; this way it builds a bridge between the column material and the nucleic acid. Separation of nucleic acids bound to the column is achieved at an optimized temperature by increasing the concentration of an organic solvent (e.g., acetonitrile) in the elution buffer, which competes with the hydrophobic interactions between the column and TEAA, resulting in the elution of the nucleic acid [125].
The presence of genetic variants found in the FN3K gene of diabetic patients by direct sequencing was verified in control subjects by using DHPLC technique. Exons 1 and 4 of the FN3K gene were amplified using protocols described in the section "DNA Amplification". PCR products were denatured at 95 °C for 10 minutes and allowed to cool to 25 °C for the formation of homo- and heteroduplices. DHPLC was carried out using a Transgenomic DHPLC WAVE-MD 4000 Plus instrument and a DNASep column with buffer A (0.1 MTEAA solution in water pH 7.0) and buffer B (0.1 M TEAA solution in water 25 % acetonitrile pH 7.0). Elution of DNA from the column was determined by adsorbance at260 nm. The optimum temperatures for analysis of each fragment were predicted with the Navigator™ software (Transgenomic) and confirmed empirically. DHPLC conditions used for the analysis of each amplicon are described in Table 3: each amplicon was screened at different temperatures.

AMPLICON	LENGHT (bp)	% GC	TEMPERATURE	BUFFER B (%)
Exon 1	244	74.2	68.6	50.7
Exon 4	251	55	63.7	50.7

Table 3. DHPLC conditions for the analysis of the 2 amplicons of the FN3K gene.

The heterozygous profiles were identified by visual inspection of the chromatograms on the basis of the appearance of additional, earlier-eluting peaks; corresponding homozygous profiles showed only one peak. Where necessary, samples were analyzed both alone and with the addition of 50 % wild-type DNA to ensure the detection of homozygous mutations. Samples displaying abnormal eluition profiles were re-amplified and subjected to direct sequencing.

4.3 RNA ANALYSIS

4.3.1 RNA extraction

RNA was extracted using the PAXgene Blood RNA System (PreAnalytiX), consisting of a blood collectiontube (PAXgene™ Blood RNA Tube) and nucleic acid purification kit (PAXgene Blood RNA Kit, QIAGEN).

For the extraction, 2.5 mL of whole blood of subjects of interest were collected in a PAXgene™ Blood RNA Tube. RNA was extracted using the manual procedure as described below:

1. PAXgene™ Blood RNA Tube was centrifuged for 10 minutes at 3000 x g using a swing-out rotor;

2. The supernatant was removed by pipetting and the pellet was resuspended in 4 mL RNase-free water;

3. The pellet was dissolved by vortexing, and centrifuged for 10 minutes at 3000 x g;

4. The supernatant was removed and discarded by pipetting;

5. The pellet was resuspended in 350 µL Buffer BR1 and dissolved by vortexing;

6. The sample was placed in a 1.5 ml microcentrifuge tube, and 300 µL Buffer BR2 plus 40 µL proteinas K were added;

7. After vortexing for 5 minutes, the sample was incubated for 10 minutes at 55°C;

8. The lysate was then placed into a PAXgeneShadder spin column (lilac) placed in a 2 mL processing tube, and centrifuged for 3 minutes at 10000 x g;

9. The entire supernatant of the flow-through fraction was transferred to a fresh 1.5 mL microcentrifuge tube without disturbing the pellet in the processing tube;

10. 350 µL of ethanol (100%) were added, and 1-2seconds centrifugation was performed to remove drops from the inside of the tube lid;

11. 700 µL of sample were pipetted into a PAXgene RNA spin column placed in a 2 mL processing tube and centrifuged for 1 minute at 10000 x g. The spin column was then placed in a new 2 mL processing tube and the old processing one containing flow-through was discarded;

12. 350 µL of Buffer BR3 were added into the PAXgene RNA spin column and centrifuged for 1 minute at 10000 x g. The spin column was then placed in a new 2 mL processing tube and the old processing one containing flow-through was discarded;

13. 80 µL of DNase incubation mix (10 µL of DNase I stock solution + 70 µL of Buffer RDD) were added directly onto the PAXgene RNA spin column membrane, and placed on the bench top (20-30°C) for 15 minutes;

14. 350 µL of Buffer BR3 were added into the PAXgene RNA spin column and centrifuged for 1 minute at 10000 x g. The spin column was then placed in a new 2 mL processing tube and the old processing one containing flow-through was discarded;

15. 500 µL of Buffer BR4 were added into the PAXgene RNA spin column and centrifuged for 1 minute at 10000 x g. The spin column was then placed in a new 2 mL processing tube and the old processing one containing flow-through was discarded;

16. Another 500 µL of Buffer BR4 were added into the PAXgene RNA spin column and centrifuged for 3 minutes at 10000 x g;

17. The PAXgene RNA spin column was placed in a 1.5 mL microcentrifuge tube, and 40 µL of Buffer BR5 was added directly onto the PAXgene RNA spin column membrane.

18. After 1 minute centrifugation at 10000 x g, the elution step (step 17) was repeated;

19. The eluate was incubated for 5 minutes at 65°C and then immediately chilled on ice.

The RNA sample obtained with this procedure was stored at -80°C until its use.

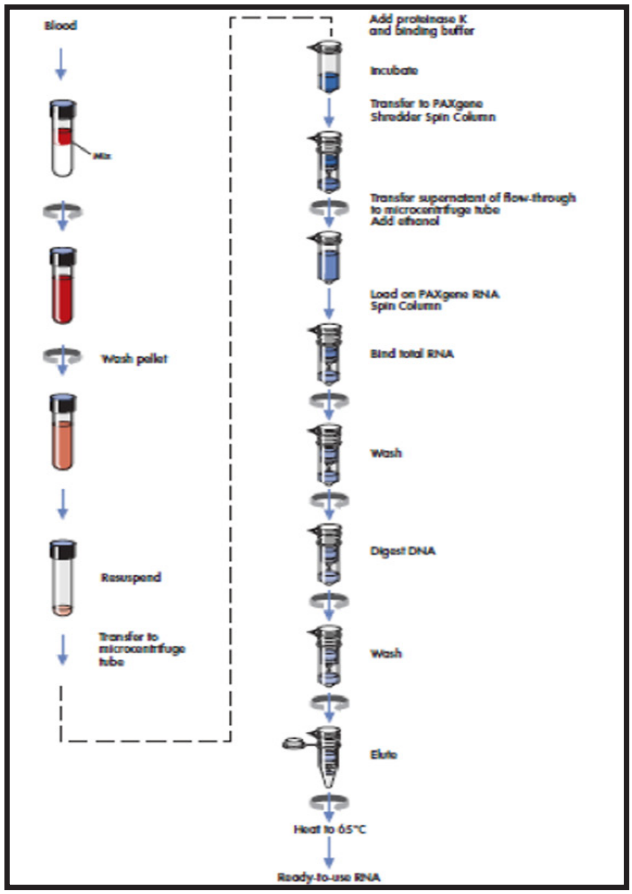


Figure 22. Schematic representation of the manual PAXgene Blood RNA procedure.

4.3.2 Reverse Trancription

Reverse transcription reaction was prepared in a total volume of 20 µL adding the following reagents:

- 1 µg RNA solution
- 1 µL40 µM random primers
- Nuclease free H2O for a total volume of 11 µL

The reaction was incubated at 65°C for 5 minutes and at room temperature for additional 5 minutes, to allow extension of the primers. Then were added:

- 2 µL PCR Buffer II (100 mM Tris-HCl, pH 8.3, 500 mM KCl)
- 4 µL 25 mM MgCl2
- 1 µL 2.5 mM dNTPs
- 1 µL RNAse inhibitor
- 1 µL MuLV reverse transcriptase (Roche)

The sample was incubated at 42°C for one hour. The enzyme was then inactivated heating the sample at 95°C for 5 minutes.

The cDNA sample obtained with this procedure was stored at -20°C until its use.

4.3.3 cDNA Amplification

cDNA was amplified by PCR. Primers were designed to amplify the region containing the new mutations found. The main characteristics are listed in Table 4.

Mutation	Primer		PCR product (bp)	Tm (°C)
c.2 T>A	forward	CTTCCGAGCGAGCA	169	58
	revense	CTGCGGTTGACTTTGA		57
c.465 G>A	forward	GAAGAGGTGAGGGTGC	167	57
	revense	CCTCTCGGTCAGCATAG		57
c.559 C>T c.716 A>G	forward	GAAGAGGTGAGGGTGCTGAG	487	65
	revense	TGGTTCCAGTGGTTCAGGTA		64

Table 4. Shows the length of PCR products and some characteristics of the primers (sequence, melting, temperature) used in each PCR amplification.

Each PCR reaction was performed in a total volume of 25µL.

PCR mixture to amplify the fragments with the c.2 T>A, and both c.559 C>T and c.716 A>G mutations contained:

- 16,1 µL H2O
- 2,5 µL PCR Buffer II (100 mM Tris-HCl, pH 8.3, 500 mM KCl)
- 1 µL 25 mM MgCl2
- 2 µL 2.5 mMdNTPs
- 1 µL 15mM primer forward
- 1 µL 15mM primer reverse
- 0,4 µL FastStartTaq® DNA Polymerase
- 1 µL cDNA

PCR mixture to amplify the fragment with c.465 G>A mutation contained:

- 16,6 µL H2O
- 2,5 µL PCR Buffer II (100 mM Tris-HCl, pH 8.3, 500 mM KCl)
- 0.5 µL 25 mM MgCl2
- 2 µL 2.5 mMdNTPs
- 1 µL 15 mM primer forward
- 1 µL 15 mM primer reverse
- 0,4 µL FastStartTaq® DNA Polymerase
- 1 µL cDNA

The program on thermal cycler was settled as follow:1 cycle at 95 °C for 4 minutes; 95 °C for 30 seconds, 57°C for 30 seconds; 72 °C for 30 seconds, for35 cycles; 1 cycle at 72 °C for 10 minutes.
For each amplification protocol, a negative control consisted of a PCR mixture aliquot with no cDNA added was used, to give indication of the presence or absence of contamination in the reaction. Once the amplification was completed, 2% agarose gels were prepared in order to test the quality of the PCR products.

4.4 miRNA ANALISYS

4.4.1 Patients

Ten subjects were selected from a randomized double-blind, parallel study undertaken with 200 men with T2DM, recruited after screening 412 Caucasian patients at Diabetes Research Centre in Kingston upon Hull, UK. Subjects aged between 45 to 75 years with a total testosterone level ≤12 nmol/L and normal gonadotrophins. They were randomized and administered either 30g soy protein with 66mg of isoflavones (SPI) per day, or 30g soy protein alone without isoflavones (less than 300 parts per billion) (SP) for 12 weeks.

Selected subjects were all under active treatment with soy and isoflavones.

All patients were on stable medications for their T2DM, hyperlipidemia and hypertension for at least 3 months prior to the study. Patients with significant hepatic or renal impairment, who were allergic to soy products, receiving testosterone replacement or who had received antibiotics 3 months prior to the study were excluded.

At randomization and during study visits, subjects were instructed to maintain their level of physical activity throughout the study. In addition, they were required to avoid food products containing soy, alcohol, vitamin or mineral supplementation, and over-the-counter medications. Dietary reinforcement was undertaken at each visit, together with measurement of plasma isoflavone levels to ensure compliance. Compliance with study preparation was calculated by counting the returned bars. All subjects gave their written informed consent. The study was given ethical approval by the Research Ethics Committee (East Yorkshire & North Lincolnshire Research Ethics Committee, ref:09/H1304/45).

4.4.2 miRNA extraction

RNA isolation was performed from plasma samples using the mirVana™ PARIS™ Kit (Life Technologies). The procedure consisted of the following steps:

1. each sample was thawed; 350 µL of plasma sample were placed in a 1.5 mL tube together with 350 µL of denaturing solution (containing 2-mercaptoethanol) and incubated on ice for 5 minutes;

2. 6 µL of 1:200 dilution stock solution of C. elegans miR-93-3p spike-in standard were added to the each sample;

3. 700 µL of acid-phenol:chloroform were added and mixed by vortexing for 30 seconds;

4. a centrifugation at 10000 g for 5 minutes at room temperature was performed;

5. the aqueous layer was removed without touching the interphase or organic layers and placed in a new 1.5 mL tube;

6. the same volume of RNase-free water removed in the previous step, was added to the interphase or organic layers, and mixed by vortexing for 30 seconds;

7. a centrifugation at 10000 g for 5 minutes at room temperature was performed;

8. the aqueous layer was removed and combined with the previous one collected;

9. 1.25 volumes of 100% ethanol were added to the combined aqueous phases and mixed thoroughly;

10. 700 µL lysate/ethanol mix were added to a filter column placed in a collection tube;

11. a centrifugation at 10000 g for 30 seconds at room temperature was performed;

12. the flow through was discarded, and 700 µL of miRNA wash solution 1 were added;

13. a centrifugation at 10000 g for 15 seconds at room temperature was performed;

14. the flow through was discarded, and 500 µL of miRNA wash solution 2/3 were added;

15. a centrifugation at 10000 g for 15 seconds at room temperature was performed;

16. steps 14) and 15) were repeated;

17. the flow through was discarded and the sample was centrifuged at 10000 g for 1 minute;

18. the column was placed in a fresh1.5 mL tube and 10µL of preheated to 95°C RNase-free water were added to the filter to elute miRNA;

19. a centrifugation at 10000 g for 30 seconds at room temperature was performed;

20. steps 18) and 19) were repeated.

The 20 µL of RNA sample obtained with this procedure were stored at -80°C until required.

4.4.3 Screening of miRNA

miRNA detection experiments were performed on Serum/Plasma Focus microRNA PCR Panel I+II (V6.0) using the miRCURY Locked Nucleic Acid (LNA™) Universal RT microRNA PCR System (Exiqon). This microRNA-specific, LNA™-basedsystem is designed for sensitive and accurate detection of microRNA by quantitative real-timePCR using SYBR® Green. The method is based on universal reverse transcription (RT) followed by real-time PCR amplification with LNA™ enhanced primers (Fig. 23).

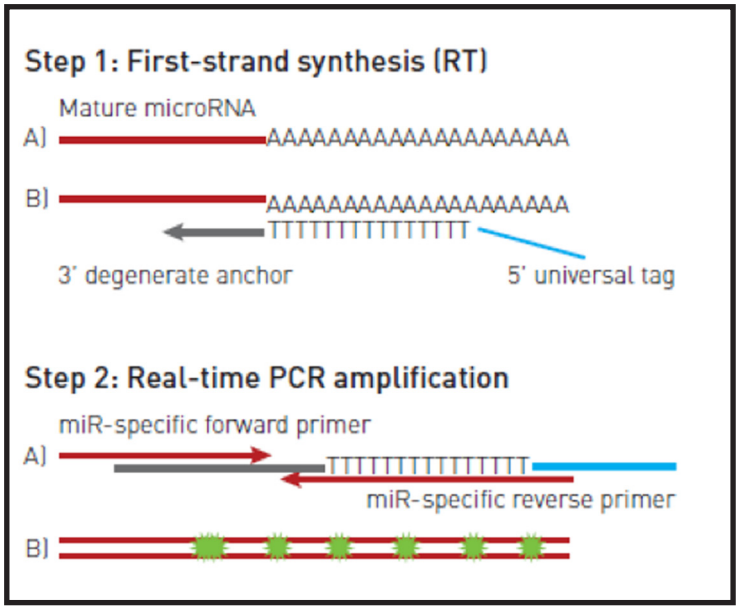


Figure 23. Schematic outline of the miRCURY LNA™ Universal RT microRNA System. A polyA tail is added to the mature microRNA template (step 1A). cDNA is synthesized using a polyT primer with a 3' degenerate anchor and a 5' universal tag (step 1B). the cDNA template is then amplified using microRNA-specific and LNA™-enhanced forward and reverse primers (step 2A). SYBER® Green is used for detection (step 2B).

4.4.4 cDNA synthesis (RT)

The reverse transcription reaction was performed in a total volume of 40 µL.
The reaction setup contained:

- 20 µL nuclease free water;
- 8 µL 5x Reaction buffer;
- 4 µL Enzyme mix;
- 8 µL RNA template (5 ng/µL)

The RT-PCR was settled as follows: incubated at 42°C for 60 minutes, heat-inactivated at 95°C for 5 minutes, and immediately cooled to 4°C using a thermalcycler (GeneAmp® PCR System 9700, AppliedBiosystems). The cDNA was then stored at -80°C or used for PCR reaction.

4.4.5 Real-time PCR amplification

Before starting, the cDNA, the nuclease free water and the ExiLENT SYBR® Green master mix 2x concentrated were placed on ice for 15-20 minutes, and plates were spinned down in a plate centrifuge. The qPCR protocol for two plates of 96 wells was performed as follows:

- 1050 µL of ExiLENT SYBR® Green master mix were combined with 1039.5 µL of nuclease free water and10.5 µL of cDNA;
- 10 µL of PCR Master mix:cDNA mix were added to each well.

The experiment can be paused at this point and stored, protected from light, at 4°C for up to 24 hours.
The amount of PCR product was detected by evaluating the level of fluorescence emitted by SYBR® Green. Quantitative PCR was carried out on a StepOnePlus™ 96-well Real-Time PCR instrument (Life Technologies) according to the following instruction: 95°C for 10 minutes, 40 amplification cycles at 95°C for 10 seconds, 60°C for 1 minute. Quantitative PCR reactions were concluded with a melting curve to disassociate the SYBR Green dye from amplification products, allowing the specificity of the amplification to be assessed.

4.4.6 Serum/Plasma Focus miRNAPCR Panels I+II

Serum/Plasma Focus microRNA PCR Panels I+II (V6.0) (Exiqon) cover the analysis of 179 human miRNAs and contain: 179 LNA™ primer sets for the amplification of human micro RNA; 2x3 inter plate calibrators, 5 RNA spike-in control primer sets; two blank wells, one in each plate (Fig. 24, 25).

	1	2	3	4	5	6	7	8	9	10	11	12
A	hsa-miR-652-3p	UniSp3 IPC	hsa-miR-221-3p	hsa-let-7f-5p	hsa-miR-27b-3p	hsa-miR-374b-5p	hsa-miR-93-5p	hsa-miR-200a-3p	UniSp2	hsa-miR-484	hsa-miR-181a-5p	hsa-miR-106a-5p
B	hsa-miR-145-5p	UniSp3 IPC	hsa-miR-486-5p	hsa-miR-26a-5p	hsa-miR-143-3p	hsa-miR-30a-5p	hsa-miR-222-3p	hsa-miR-342-3p	UniSp4	hsa-miR-324-5p	hsa-miR-185-5p	hsa-let-7e-5p
C	hsa-miR-139-5p	UniSp3 IPC	hsa-miR-451a	hsa-let-7a-5p	hsa-miR-128-3p	hsa-miR-125a-5p	hsa-miR-485-3p	hsa-let-7b-3p	UniSp5	hsa-miR-25-3p	hsa-miR-375	hsa-miR-33a-5p
D	hsa-miR-16-5p	hsa-let-7b-5p	hsa-miR-152-3p	hsa-miR-30c-5p	hsa-miR-197-3p	hsa-miR-99b-5p	hsa-miR-30e-5p	hsa-miR-146a-5p	cel-miR-39-3p	hsa-miR-424-5p	UniSp6	hsa-miR-107
E	hsa-miR-148b-3p	hsa-miR-20a-5p	hsa-miR-17-5p	hsa-miR-30d-5p	hsa-miR-186-5p	hsa-let-7i-5p	hsa-miR-26b-5p	hsa-miR-34a-5p	hsa-miR-320b	hsa-miR-590-5p	hsa-miR-191-5p	hsa-miR-99a-5p
F	hsa-miR-301a-3p	hsa-miR-151a-5p	hsa-miR-122-5p	hsa-miR-423-5p	hsa-miR-101-3p	hsa-miR-365a-3p	hsa-miR-23a-3p	hsa-miR-215-5p	hsa-miR-320a	hsa-miR-338-3p	hsa-miR-223-3p	hsa-miR-103a-3p
G	hsa-miR-331-3p	hsa-miR-142-3p	hsa-let-7d-3p	hsa-miR-126-5p	hsa-miR-19a-3p	hsa-miR-144-3p	hsa-miR-126-3p	hsa-miR-148a-3p	hsa-miR-10b-5p	hsa-miR-195-5p	hsa-miR-125b-5p	hsa-miR-192-5p
H	hsa-miR-18b-5p	hsa-miR-335-5p	hsa-miR-320d	hsa-let-7g-5p	hsa-miR-22-3p	hsa-miR-199a-3p	hsa-miR-19b-3p	hsa-miR-29a-3p	hsa-miR-150-5p	hsa-miR-30b-5p	hsa-miR-24-3p	Blank (H2O)

Figure 24. Serum/Plasma Focus microRNA PCR Panel I, 96-well plate layout.

	1	2	3	4	5	6	7	8	9	10	11	12
A	hsa-miR-502-3p	hsa-miR-339-3p	hsa-miR-409-3p	hsa-miR-154-5p	hsa-miR-155-5p	hsa-miR-140-3p	hsa-miR-92b-3p	hsa-miR-505-3p	hsa-miR-23b-3p	hsa-miR-141-3p	hsa-miR-361-5p	hsa-miR-27a-3p
B	hsa-miR-1	hsa-miR-382-5p	hsa-miR-32-5p	hsa-miR-133b	hsa-let-7d-5p	hsa-miR-133a-3p	hsa-miR-20b-5p	hsa-miR-106b-5p	hsa-miR-328-3p	hsa-miR-532-3p	hsa-miR-15a-5p	hsa-miR-532-5p
C	hsa-miR-194-5p	hsa-miR-660-5p	hsa-miR-574-3p	hsa-miR-320c	hsa-miR-130a-3p	hsa-miR-28-5p	hsa-miR-497-5p	hsa-miR-425-3p	hsa-miR-132-3p	hsa-let-7c-5p	hsa-miR-18a-5p	hsa-miR-29b-3p
D	hsa-miR-136-5p	hsa-miR-1260a	UniSp3 IPC	hsa-miR-15b-3p	hsa-miR-142-5p	hsa-miR-100-5p	hsa-miR-326	hsa-miR-362-3p	hsa-miR-421	hsa-miR-223-5p	hsa-miR-146b-5p	hsa-miR-205-5p
E	hsa-miR-339-5p	Blank (H2O)	UniSp3 IPC	mmu-miR-378a-3p	hsa-miR-425-5p	hsa-miR-454-3p	hsa-miR-874-3p	hsa-miR-193a-5p	hsa-miR-885-5p	hsa-miR-127-3p	hsa-miR-208a-3p	hsa-miR-16-2-3p
F	hsa-miR-140-5p	hsa-miR-130b-3p	UniSp3 IPC	hsa-miR-629-5p	hsa-miR-200c-3p	hsa-miR-501-3p	hsa-miR-423-3p	hsa-miR-376a-3p	hsa-miR-22-5p	hsa-miR-2110	hsa-miR-376c-3p	hsa-miR-93-3p
G	hsa-miR-144-5p	hsa-miR-210-3p	hsa-miR-199a-5p	hsa-miR-766-3p	hsa-miR-584-5p	hsa-miR-92a-3p	hsa-miR-363-3p	hsa-miR-374a-5p	hsa-miR-483-5p	hsa-miR-7-1-3p	hsa-miR-877-5p	hsa-miR-151a-3p
H	hsa-miR-28-3p	hsa-miR-324-3p	hsa-miR-136-3p	hsa-miR-15b-5p	hsa-miR-106b-3p	hsa-miR-29c-3p	hsa-miR-335-3p	hsa-miR-21-5p	hsa-miR-30e-3p	hsa-miR-543	hsa-miR-7-5p	hsa-miR-495-3p

Figure 25. Serum/Plasma Focus microRNA PCR Panel II, 96-well plate layout.

4.4.7 miRNA measurement

For calibration between PCR plate runs were used inter-plate calibrators. Each panel contain three inter-plate calibrators assay (annotated as UniSp3IPC in the plate layout, Fig. 24, 25). Each of these wells contain pre-aliquoted primer pair and DNA template, with a minimal variation from well-to-well and from plate to plate.
Inter-plate calibration (IPC) was performed manually using the IPC assay replicates as follows. For each plate was verified that the replicates had Ct standard deviation within 0.5. The average of the replicates for each plate, the overall average (average of IPC values from all plates) were calculated. The calibration factor was calculated as the differences between plate average and overall average for each plate (calibration factor= IPC plate – IPC overall). Finally, each plate was calibrated by subtracting the calibration factor from all Ct values in the plate. Threshold for expression was set as Ct<37, and miRNAs not expressed under this criteria were removed from consideration.
Quantification of relative miRNA expression was performed with the comparative Ct method using the formula: 2^-ΔΔCT where ΔCT = [(CT gene of interest-CT reference gene) sample A - (CT gene of interest-CT reference gene) sample B] by using *cel-miR-93-3p* as the reference gene.

4.5 Statistical analysis

Student's t-test was used to examine the differences between two groups and ANOVA test to evaluate differences between more groups. Differences were considered significant for P-value <0.05. Statistical analysis were performed using GraphPad Prism® 6 sofware.

5. RESULTS

5.1 Fructosamine 3-kinase screening in an Italian population of diabetic subjects

A total 150 diabetic patients and 33 healthy controls were analyzed for the FN3K gene. Diabetic subjects were recruited in different times at Policlinic Hospital of Padua and divided into two different cohorts: cohort 1 and cohort 2. Diabetic state was assessed using WHO criteria [24], according to which fasting plasma glucose level had to be more than 126 mg/dL. The mean values of HbA1c were above the reference threshold (4-6% in healthy people) [68] and correlates with presence of diabetic complications. The same criteria were used to exclude the presence of diabetes in healthy controls.

5.1.1 Analysis of FN3K promoter region in cohort 1

Cohort 1 was made up of 70 patients, 35 T1DM and 35 T2DM, and 33 healthy controls. This cohort was previously analyzed for the FN3K gene [76].
The baseline characteristics of patients and control subjects are reported in Table 5.

	T1DM	T2DM	CONTROLS
n	35	35	33
Gender (M/F)	15/20	17/18	23/10
Age (years)	40 ± 20	59 ± 20	42 ± 13
Duration disease (years)	16 ± 9	14 ± 9	/
Fasting glycemia (mmol/L)	12 ± 1	9.5 ± 3.2	4.2 ± 0.6
HbA1c (%)	8.4 ± 16	7.8 ± 14	5.6 ± 0.3

Table 5. Clinical characteristic of cohort 1. The data are presented as mean ± standard deviation.

The first screening analyzed the coding part of the FN3K gene (6 exons and corresponding intron/exon boundaries), identifying two new mutations (c.716 G>A, c. 559 C>T) and confirming the presence of some variants previously reported [76].
Here the molecular investigation of FN3K gene was completed by the analysis of its promoter region. The portion of promoter region was -470 bp with the respect of ATG. PCR conditions were optimized and the fragment was amplified following protocol described in the section “DNA amplification” of Material and Methods. PCR products were purified and both forward and reverse strands were analysed though direct sequencing.
Genotypes and allele frequencies distributions of genetic variants detected in the promoter region of FN3K gene in T1DM, T2DM and controls are reported in Table 6.

PROMOTER REGION VARIANTS		T1DM (n=35)	T2DM (n=35)	CONTROLS (n=33)
c.-385 A>G	AA	0.20	0.23	0.33
(rs3859206)	AG	0.60	0.54	0.55
	GG	0.20	0.23	0.12
	G-allele	0.50	0.50	0.39
c.-232 A>T	AA	0.17	0.20	0.27
(rs2256339)	AG	0.69	0.63	0.49
	TT	0.14	0.17	0.24
	T-allele	0.49	0.49	0.48
c.-421 C>T	CC	0.97	1.0	1.0
	CT	0.03	/	/
	TT	/	/	/
	T-allele	0.01	0	0
c.-429delATCGGAG	+/+	0.97	1.0	1.0
	+/del	0.03	/	/
	del/del	/	/	/
	del	0.01	0	0

Table 6. Genotypes and allele frequencies of FN3K promoter genetic variants identified in cohort 1.

The values for genotypes and for the rare alleles are frequencies.
T1DM, type 1 diabetes; T2DM, type 2 diabetes; Controls, healthy subjects.
rs, RefSNP ID: http://www.ncbi.nlm.nih.gov/snp/.
The ''+'' symbol indicates the wild type allele and ''del'' the allele characterized by the deletion.

The presence of two polymorphisms, the c.-385A>G (rs3859206) and the c.-232A>T (rs2256339), known to be associated with the FN3K enzymatic activity in erythrocytes [71] was evaluated.

The allelic frequencies for each polymorphism were calculated in all groups. Hardy-Weinberg equilibrium (HWE) of the identified polymorphisms for the entire population (103subjects) and for each subpopulation (T1DM, T2DM, Controls) was estimated using the $\chi 2$-test. Results were considered significant for p-values <0.05.

The statistical analysis performed on our cohort indicates that the population was in HWE for the polymorphism c.-385A>G, while for c.-232A>T, T1DM patients and total population had values above the critical threshold of 3.84, considered at a level of significance of 5%, pointing out that HWE is not respected (Table 7).

CHI-SQUARE	T1DM (n=35)	T2DM (n=35)	CONTROLS (n=33)	TOTAL POPULATION (n=103)
c.-385 A>G (rs3859206)	1,43	0,26	0,66	1,78
c.-232 A>T (rs 2256339)	4,86	2,33	0,09	4,29

Table 7. Chi-square values calculated in diabetic and control populations

Furthermore, the screening allowed the identification of two new variants, the c.-429delATCGGAG and the c.-421C>T, in one patient with T1DM, reported here for the first time.

5.1.2 Analysis of FN3K gene in cohort 2

Cohort 2 was recruited in a second time and was made up of 80 patients with T2DM. For all enrolled subjects more clinical informations were available: BMI, waist, presence of hypertension, dyslipidemia, microalbuminuria, cerebral macroangiopathy, retinopathy and peripheral arterial disease. Moreover, several biochemical parameters were measured: HbA1c, ADMA, arginin, p-mda, er-MDA, omocistein, vitamin A and vitamin E. The characteristics of patients belonging to cohort 2 are reported in Table 8.

	COHORT 2
n	80
Gender (M/F)	41/39
Age (years)	63,10 ± 7,24
BMI	30,26 ± 5,29
Waist	102,39 ± 10,85
Hypertension (Yes/No)	60/20
Dislipidemia (Yes/No)	55/25
Microalbuminria (Yes/No)	17/63
Cerebral Macroangiopathy (Yes/No)	30/50
Rettinopathy (Yes/No)	22/58
Peripheral arterial disease (Yes/No)	set-61
HbA1c (%)	7,30 ± 0,84
ADMA (µmol/l)	0,48 ± 0,11
Arginine (µmol/l)	32,41 ± 13,31
p-mda (µmol/l)	0,15 ± 0,04
er-MDA (nmol/gr Hb)	14,90 ± 4,86
Homocysteine (µmol/l)	16,22 ± 5,65
Vitamin A (µg/dl)	51,32 ± 14,81
Vitamin E (mg/dl)	1,32 ± 0,29

Table 8. Clinical characteristic of cohort 2.
Data are expressed in mean + standard deviation where applicable.

All subjects were analyzed for the entire FN3K gene, all 6 exons with corresponding intro-exon boundaries and the promoter region. DNA was amplified with specific primers following protocol described in Material and Methods. PCR products were purified and both forward and reverse strands were analysed though direct sequencing. The screening identified 13 variants within the FN3K gene, 3 of them (c.2T>A, IVS2-27A>G and c.465G>A) being never reported so far. Furthermore, variants previously identified in a single patients with T1DM (c.-421C>T, c.-429delATCGGAG) were found in one subject with T2DM (Table 9).

The allelic frequencies for each polymorphism were calculated in all groups. HWE of the identified polymorphisms was estimated using the χ2-test. Results were considered significant for p-values <0.05. The statistical analysis performed on our cohort indicates that HWE for all detected polymorphisms was respected.

PROMOTER REGION VARIANTS		T2DM (n=80)	CONTROLS (n=33)	EXONIC/INTRONIC VARIANTS		T2DM (n=80)	CONTROLS (n=33)
c.-385 A>G	AA	0,31	0,33	**c.2 T>A**	TT	0,99	1
(rs3859206)	AG	0,54	0,55		AT	0,01	/
	GG	0,15	0,12		AA	/	/
	G-allele	0,42	0,39		A-allele	0,01	/
c.-232 A>T	AA	0,27	0,27	**c.187 A>C**	AA	/	/
rs 2256339	AT	0,58	0,48	**(rs2253149)**	AC	/	/
	TT	0,15	0,25		CC	1	1
	T-allele	0,44	0,48		C-allele	0,01	0,03
c.-421 C>T	CC	0,99	1	**IVS2-27 A>T**	AA	0,99	1
	CT	0,01	/		AT	0,01	/
	TT	/	/		TT	/	/
	T-allele	0,01	/		T-allele	0,01	/
c.-429delATCGGAG	+/+	0,99	1	**IVS2+26 A>G**	AA	1	1
	+/del	0,01	/	**(rs2253132)**	AG	/	/
	del/del	/	/		GG	/	/
	del	0,01	/		A-allele	0,01	0,03
				IVS2+31 A>T	AA	/	/
				(rs2253131)	AT	0,19	0,15
					TT	0,81	0,85
					A-allele	0,09	0,08
				c.465 G>A	GG	0,99	1
					AG	0,01	/
					AA	/	/
					A-allele	0,01	/
				IVS4-9delTTG	+/+	0,74	0,67
				(rs72318398)	+/del	0,25	0,33
					del/del	0,01	/
					del	0,14	0,17
				c.900 C>G	CC	0,15	0,09
				(rs1056534)	GC	0,48	0,52
					GG	0,37	0,39
					C-allele	0,39	0,35
				c.906 C>T	CC	0,98	1
				(rs149413139)	CT	0,03	/
					TT	/	/
					T-allele	0,01	/

Table 9. Genotypes and allele frequencies of FN3K genetic variants identified in cohort 2.

The values for genotypes and for the rare alleles are frequencies.
T2DM, type 2 diabetes; Controls, healthy subjects.
rs, RefSNP ID: http://www.ncbi.nlm.nih.gov/snp/.
The ‘‘+’’ symbol indicates the wild type allele and ‘‘del’’ the allele characterized by the deletion.

5.1.3 Identification of FN3K genetic variants

The characterization of the FN3K gene though direct sequencing in both cohort 1 and 2 revealed the presence of 15 different genetic variants reported in Table 10: 4 located in the promoter region; 7 in coding regions (exons 1, 2, 4, 5, 6); while 4 in non-coding portions (introns 2 and 4).

Of these 15 variants, 5 were unpublished variants, 2 were previously published by our group [76], while the remaining 8 were polymorphisms already reported in NCBI SNPs database.

VARIANT	AA[a] CHANGE	REFSNP ID[b]
c.-232 A>T	/	rs 2256339
c.-385 A>G	/	rs 3859206
c.-421 C>T	/	/
c.-429delATCGGAG	/	/
c.2 T>A	M1?	/
c.187 A>C	R63R	rs 2253149
IVS2-27 A>T	/	/
IVS2+26 A>G	/	rs 2253132
IVS2+31 A>T	/	rs 2253131
c.465 G>A	P155P	/
IVS4-9delTTG	/	rs 72318398
c.559 C>T	R187X	/
c.716 A>G	Y239C	/
c.900 C>G	S300S	rs 1056534
c.906 C>T	G302G	rs 149413139

Table 10. List of FN3K genetic variants identified in the studied groups.

a AA:aminoacid change

b RefSNP ID: http://www.ncbi.nlm.nih.gov/snp/

5.1.4 Characterization of genetic variants

Variant c.-232 A>T

The polymorphism c.-232 A>T was identified in the promoter region (Fig. 26 A).
It was detected in heterozygous state in 92 patients (Fig. 26 B): of these 24 were T1DM, 22 T2DM of cohort 1 and 46 T2DM of cohort 2. The homozygous state was detected in 35 patients (Fig. 26 C): of these .6 were T1DM, 7 T2DM of cohort 1 and 22 T2DM of cohort 2.
The polymorphism was already described in literature (rs2256339 in the NCBI SNPs database).

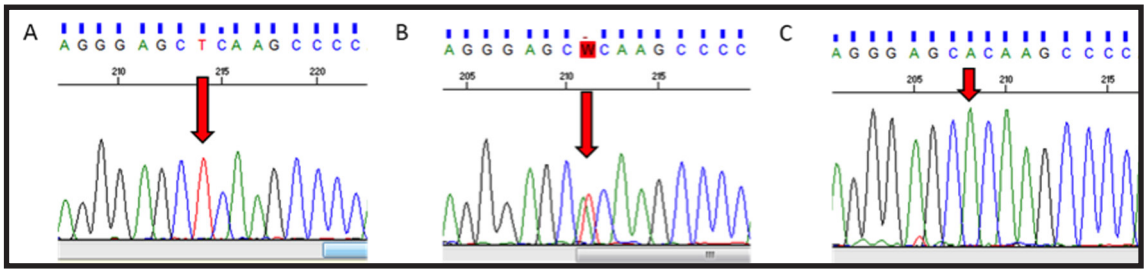


Figure 26. Position of sequencing analysis of promoter region. The red arrow shows the position of the variant c.-232 A>T: **A**) wild type (single red peak); **B**) heterozygous state (double peak); **C**) homozygous state (single green peak).

Variant c.-385 A>G

The polymorphism c.385 A>G was identified in the promoter region (Fig. 27 A).
It was detected in heterozygous state in 83 patients (Fig. 27 B): of these 21 were T1DM, 19 T2DM of cohort 1 and 43 T2DM of cohort 2. The homozygous state was detected in 40 patients (Fig. 27 C): of these 7 were T1DM, 8 T2DM of cohort 1 and 25 T2DM of cohort 2.
The polymorphism was already described in literature (rs3859206 in the NCBI SNPs database).

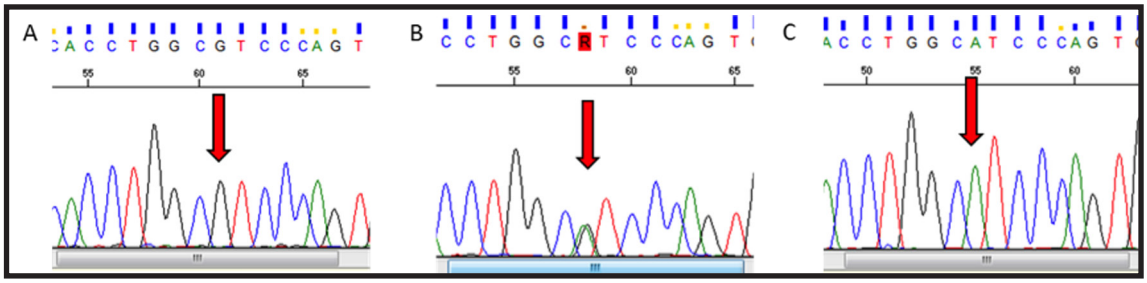


Figure 27. Position of sequencing analysis of promoter region. The red arrow shows the position of the variant c.385 A>G: **A**) wild type (single black peak); **B**) heterozygous state (double peak); **C**) homozygous state (single green peak).

Variant c.-421 A>T and Variant c.-429delATCGGAG

These variants were identified in the promoter region (Fig. 28) in two patients one with T1DM of cohort 1 and one with T2DM of cohort 2 both in heterozygous state. The first one, c.-421 A/T, caused the substitution of an A with a T at position -421; while the second one, variant c. -429delATCGGAG, consisted of a 7 nucleotides deletion.
These variants were previously unpublished.

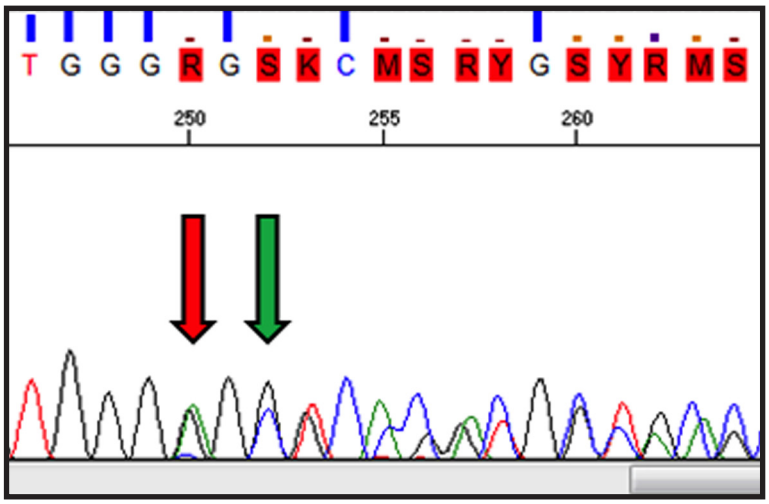


Figure 28. Position of sequencing analysis of promoter region. The red arrow shows the position of the variant c.-421 A>T. The green arrow indicates the starting point of the c. -429delATCGGAG deletion. The double peaks demonstrate the heterozygous state.

Variant c.2 T>A

Variant c.2 T>A was detected in heterozygous state in exon 1 (Fig. 29), in only a patient affected by T2DM: it changed the T forming the first triplette ATG for methionine, with an A. This variant causes the abolition of the translation starting site and probably the activation of a new starting site downstream the N-terminal region (p.Met1_Gln50del).

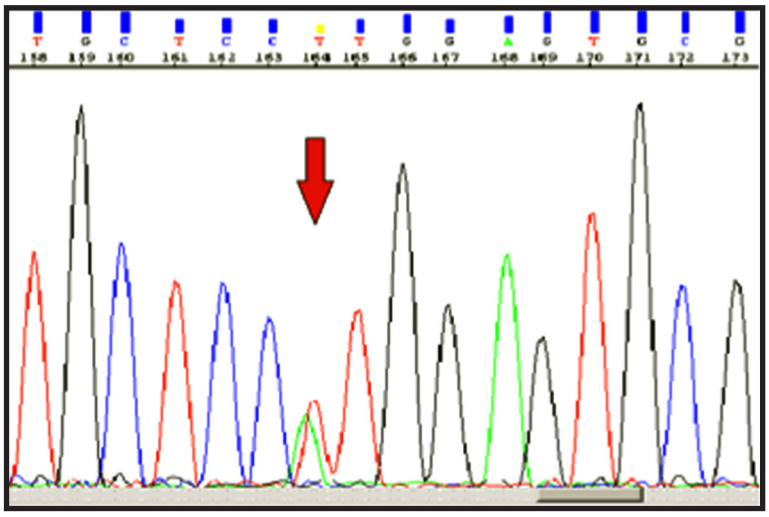


Figure 29. Position of sequencing analysis of exon 1. The red arrow shows the position of the variant c.2 T>A; the double peak demonstrates that it is present in heterozygous state (T/A).

Variant c.187 A>C

The polymorphism c.187 A>C was identified in exon 2 (Fig. 30): it didn't change the deduced sequence of the protein (p.R63R).It was present in all the diabetic patients in homozygous state (C/C).

This polymorphism was already reported in literature (rs2253149 in the NCBI SNPs database).

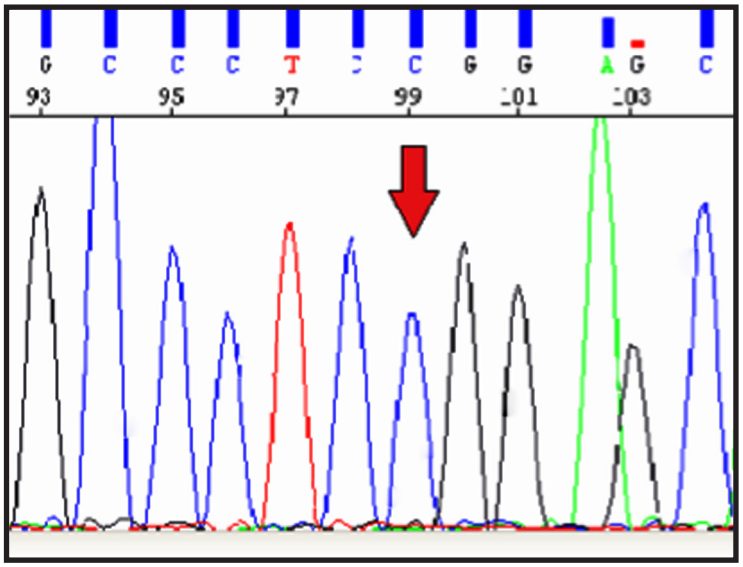


Figure 30. Position of sequencing analysis of exon 2. The red arrow shows the position of the variant c.187 A>C; the single blue peak demonstrates that it is present in homozygous state (C/C).

Variant IVS2+26 G>A in intron 2

The polymorphism IVS2+26 G>A was detected in intron 2 (Fig.31): it was presented in all diabetic subjects in homozygous state (A/A).

This polymorphism was already reported in literature (rs2253132 in the NCBI SNPs database).

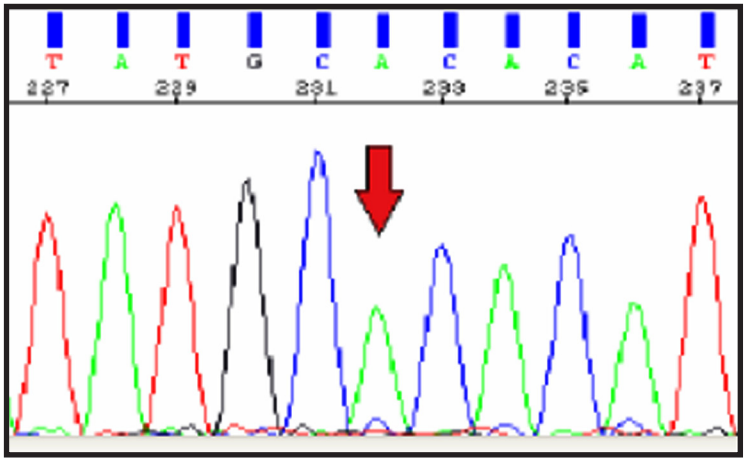


Figure 31. Position of sequencing analysis of intron 2. The red arrow shows the position of the variant IVS2+26 G>A; the green peak demonstrates that it is present in homozygous state (A/A).

Variant IVS2+31 A>T in intron 2

The polymorphism IVS2+31 A>T was identified in intron 2 (Fig. 32): it was detected in heterozygous state in 15 T2DM patients.

This polymorphism was already reported in literature (rs2253131 in the NCBI SNPs database).

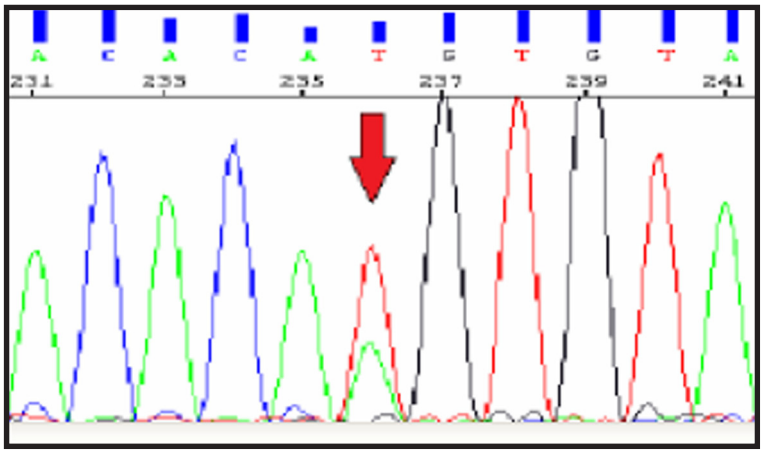


Figure 32. Position of sequencing analysis of intron 2. The red arrow shows the position of the polymorphism IVS2+31 A>T; the double peak demonstrates that it is present in heterozygous state (A/T)

Variant IVS2-27 A>T in intron 2

The variant IVS2-27 A>T was identified in intron 2 (Fig. 33) in a single patients with T2DM in heterozygous state. This variant was previously unpublished and caused the substitution of a G with an A 27 nucleotides upstream exon 3.

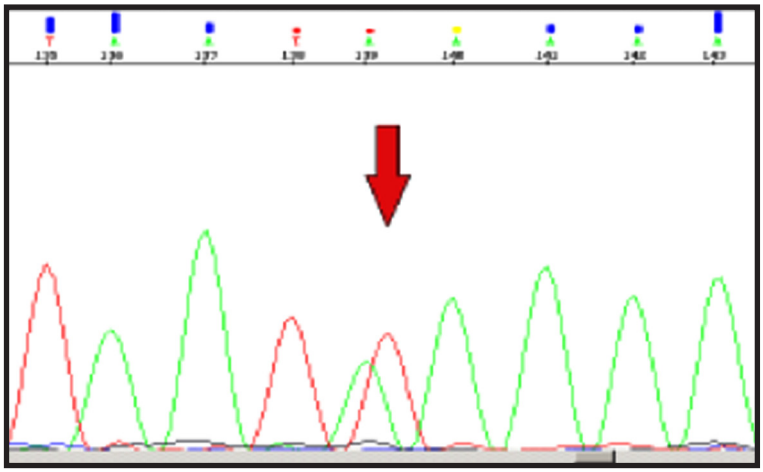


Figure 33. Position of sequencing analysis of intron 2. The red arrow shows the position of the variant IVS2-27 A>T; the double peak demonstrates that it is present in heterozygous state (T/A).

Variant c.465 G>A

The variant c.465 G>A was identified in heterozygous state in exon 4 (Fig. 34) in only one patient with T2DM. It caused a synonymous substitution at codon 155 (p.P155=) of a proline, that is located few nucleotides from the splicing site region possibly altering the splicing process. This mutation was unpublished.

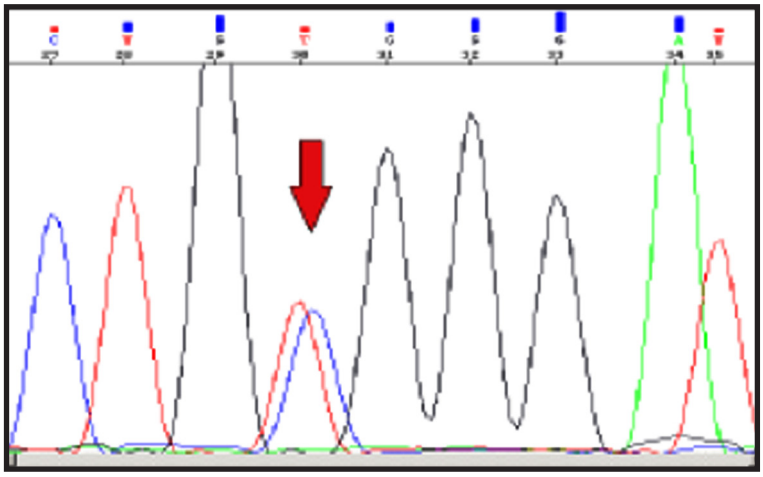


Figure 34. Position of sequencing analysis of exon 4. The red arrow shows the position of the variant c.465 G>A; the double peak demonstrates that it is present in heterozygous state (G/A).

Variant IVS4-9-11delTTG

The genetic variant IVS4-9-11delTTG was identified in intron 4 (Fig. 35 A): it consisted of a deletion of 3 nucleotides (TTG) and was found in 20 patients in heterozygous state (Fig. 35 B) and in one patient in homozygous state (Fig. 35 C).
This polymorphism was already reported in literature (rs72318398 in the NCBI SNPs database).

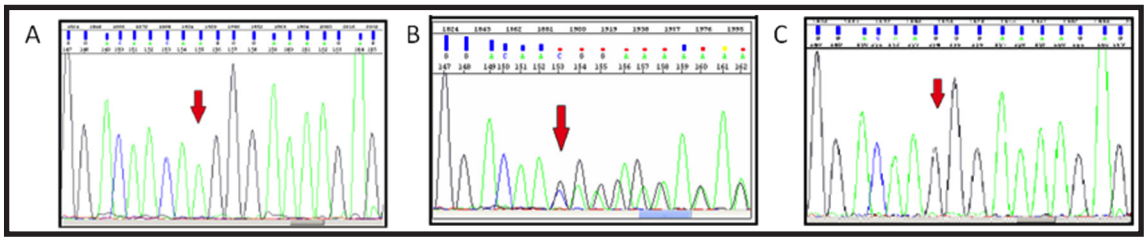


Figure 35. Position of sequencing analysis of intron 4. The variant IVS4-9-11delTTG is showed in: **A**) wild type state; **B**) heterozygous state; **C**) homozygous state. The red arrow shows the starting position of the deletion.

Variant c.559 C>T

Variant c.559 C>T was identified in exon 5 (Fig. 36) and caused a substitution of an arginine into a premature stop codon at codon 187 (p.R187*). It was found in on patient affected by T2DM (cohort 1) in heterozygous state. This variant was previously published by our group [76].

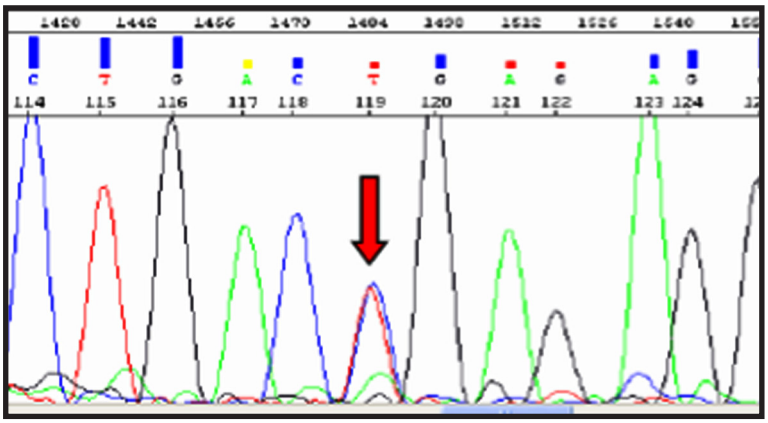


Figure 36. Position of sequencing analysis of exon 5. The red arrow shows the position of the variant c.559 C>T; the double peak demonstrates that it is present in heterozygous state (C/T).

Variant c.716 A>G

The variant c.716 A>G was detected in exon 6 (Fig. 37) and caused the substitution of a tyrosine for a cysteine at codon 239 (p.Y239C). It was found in only one subject with T2DM (cohort 1) in heterozygous state. This variant was previously published by our group [76].

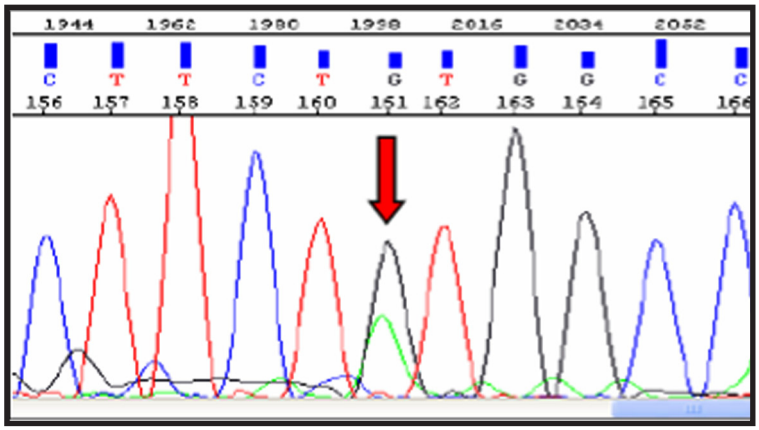


Figure 37. Position of sequencing analysis of exon 6. The red arrow shows the position of the variant c.716 A>G; the double peak demonstrates that it is present in heterozygous state (A/G).

Variant c.900 C>G

The variant c.900 C>G was identified in exon 6 (Fig. 38): it caused a synonymous substitution of serine at codon 300 (p.S300S). It was found in heterozygous state and in homozygous state respectively in the 38 and 30 patients with diabetes.
The polymorphism was already described in literature (rs1056534 in the NCBI SNPs database).

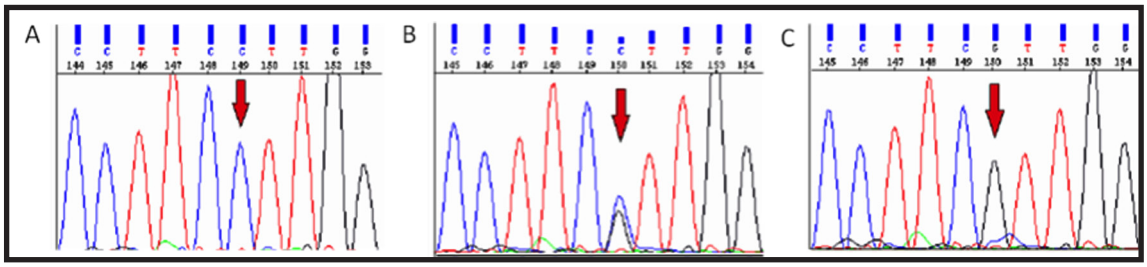


Figure 38. Position of sequencing analysis of exon 6. The red arrow shows the position of the variant c.900 C>G: **A**) wild type (single blue peak); **B**) heterozygous state (double peak); **C**) homozygous state (single black peak) .

Variant c.906 C>T

The variant c.906 C>T was identified in exon 6 (Fig. 39) and caused the substitution of a C with a T at codon 302 without changing the amino acid (G302=). This variant was found in heterozygous state in 2 diabetic subjects with T2DM.
This polymorphism was already described in literature (rs149413139 in the NCBI SNPs database).

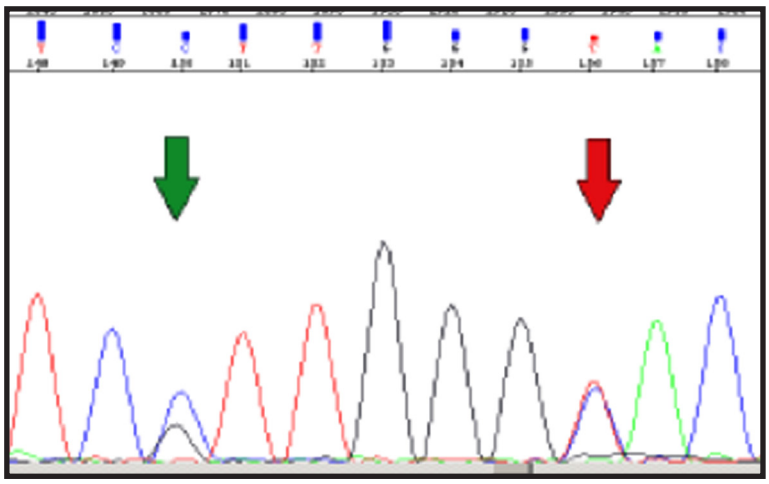


Figure 39. Position of sequencing analysis of exon 6. The red arrow shows the position of the variant c.900 C>G. The green arrow indicates the polymorphism c.900 C>T previously described. Double peaks demonstrate the heterozygous state for both polymorphisms.

5.1.5 Screening of control subject using DHPLC

DHPLC analysis was performed on 33 healthy control subjects in order to exclude presence of new mutations found in coding regions of diabetic patients: c.2 T>A, c.465 G>A. Analysis was conducted on PCR products of exons 1 and 4 on both control and diabetic subjects harboring mutations detected through direct sequencing.
A previously sequenced sample without any mutation was used as internal control. None of the mutations were never found in these control group (Fig.40).

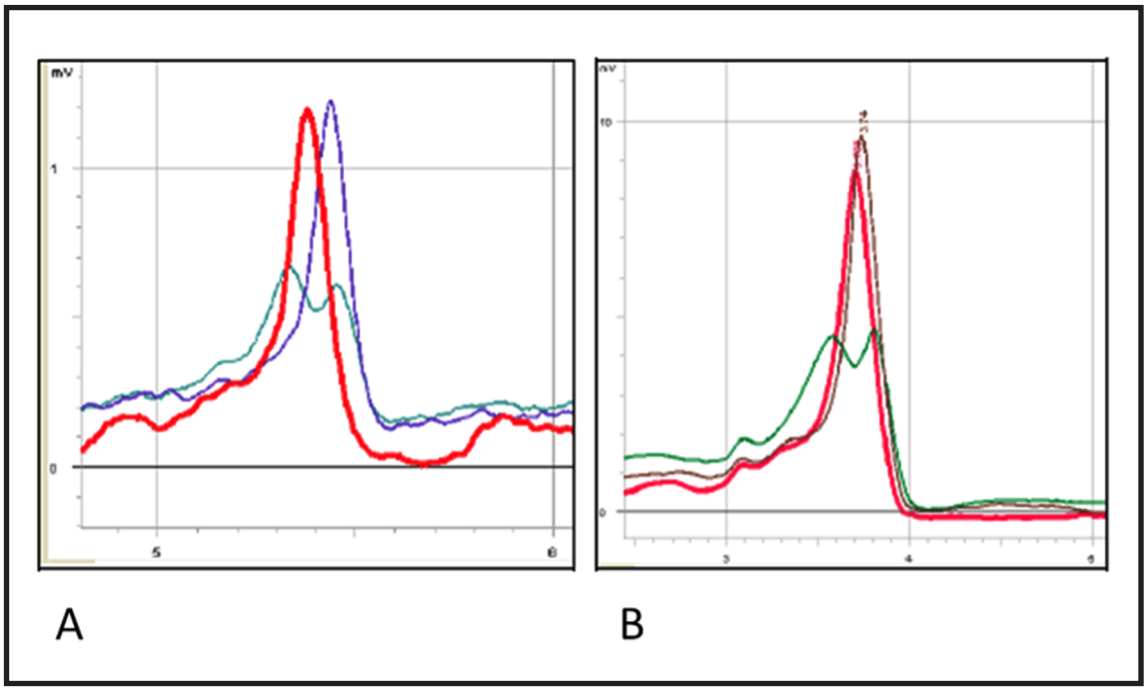


Figure 40. DHPLC chromatograms for control subjects screening. Retention times of homoduplex and heteroduplex fractions are indicated on the x-axis. **A**) Analysis of the variant c.2 T>A: green double pick represent heterozygous sample. **B**) Analysis of the variant c.465 G>A: green double pick represent heterozygous sample.

5.1.6 Genotype and phenotype correlation

Referring to the literature, a genotype composed by the combination of polymorphisms mostly associated with a variation in the enzymatic activity of FN3K in erythrocytes [71] and with a variation of HbA1c values [73] was considered. These polymorphisms are: rs3859206, rs2256339 (in the promoter region) and rs1056534 (in exon 6). T1DM subjects of cohort 1 were excluded since HWE was not respected for polymorphism rs2256339.
Thirteen genotypes have been identified and summarized in Table 11.

GENOTYPES	c.-385 A>G (rs3859206)	c.-232 A>T (rs2256339)	c.900 C>G (rs1056534)	Cohort 1 (n=35)	Cohort 2 (n=80)	CONTROLS (n=33)
A	GG	TT	CC	0,1429	0,1000	0,0303
B	GA	TA	CG	0,4286	0,2875	0,2424
C	AA	AA	GG	0,0857	0,1375	0,1212
D	AA	TA	CG	/	0,0375	0,0909
E	AA	TA	GG	0,0857	0,1000	0,0606
F	GG	TA	CG	0,0857	0,0500	/
G	GA	AA	GG	0,0857	0,0750	0,0909
H	GA	TA	GG	0,0286	0,0500	0,0606
I	GA	TA	CC	/	0,0500	/
J	AA	AA	CG	0,0286	0,0250	0,0606
K	AA	TT	GG	0,0286	0,0125	/
L	GA	TT	CG	/	0,0375	0,0606
M	GA	AA	CG	/	0,0375	/

Table 11. Genotypes and relative frequencies in the studied groups.

The values for genotypes are frequencies.
Cohort 1, (n= number of T2DM subjects);Cohort 2, (n= number of T2DM subjects); Controls, (n= healthy subjects).
rs, RefSNP ID: http://www.ncbi.nlm.nih.gov/snp/.

The genotypes GG at position -385, TT at position -232 and CC at c.900 have been associated with a better performance of the FN3K enzymatic activity [71]. Moreover, the latter has been also associated with lower HbA1c values [73].
The previous study performed on cohort 1 [76] confirmed the observation that patients carrying the CC allele for polymorphism c.900 C>G had lower HbA1c level compared with those presenting GG alleles. Adding the promoter region results, it was interesting to note that these patients presented also the GG genotype for the polymorphism -385A>G.
The same result was found also on cohort 2 and on gathering cohort 1 to cohort 2 (Total). Subjects carrying favorite alleles for the 3 mentioned polymorphisms, genotype A, displayed HbA1c values lower than those with genotype C (Fig.41).

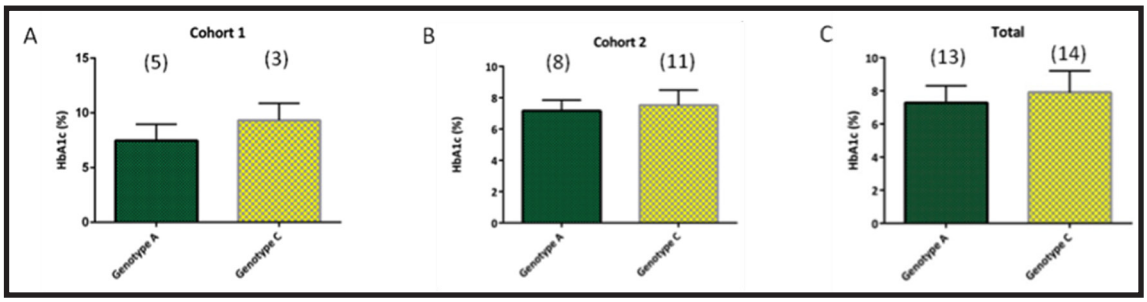


Figure 41. HbA1c concentration in Cohort 1 (**A**), Cohort 2 (**B**) and Total T2DM (**C**) subjects according to genotype A (green) and C (yellow). Genotypes composition are reported in Table 11. HbA1c values are expressed as mean + standard deviation. The number of patients for genotype is indicated in brackets.

However, the difference between glycated hemoglobin mean values and different genotypes was not statistically significant (Cohort 1 p= 0.1416; Cohort 2 p=0.3779; Total p=0.1749).

According to the presence of microvascular (microalbuminuria and rethynopathy) or macrovascular (cerebral macroangiopathy and peripheral arterial disease) complications, patients of cohort 2 were subdivided into two groups. Patients belonging to the two groups were categorized based on genotypes and HbA1c mean values were compared. Genotypes represented by a single patient have been excluded from the analysis (Fig. 42).

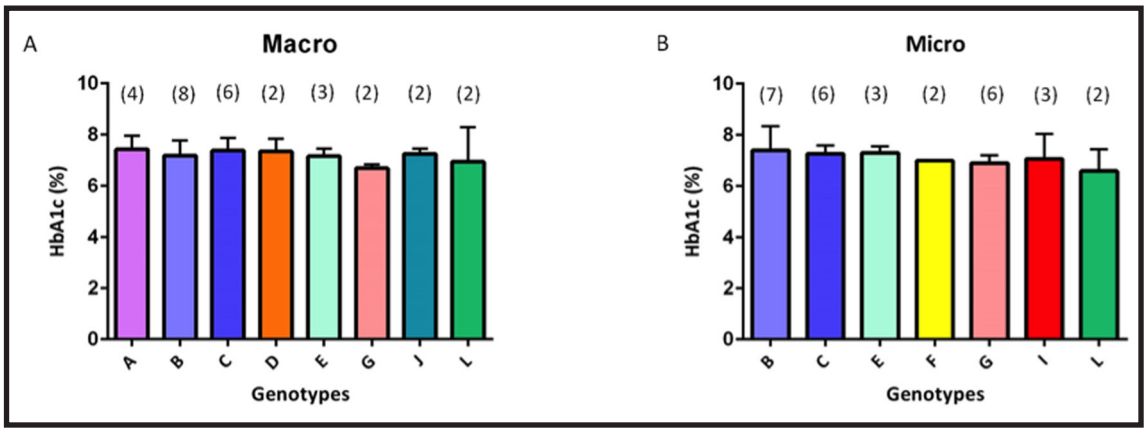


Figure 42. HbA1c concentration in Macrovascular (**A**) and Microvascular (**B**) groups according to genotype. Genotypes composition are reported in Table 11. HbA1c values are expressed as mean + standard deviation. The number of patients for genotype is indicated in brackets.

No statistically significant difference between glycated hemoglobin mean values and different genotypes have been observed using ANOVA test (Macro p= 0.6947; Micro p=0.4980). However, patients with macorvascularar complications displayed different genotypes from those with microvascular ones.

Finally, the C allele of polymorphism in exon 6 (c.900 C>G, rs1056534) has been associated with progression of diabetic nephropathy and cardiovascular morbidity and mortality [74]. Seventeen subjects belonging to cohort 2 presented nephropathy. However, only 3 patients out of the 17 with nephropathy presented the CC alleles for the polymorphism in exon 6, highlighting the absence of a significant association (p=0.7305) of the polymorphism with diabetic nephropathy in this cohort.

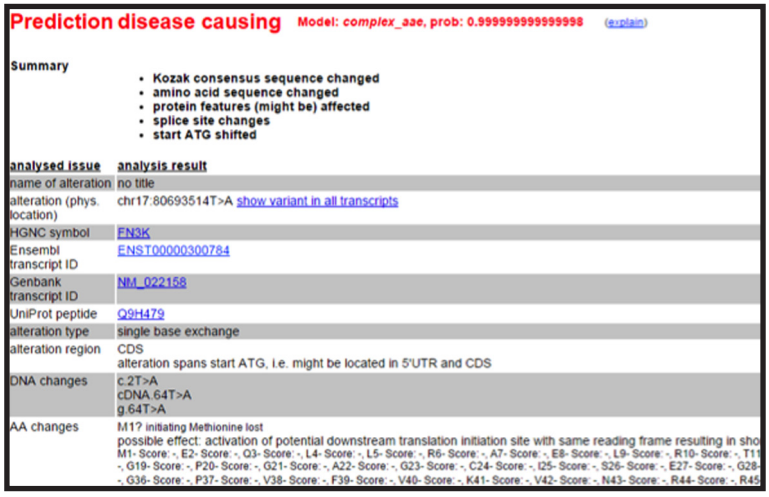
5.2 Study of new mutations identified

5.2.1 In silico analysis of new variants

New variants identified by sequencing were analysed using software Mutation Taster algorithm [127] to explore the likelihood of associated disruption of the FN3K protein structure and concomitant disease induction. Mutation Taster is a free web-based application for rapid evaluation of disease causing potential of DNA sequence alterations. MutationTaster integrates information from different biomedical databases and uses established analysis tools [127]. Analyses comprise evolutionary conservation, splice-site changes, loss of protein features and changes that might affect the amount of mRNA. Test results are then evaluated by a naive Bayes classifier, which predicts the disease potential.

The deleterious effects of missense mutations on FN3K protein were confirmed using Polyphen2 [128] and SIFT [129] *in silico* softwares.

Among 7 new genetic variants, 4 were found to be pathogenetic: c.2 T>A; c.465 G>A; c.559 C>T and c.716 A>G.



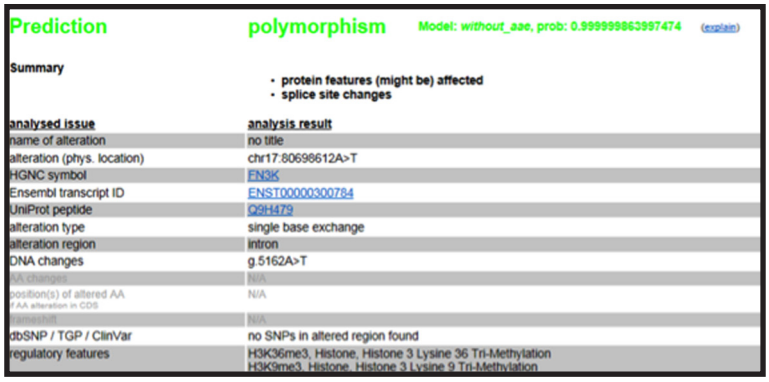
Prediction disease causing Model: *complex_aae*, prob: 0.999999999999998 (explain)

Summary

- **Kozak consensus sequence changed**
- **amino acid sequence changed**
- **protein features (might be) affected**
- **splice site changes**
- **start ATG shifted**

analysed issue	analysis result
name of alteration	no title
alteration (phys. location)	chr17:80693514T>A show variant in all transcripts
HGNC symbol	FN3K
Ensembl transcript ID	ENST00000300784
Genbank transcript ID	NM_022158
UniProt peptide	Q9H479
alteration type	single base exchange
alteration region	CDS alteration spans start ATG, i.e. might be located in 5'UTR and CDS
DNA changes	c.2T>A cDNA 64T>A g.64T>A
AA changes	M1? initiating Methionine lost possible effect: activation of potential downstream translation initiation site with same reading frame resulting in sho M1- Score: -, E2- Score: -, Q3- Score: -, L4- Score: -, L5- Score: -, R6- Score: -, A7- Score: -, E8- Score: -, L9- Score: -, R10- Score: -, T11 -, G19- Score: -, P20- Score: -, G21- Score: -, A22- Score: -, G23- Score: -, C24- Score: -, I25- Score: -, S26- Score: -, E27- Score: -, G28- -, G36- Score: -, P37- Score: -, V38- Score: -, F39- Score: -, V40- Score: -, K41- Score: -, V42- Score: -, N43- Score: -, R44- Score: -, R45

Figure 43. Mutation Taster prediction output for variant c.2 T>A.



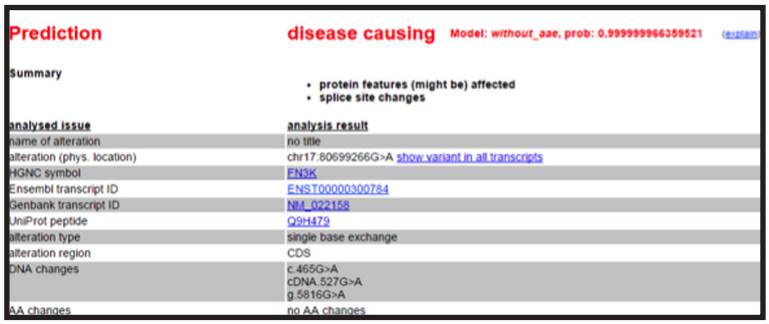
Prediction **polymorphism** Model: *without_aae*, prob: 0.999999863997474 (explain)

Summary

- **protein features (might be) affected**
- **splice site changes**

analysed issue	analysis result
name of alteration	no title
alteration (phys. location)	chr17:80698612A>T
HGNC symbol	FN3K
Ensembl transcript ID	ENST00000300784
UniProt peptide	Q9H479
alteration type	single base exchange
alteration region	intron
DNA changes	g.5162A>T
AA changes	N/A
position(s) of altered AA in AA alteration in CDS	N/A
conservation	N/A
dbSNP / TGP / ClinVar	no SNPs in altered region found
regulatory features	H3K36me3, Histone, Histone 3 Lysine 36 Tri-Methylation H3K9me3, Histone, Histone 3 Lysine 9 Tri-Methylation

Figure 44. Mutation Taster prediction output for variant IVS2-27 A>T.



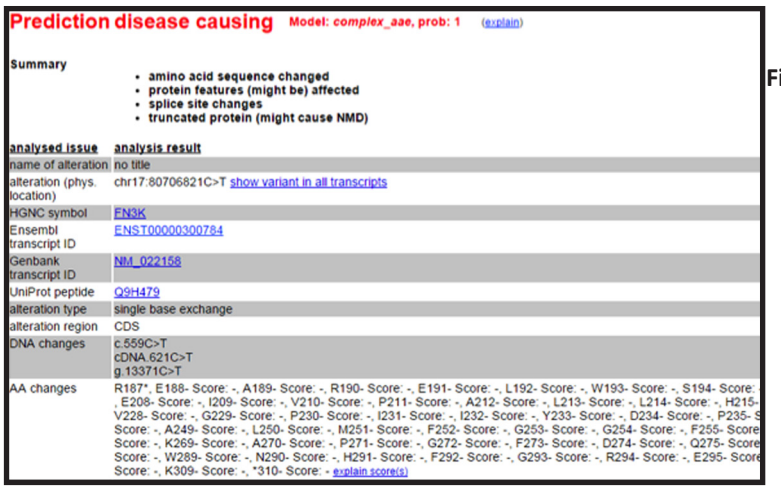
Prediction **disease causing** Model: *without_aae*, prob: 0.999999966359521 (explain)

Summary

- **protein features (might be) affected**
- **splice site changes**

analysed issue	analysis result
name of alteration	no title
alteration (phys. location)	chr17:80699266G>A show variant in all transcripts
HGNC symbol	FN3K
Ensembl transcript ID	ENST00000300784
Genbank transcript ID	NM_022158
UniProt peptide	Q9H479
alteration type	single base exchange
alteration region	CDS
DNA changes	c.465G>A cDNA.527G>A g.5816G>A
AA changes	no AA changes

Figure 45. Mutation Taster prediction output for variant c.465 G>A.



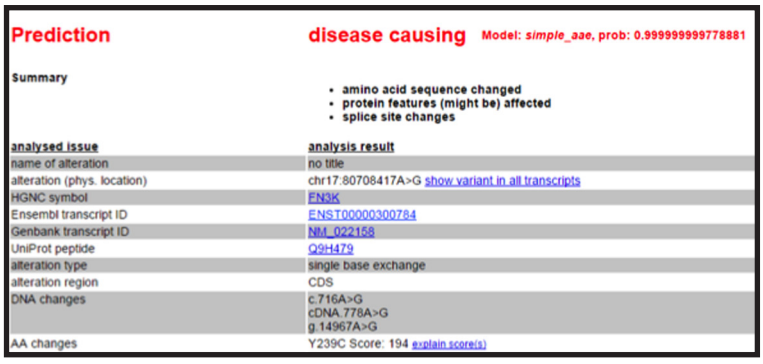
Prediction disease causing Model: *complex_aae*, prob: 1 (explain)

Summary

- **amino acid sequence changed**
- **protein features (might be) affected**
- **splice site changes**
- **truncated protein (might cause NMD)**

analysed issue	analysis result
name of alteration	no title
alteration (phys. location)	chr17:80706821C>T show variant in all transcripts
HGNC symbol	FN3K
Ensembl transcript ID	ENST00000300784
Genbank transcript ID	NM_022158
UniProt peptide	Q9H479
alteration type	single base exchange
alteration region	CDS
DNA changes	c.559C>T cDNA.621C>T g.13371C>T
AA changes	R187*, E188- Score: -, A189- Score: -, R190- Score: -, E191- Score: -, L192- Score: -, W193- Score: -, S194- Score: , E208- Score: -, I209- Score: -, V210- Score: -, P211- Score: -, A212- Score: -, L213- Score: -, L214- Score: -, H215- V228- Score: -, G229- Score: -, P230- Score: -, I231- Score: -, I232- Score: -, Y233- Score: -, D234- Score: -, P235- S Score: -, A249- Score: -, L250- Score: -, M251- Score: -, F252- Score: -, G253- Score: -, G254- Score: -, F255- Score Score: -, K269- Score: -, A270- Score: -, P271- Score: -, G272- Score: -, F273- Score: -, D274- Score: -, Q275- Score Score: -, W289- Score: -, N290- Score: -, H291- Score: -, F292- Score: -, G293- Score: -, R294- Score: -, E295- Score Score: -, K309- Score: -, *310- Score: - explain score(s)

Figure 46. Mutation Taster prediction output for variant c.559 C>T.



Prediction **disease causing** Model: *simple_aae*, prob: 0.999999999778881

Summary

- **amino acid sequence changed**
- **protein features (might be) affected**
- **splice site changes**

analysed issue	analysis result
name of alteration	no title
alteration (phys. location)	chr17:80708417A>G show variant in all transcripts
HGNC symbol	FN3K
Ensembl transcript ID	ENST00000300784
Genbank transcript ID	NM_022158
UniProt peptide	Q9H479
alteration type	single base exchange
alteration region	CDS
DNA changes	c.716A>G cDNA.778A>G g.14967A>G
AA changes	Y239C Score: 194 explain score(s)

Figure 47. Mutation Taster prediction output for variant c.716 A>G.

5.2.2 RNA analysis

In order to confirm hypothesis on pathogenicity of these four mutations (c.2 T>A, c.465 G>A, c.559 C>T and c. 716 G>A) found at genomic level, an RNA expression study was performed. All subjects carrying a predicted pathological mutation, a total of 4 subjects, have been processed. RNA was extracted from peripheral blood using Paxgene Blood RNA kit. Coding DNA (cDNA) was obtained retro-transcribing mRNA using MuLV reverse transcriptase enzyme. A cDNA previously sequenced and not carrying any mutations was used as control subject. cDNA was amplified by PCR with specific primers flanking regions containing alterations (see Materials and Methods).

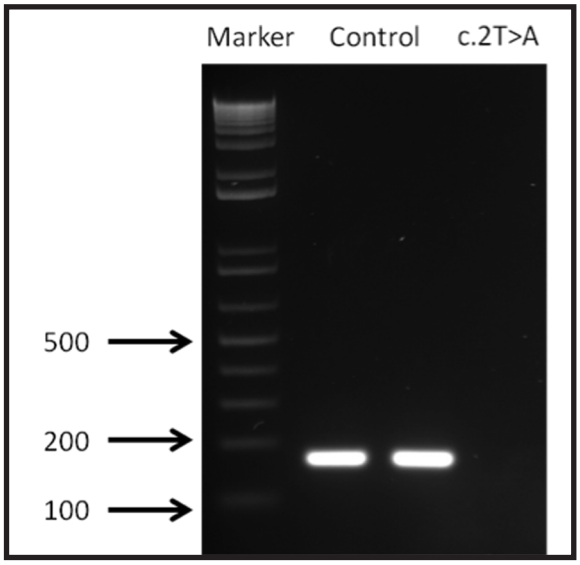


Figure 48. cDNA amplification fragment containing c.2 T>A mutation.
Length of the amplicon was 169 bp.

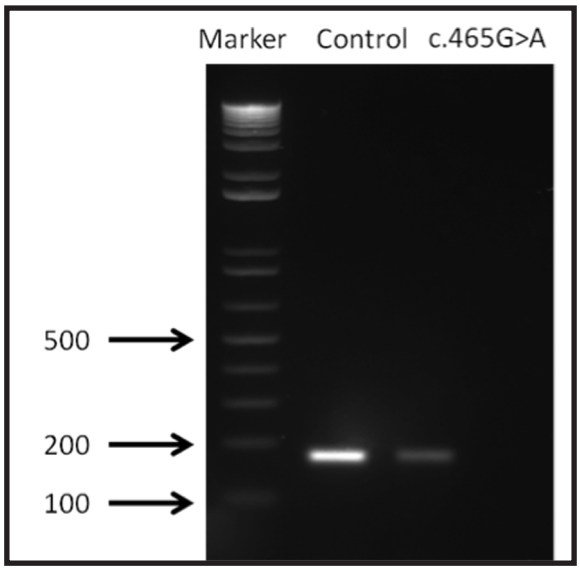


Figure 49. cDNA amplification fragment containing c.465 G>A mutation.
Length of the amplicon was 167 bp.

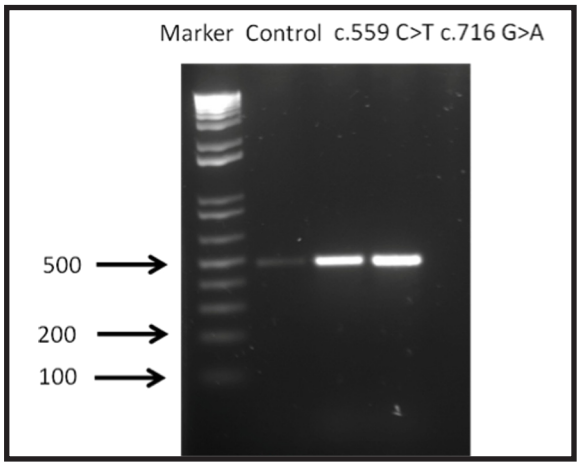


Figure 50. cDNA amplification fragment containing c.559 C>T and c.716 G>A mutations. Length of the amplicon was 487 bp.

The analysis didn't find out any changes in amplicons lengths for all variants taken into consideration compared to control sample.

5.2.3 Familial study of some new genetic variants in the FN3K gene

Family 1

Here is described the family harbouring the c.716 C>T (p.Y239C) mutation in the FN3K gene (Fig. 51). The index case (II:3) is a 61 year-old woman with a history of T2DM. Three members of the index case family were available for the genetic test: her 58 year-old brother (II:1) and two sisters, one of 59 years old (II:2) and the elder sister of 62 years old (II:4) (Fig. 51).

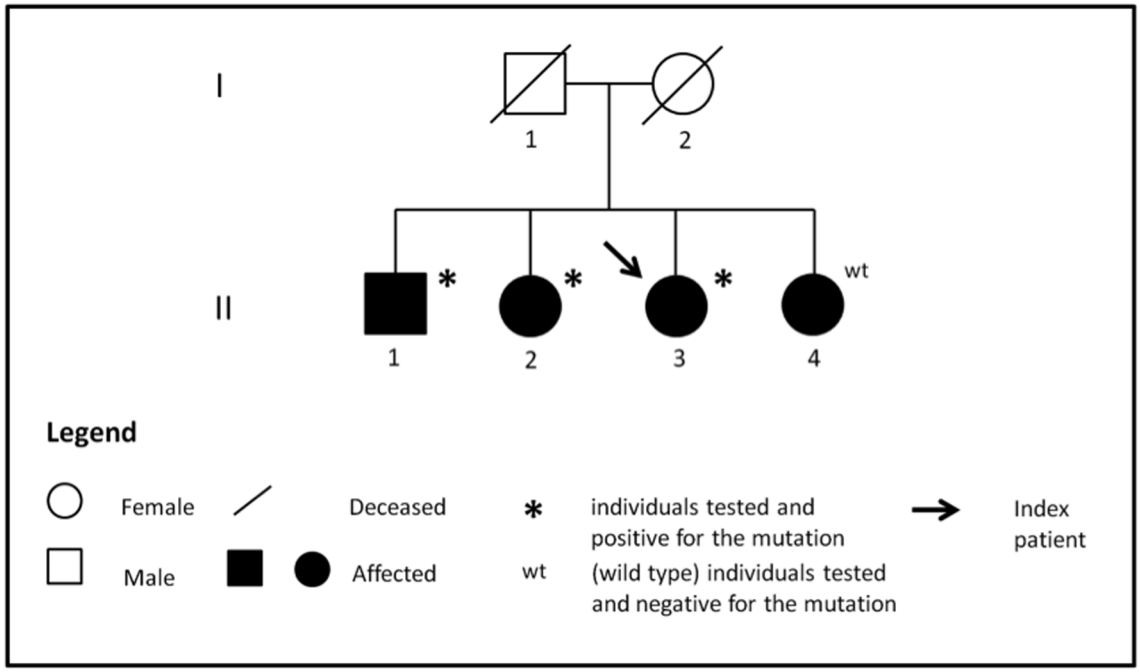


Figure 51. Pedigree of family 1 with the Y239C mutation in the FN3K gene.

All 3 subjects had been screened for the presence of the mutation c.716 C>T in exon 6. Moreover, they were screened also for the presence of the 3 polymorphisms associated with some peculiar aspects of diabetes: c.-385 G>A (rs3859206) and c. -233 T>A (rs2256339) in the promoter region, and c.900 G>C (rs1056534) in exon 6. Results of genetic testing are summarized in Table 12.

	FN3K RESULTS			
SUBJECTS	c.-385 A>G (rs3859206)	c.-232 A>T (rs2256339)	c.900 C>G (rs1056534)	c.716 A>G
II:1	GA	AA	GG	AG
II:2	GA	AA	GG	AG
II:3	GA	AA	GG	AG
II:4	GA	TT	GC	AA

Table 12. Results of FN3K molecular analysis of family 1.

Two out of three siblings carrying the mutation c.716 A>G. It is interesting to note that all subjects carrying the mutation presented the same genotypes for the other polymorphisms examined.
All subjects were clinically investigated and, except for subject II:2 that has T1DM, the remaining three presented T2DM. Clinical data are listed in Table 13.

Tests	II:1	II:2	II:3	II:4
HbA1c (%)	8.2	6.8	6.5	7.1
Fasting Glucose (mg/dL)	188	110	140	117
Cholesterol (mg/dL)	256	206	208	200
c-HDL (mg/dL)	58	83	81	106
c-LDL (mg/dL)	154	110	98	88
Creatinine (mg/dL)	0.9	0.73	0.73	0.69
Microlbuminuria	N	N	N	N
Rethinopathy	N	N	N	N
Neuropathy	N	N	N	N
Cardiovascular Disease	Y	N	N	N
Insulin Treatment	Y	Y	N	N

Table 13. Clinical characteristic of family 1.

The RNA expression study was performed on mutated subjects. As for the index case, the analysis didn't find out any change in the amplicons lengths.

Family 2

Here is described the family with the c.559 C>T (p.R187*) mutation in the FN3K gene (Fig. 52). The index case (II:2) is a 75 year-old woman with a history of T2DM. Her brother (II:1), a 66 year-old man without diabetes symptoms, gave his consent for the genetic test.

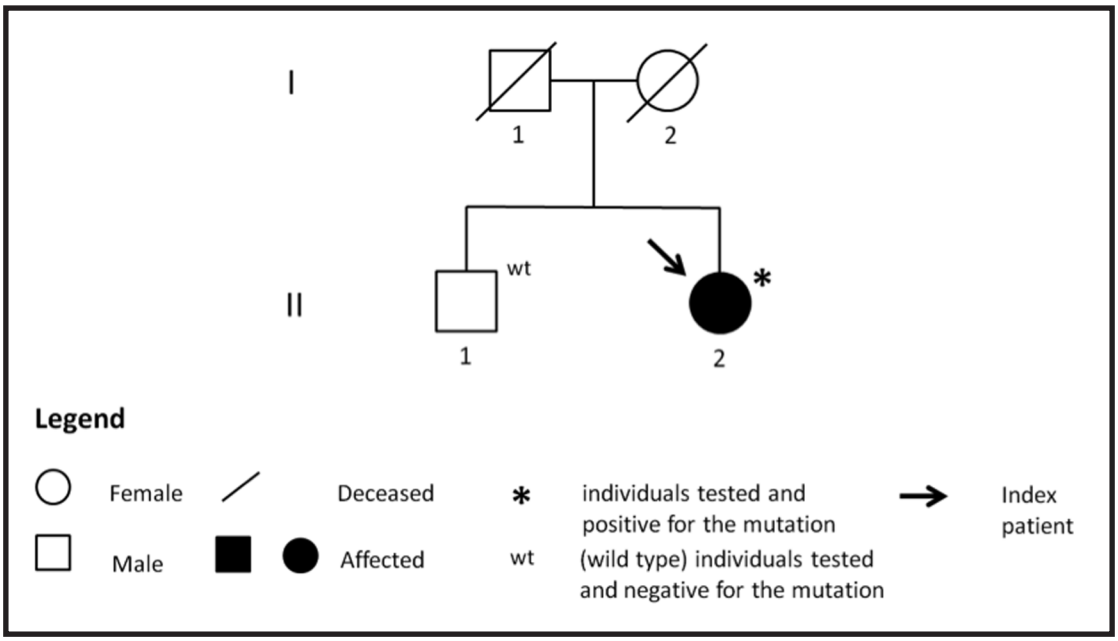


Figure 52. Pedigree of family 2 with R187* mutation in the FN3K gene.

Proband's brother has been screened for the presence of the mutation c.559 A>G in exon 5. Moreover, we tested him also for the presence of the 3 polymorphisms associated with peculiar aspects of diabetes: c.-385 G>A (rs3859206), c.-233 T>A (rs2256339) and c.900 G>C (rs1056534). Results of genetic testing are summarized in Table 14.

	FN3K RESULTS			
SUBJECTS	c.-385 A>G (rs3859206)	c.-232 A>T (rs2256339)	c.900 C>G (rs1056534)	c.559 C>T
II:1	GG	TT	CC	CC
II:2	GG	TT	CC	CT

Table 14. Results of FN3K molecular analysis of family 2.

FN3K analysis pointing out that the proband's brother (II:1) didn't have the mutation c. 559 C>T. However, in this family, is interesting to note that siblings presented the same genotypes for the 3 polymorphisms, even if only the proband (II:2) has diabetes. All subjects were clinically investigated. Clinical data are listed in Table 15.

Tests	II:1	II:2
HbA1c (%)	5.5	8.8
Fasting Glucose (mg/dL)	105	162
Cholesterol (mg/dL)	240	164
c-HDL (mg/dL)	72	53
c-LDL (mg/dL)	153	71
Creatinine (mg/dL)	1.07	0.92
Microlbuminuria	N	N
Rethinopathy	N	Y
Neuropathy	N	N
Cardiovascular Disease	N	N
Insulin Treatment	N	Y

Table 15. Clinical characteristic of family 2.

5.3 Plasma profile of microRNAs after soy supplementation in patients with T2DM and subclinical hypogonadism

A total of 10 male T2DM subjects, with a total testosterone level ≤12 nmol/L and compensated hypogonadism, were analyzed to identify the differentially expressed miRNA in the peripheral plasma samples before and after treatment with soy and phytoestrogens.

All patients were under active treatment with soy and isoflavons for twelve week and were divided into two groups: before (visit1) and after administration (visit 2). Plasma samples fromT2DM patients were collected on both occasions. Clinical characteristics of the individuals included in the analysis are summarized in Table 16.

Tests (Reference intervals)	**Visit 1**	**Visit 2**
Fasting Glucose, mmol/L (3.9 - 5.8)	7.5 ± 2.5	7.3 ± 2.2
HbA1c, mmol/mol (20 - 42)	50.9 ± 11.2	54.8 ± 15.7
Cholesterol, mmol/L (4.9 - 7.1)	3.8 ± 0.6	3.5 ± 0.6
Triglycerides, mmol/L (0.6 - 2.9)	1.1 ± 0.2	0.9 ± 0.2
c-HDL, mmol/L (0.8 -1.9)	1.2 ± 0.3	1.1 ± 0.2
c-LDL, mmol/L (2.5 - 5.4)	2.2 ± 0.7	1.9 ± 0.6
C-Reactive Protein, mg/L (0.0 - 0.8)	3.9 ± 5.5	1.4 ± 1.9
Hemoglobin, g/dL (12.6 - 17.4)	14.9 ± 1.2	13.9 ± 1.0
White Cell Count, 10^9/L (4.5 - 11.0)	6.4 ± 1.9	6.3 ± 1.7
Platelets, 10^9/L (150 - 400)	238 ± 85.7	265.9 ± 119.1
Sodium, mmol/L (135.0 - 148.0)	138 ± 1.6	139 ± 2.4
Potassium, mmol/L (3.7 - 4.6)	4.32 ± 0.2	4.51 ± 0.4
Chloride, mmol/L (100.0 - 109.0)	103 ± 1.2	102.8 ± 2.0
Bicarbonate, mmol/L (22 - 26)	28.2 ± 2.0	28.7 ± 2.0
Urea, mmol/L (2.8 - 7.7)	4.8 ± 0.5	4.7 ± 0.6
Creatinine, µmol/L (71.0 - 124.0)	72.1 ± 11.8	66.8 ± 9.2
Calcium, mmol/L (2.2 - 2.6)	2.2 ± 0.1	2.3 ± 0.1
Phosphate, mmol/L (0.7 - 1.2)	1.1 ± 0.1	1.1 ± 0.2
Bilirubin, µmol/L (3.0 - 19.0)	13.3 ± 3.6	12.2 ± 2.3
Alkaline Phosphatase, U/L (32 - 92)	59.3 ± 13.1	57.4 ± 15.5
Albumin, g/L (32 - 46)	40.2 ± 2.1	38.7 ± 2.4
Total Protein, µmol/L (63.0 - 80.0)	64.4 ± 2.7	65.3 ± 2.2
TSH, mU/L (0.5 - 8.9)	1.7 ± 0.6	3.5 ± 1.1
Free Thyroxine, µg/dL (5 - 10)	12.2 ± 1.2	10.5 ± 1.1
Free Triiodothyronine, pmol/L (3.0 - 5.2)	4.5 ± 0.4	4.5 ± 0.3
Insulin, µU/ml (0.7 - 25.0)	16.7 ± 12.8	9.0 ± 4.6
Prolactin, mIU/L (48.0 - 419)	125 ± 63.5	136.4 ± 87.1
DHEAS, µmol/ml (3.9 - 14.9)	3.0 ± 2.3	2.3 ± 2.0
FSH, IU/L (1.42 - 15.4)	5.1 ± 2.6	5.1 ± 3.0
LH, IU/L (1.24 - 7.8)	3.0 ± 1.4	3.6 ± 1.5
Daidzein ng/mL	1.8 ± 0.9	30.2 ± 31.0
Genistein ng/mL	3.0 ± 2.2	92.5 ± 80.1

Table 16. Clinical characteristics of the study participants (n=10 adult male subjects with T2DM, mean age 66.56 ± 10.92). The data are presented as mean ± standard deviation.

Based on the qRT-PCR platform, the Exiquon miRCURY Serum/Plasma Focus microRNA PCR Panels I+II (V6.0) were used for miRNAs detection.

Each miRNA with cycle threshold (Ct) value less than 37 in the panel were included in data analysis. Furthermore, fold changes >2 or below -2 where considered for further investigations.

The internal spike-in (*cel-miR-93-3p*) to normalize samples, was not detected in one patient, thus it was excluded from clinical records.

Totally, 77% of miRNAs in panels belonging to patients before starting soy supplementation, and 81% of miRNAs in patients after treatment were detected. As given in Table17, 57 circulating miRNAs differed for about 2 fold between visit 1 and visit 2, including 7 up-regulated and 50 suppressed miRNAs. Among them miR-34a-5p, miR-144-3p and miR-19b-3p showed a P-value<0.05. Other three miRNA showed a statistically significant difference (miR-10b-5p, miR-193a-5p and miR-20b-5p), however their fold change is less than 2.

MicroRNA	Fold change	P-value	Up/Down regulation	MicroRNA	Fold change	P-value	Up/Down regulation
hsa-miR-34a-5p	7.92	0.05	Up	**hsa-miR-139-5p**	-2.81	ns	Down
hsa-let-7e-5p	3.43	ns	Up	**hsa-miR-155-5p**	-2.88	ns	Down
hsa-miR-144-3p	2.41	0.04	Up	**hsa-miR-130a-3p**	-2.93	ns	Down
hsa-miR-122-5p	2.41	ns	Up	**hsa-miR-154-5p**	-3.03	ns	Down
hsa-miR-19b-3p	2.21	0.03	Up	**hsa-miR-454-3p**	-3.27	ns	Down
hsa-miR-451a	2.10	ns	Up	**hsa-miR-335-3p**	-3.35	ns	Down
hsa-miR-101-3p	2.03	ns	Up	**hsa-miR-210-3p**	-3.36	ns	Down
hsa-miR-223-5p	-2.01	ns	Down	**hsa-miR-140-5p**	-3.42	ns	Down
hsa-miR-22-5p	-2.01	ns	Down	**hsa-miR-324-5p**	-3.55	ns	Down
hsa-miR-425-3p	-2.03	ns	Down	**hsa-miR-28-3p**	-3.57	ns	Down
hsa-miR-382-5p	-2.05	ns	Down	**hsa-miR-181a-5p**	-3.59	ns	Down
hsa-let-7f-5p	-2.08	ns	Down	**hsa-miR-766-3p**	-3.75	ns	Down
hsa-miR-195-5p	-2.13	ns	Down	**hsa-miR-328-3p**	-3.82	ns	Down
hsa-miR-146b-5p	-2.17	ns	Down	**hsa-miR-409-3p**	-3.88	ns	Down
hsa-miR-18a-5p	-2.18	ns	Down	**hsa-miR-106b-3p**	-4.20	ns	Down
hsa-miR-30a-5p	-2.25	ns	Down	**hsa-miR-133b**	-4.32	ns	Down
hsa-let-7d-5p	-2.29	ns	Down	**hsa-miR-301a-3p**	-4.77	ns	Down
hsa-miR-652-3p	-2.34	ns	Down	**hsa-miR-132-3p**	-4.78	ns	Down
hsa-miR-376c-3p	-2.34	ns	Down	**hsa-miR-145-5p**	-5.02	ns	Down
hsa-miR-136-3p	-2.43	ns	Down	**hsa-miR-144-5p**	-5.05	ns	Down
hsa-miR-128-3p	-2.44	ns	Down	**hsa-miR-339-5p**	-5.31	ns	Down
hsa-miR-320d	-2.46	ns	Down	**hsa-miR-532-3p**	-5.80	ns	Down
hsa-miR-151a-3p	-2.51	ns	Down	**hsa-miR-136-5p**	-6.35	ns	Down
hsa-miR-200c-3p	-2.52	ns	Down	**hsa-miR-421**	-6.70	ns	Down
hsa-miR-125a-5p	-2.54	ns	Down	**hsa-miR-28-5p**	-6.89	ns	Down
hsa-miR-374b-5p	-2.63	ns	Down	**hsa-miR-543**	-8.57	ns	Down
hsa-miR-501-3p	-2.63	ns	Down	**hsa-miR-152-3p**	-22.95	ns	Down
hsa-miR-93-3p	-2.78	ns	Down				

Table 17. Differential expression of circulating miRNAs in T2DM and compensated hypogonadism subjects.

6. DISCUSSION

In the last few decades the notion of diabetes has widened, ascertain that diabetes is a much more heterogeneous disease than the present subdivision into type 1 and type 2 diabetes assumes. Both type 1 and type 2 diabetes seem to result from a collision between genes and environment. The change toward a more affluent Western lifestyle that has taken place during the last 50 years has started a worldwide epidemic increase in the prevalence of T2DM and obesity [7][130].

All forms of diabetes are characterized by chronic hyperglycemia. Since the early 20th century, the diagnosis of diabetes has been based on the measurement of glucose concentrations in the blood [24]. In the past 25 years the measurement of HbA1c has been interpreted as routine useful integrated measure of glycemic control. However, several studies have highlighted the limits of its use in patients with underlying disorders, including diseases changing erythrocytes turnover (hemolytic anemias, chronic malaria, major blood loss or blood transfusions), as well as genetic hereditary anemias and iron storage disorders that may influence the variability of HbA1c in populations [131]. Furthermore, it has been shown that HbA1c values may not be constant among individuals despite the presence of similar blood glucose or fructosamine concentration [132].

Monitoring of diabetic status is considered as a cornerstone of diabetes care. Over the past four decades, major advances have been done in the knowledge of the human genome, and the use of molecular biology tools has brought novel insights in the pathogenesis of human diseases and diabetes. However, mechanisms other than genetic ones are also involved in the pathophysiology of diabetes. This applies for the non-enzymatic post-translational modifications, which alter structural and biological properties of proteins and living organisms.

Non-enzymatic glycation has been strongly related to hyperglycemia conditions and therefore to chronic complications associated with diabetes. Protein glycation is a spontaneous reaction where reducing sugars interact with amino groups of proteins, which, as all spontaneous reactions, cannot be prevented [133]. The reaction is initiated by the formation of a Shiff base, which then undergoes an Amadori rearrangement to become a ketoamine, fructosamine when glucose is the reactive sugar [134]. Until recently, the formation of fructosamines was thought to be irreversible, which unavoidably followed by their slow conversion into a series of compounds known as AGEs [64].

These materials accumulate in tissue and plasma during aging and diabetes [40], and numerous studies have shown an association between AGEs and many age- and diabetes-related pathologies, including atherosclerosis, cataract, nephropathy, retinopathy, and Alzheimer's disease [135].
Thus, the identification of an enzyme, FN3K, able to decompose fructosamine 3-phospates to 3-deoxyglucoseone, inorganic phosphate and an amine [61] opened the perspective that fructosamines can physiologically be removed from proteins. For this reason, FN3K appear as a part of a protein repair system and/or cellular defense [136] opposing to the consequences of hyperglycemia, controlling non-enzymatic glycation of proteins and therefore protecting from diabetic complications development.

Only Delpierre et al. proved a possible association between the variability of erythrocytes FN3K activity and genetic polymorphisms. This study, conducted on a Belgian cohort of T1DM patients, demonstrated the presence of 6 different polymorphisms: two in the promoter region and one in exon 6 of the gene, were significantly associated with the erythrocytes FN3K activity [71].

6.1 Genetic variations

In this study, the FN3K gene was analyzed by direct sequencing in a population of Italian diabetic subjects. Then the presence of variants found in this first step of the work was verified in the control healthy group by using DHPLC, a less expensive and time consuming method that is successfully used for the analysis of already known genetic variants [125].

Fifteen different genetic variants were identified: 4 in the promoter region; 7 in coding regions (exons 1, 2, 4, 5, 6); while 4 located in non-coding portions (introns 2 and 4); the detection of the latter was possible since also the intron/exon boundaries of FN3K gene were PCR amplified. Of these 15 variants, 5 were unpublished variants, 2 were previously published by our group [76], while the remaining 8 were polymorphisms already reported in NCBI SNPs database.
Among 7 new genetic variants, 4 were predicted as pathogenic: c.2 T>A; c.465 G>A; c.559 C>T and c.716 A>G using the Mutation Taster tool. The deleterious effects of these mutations on FN3K protein were confirmed using Polyphen2 [128] and SIFT [129] *in silico* softwares.

Variant c.2 T>A

The variant c.2 T/A affected the start codon in exon 1, resulting in the elimination of the amino acid methionine at position 1 (p.M1?). The abolition of the translation starting site threatening protein production, that resulted in a complete absence of the FN3K enzyme encoded by the allele carrying the mutation. However, software analysis predicts the presence of another ATG downstream the N-terminal region. This new translation starting site is in the same Open Reading Frame of the original ATG. As result, the new protein produced from the new starting site will be identical to the former protein but will lack of 50 amino acid (p.Met1_Gln50del).

The ATG triplette coding for first methionine, essential condition for protein translation, has to be located in a peculiar nucleotidic contest, namely Kozak sequence (gccRccATGG: nucleotide in capital letter are those majorly conserved; the R represent a purine). This sequence allowed the ribosome correct recruitment on the mRNA. The new ATG didn't present a similar contest around, but the protein translation from this site cannot be excluded. The mutated protein will lack of 50 amino acids in the N-terminal region (16% less of the wild type protein). Nevertheless, the catalytic site of the enzyme is located in the C-terminal region, thus, this mutated enzyme could equally exert its function if translated.

Variant c.465 G>A

The variant c.465 G>A caused the substitution of the third nucleotide of CCG triplette, coding a proline at codon 155, with an A, leading to the formation of a CCA triplette encoding for the same amino acid (p.P155=). This synonymous mutation is located few nucleotides from the splicing site region.

The pathological effect of this variant lies in the fact that it could alters the splicing process potentially affecting the mRNA maturation. In details, the splicing site located 3 nucleotides downstream the alteration will be abolished, leading to intron retention. An aberrant mRNA and a protein with an altered function would be produced or the mutated allele could not produce the protein.

Variant c.559 C>T and variant c.716 A>G

The variant c.559 C>T caused the substitution of an arginine into a premature stop codon at codon 187 (p.R187*). This premature signal for termination of translation causes a resulting protein that is abnormally shortened, lacking of part of exon 5 and all exon 6. Amino acid homologies studies between human FN3K and homologous

proteins from different organisms revealed that the most significant regions of amino acid identified are located in the carboxy termini of the proteins and that this region probably represents the ATP binding site [63]. In the case of human FN3K, this important region corresponds to exon 6 that can be therefore considered fundamental for the proper function of the enzyme. For this reason, its absence caused by mutation p.R187* is probably connected with an altered enzymatic activity. In addition we cannot exclude that the truncated protein is degraded or not targeted to its final destination.

Also the second mutation (c.716 A>G; p.Y239C) has got an important functional meaning. It is localized in a conserved region of exon 6 and involves the substitution at codon 239 of a tyrosine into a cysteine, an aminoacid containing sulfur implicated in the disulfur bonds formation which could allow the effects of structural changes within the protein and therefore an alteration of the normal enzymatic activity.

Additional three new variants have been identified: two in the promoter region (c.-421 C>T and c.-429delATCGGAG) and one in intron 2 (IVS2-27 A>G).

The intronic variant IVS2-27 caused a substitution of an A with a T, 27 nucleotides upstream exon 3. It was classified as a polymorphism by Mutation Taster. The software predicted an ineffective role of the variant in influencing the slicing since it isn't located in a consensus sequence for splicing.

Regarding the FN3K promoter region, it has been previously examined and described [63]. FN3K core promoter region was identified 252 bp upstream the transcription start site. Its sequence is consistent with those of housekeeping genes: it lacks of consensus TATA- and CAAT- boxes and it is rich in GC content, consistent with the hypothesis that FN3K gene could be co-ordinately regulated by CpG island-driven core promoters. Moreover, FN3K promoter contains two possible binding sites for NFκB transcription factor (-53 and -58 bpwith the respect of ATG) [63]. The 2 variants detected in the promoter region, c.-421 C>T and c.-429delATCGGAG, were found in 2 subjects, with T1DM and T2DM, both in heterozygous state. The first one, caused the substitution of an A in a T at position -421 (c.-421 A>T); while the second one, consisted in a 7 nucleotides deletion (variant c. -429delATCGGAG). These variants are located very far from transcription factor sites and the fact that are present in more than one subject leading to the hypothesis that such variants are ineffective polymorphisms more than mutations.

The other genetic variants identified in our study were 8 polymorphisms already described and present in the NCBI SNPs database: 2 are located in the promoter region (c.-232A>T and c.-385 A>G); 1 in exon 2 (c.187 A>C); 2 inintron 2 (IVS2+26 G>A and IVS2+31 A>T); 1 in intron 4 (IVS4-9-11delTTG), consisting in a deletion of three nucleotides very closed to the consensus sequence for splice sites, and 2 in exon 6 (c.900 C>G and c.906 C>T). It is reported that more than 90 % of SNPs are in introns or in inter-genic regions, and most of them are probably of little functional consequence. However, SNPs in promoter regions can alter the affinity of DNA-binding proteins and modify the level of gene expression; other SNPs in exon/intron boundaries result in intron retention or exon skipping, thus profoundly changing the structure of the resulting protein. It is also possible that the intron where the SNP is located can function as an RNA interference element [137].

None of the polymorphisms identified in this study resulted in a modification of the sequence of the encoded protein or was present in a consensus sequence for a splicing site or in a site for transcription factor, but a possible modulatory role in gene regulation or expression or in differential splicing[138] cannot be excluded. This hypothesis is supported by Delpierre's findings [71]: the differences in erythrocyte FN3K activity are due to sequence variations in or near the FN3K gene, which don't cause any change in the sequence of the protein. To this respect, polymorphism c.900 C>G is of particular interest, since it is the only exon-encoded polymorphism found to be significantly associated with the erythrocyte FN3K enzymatic activity [71]; in particular, the presence of the non ancestral allele C seems to cause its increase. An effective role of polymorphism c.900 C>G in diabetes has also been demonstrated in another work conducted by an Hungarian research group [73]. This polymorphism was analyzed in T2DM and in the case of the CC allelic combination, T2DM was diagnosed at a later age than in the case of CG or GG genotypes. Moreover, diabetic individuals with the CC genotype had lowerHbA1c levels compared with the others, although the N-terminal valine of the β-chains that forms HbA1c is reported as a poor substrate for FN3K [66]. Moreover, further evidences on an effective association of such polymorphism to diabetes came to light in recent years. Indeed, an association of c.900 C>G with the progression of diabetic nephropathy and cardiovascular morbidity and mortality was reported [74]. Recently, Škrha and co-workers, linked the GG genotype of polymorphism c.900 C>G to a decrease of sRAGE concentration. Therefore the genotype CC for the polymorphism c.900 C>G seems to have a protective role in T2DM. In our opinion, these findings are very useful for our research since they link the potential impact of FN3K on HbA1c levels and typical aspect of diabetes.

6.2 Association studies

Hemoglobin A1C is formed by the non-enzymatic attachment of glucose to the N-terminal valine of the β-chain of hemoglobin [139]. HbA1C reflects long-term glycemic exposure, representing the average glucose concentration over the preceding 8–12 weeks [121][140].

The A1C assay is the “gold standard” for the diagnosis of diabetes. Its formation is directly proportional to the glucose concentration to which erythrocytes are exposed during circulation, since they are permeable to glucose [141]. The life span of erythrocytes is 120 days, however, glycation of protein with a much longer half life, could lead to their inactivation and aggregation thought AGEs formation. This mechanism has been hypothesized to be one of the causes of diabetic complications development. FN3K with its deglycating function would be able to remove glycation product from proteins suggesting a protective role in the development of such complications [66].

However, variants in the FN3K gene could alter its enzymatic activity and consequently the progression of diabetic complications. Three polymorphisms have been identified in this respect: two in the promoter region (rs3859206, rs2256339) and one in exon 6 (rs1056534). In details, genotypes GG at position -385, TT at position -232 and CC at c.900 have been associated with a better performance of the FN3K enzymatic activity [71]. Moreover, the latter has been also associated with lower HbA1c values [73].

For this reason a genotype composed by the combination of these polymorphisms has been considered. Thirteen genotypes have been identified (Table 11) and HbA1c mean values for each genotype have been compared (Fig. 41). It has been observed that subjects carrying the favorite alleles for the tree polymorphisms, indicated by genotype A, displayed HbA1c values lower than those with genotype C. However, this difference was not statistically significant (Cohort 1 p= 0.1416; Cohort 2 p=0.3779; Total p=0.1749).

This genotype, even the small number of patients analyzed, seems to be associated with a lower concentration of HbA1c values in T2DM patients, consistent with the results of Mohàs et al. [73] regarding the rs1056534. The finding that these patients present also the GG genotype for rs3859206, associated with a better performance of the FN3K enzyme [71], add new useful information reinforcing the idea that these two variants could act together and influence the FN3K enzymatic activity, varying the amount of HbA1c.

Moreover, it has been observed that some genotypes identified in subjects with macro-vascular complications were absent in patients with microvascular complications and vice versa. This observation could represent a new step in early diagnosis of diabetic complications. Identification of specific polymorphisms in an early stage of the disease or during the asymptomatic period could be useful especially in familial cases allowing a prompt intervention and promising a better prognosis.

Finally, in cohort 2 a significant association of the C allele of polymorphism in exon 6 (c.900 C>G, rs1056534) with diabetic nephropathy was not observed (p=0.7305), maybe for the small number of patients presenting such complication.

6.3 Mutation pathogenicity and familial study

The FN3K gene analysis identified four likely pathogenic variants (c.2 T>A; c.465 G>A; c.559 C>T and c.716 A>G). These genetic variants were found in the heterozygous state in only one subject in all cases with T2DM. All of them were not detected in any subjects of the control group. Moreover, these mutations were never described before and could represent a new step in the identification of genetic risk factors for T2DM.

An RNA expression study was performed in order to confirm the hypothesis on pathogenicity of these mutations found at genomic level. However, this analysis didn't confirm the result predicted in silico, suggesting that these variants could exert a pathological function with other mechanisms or interacting with other variants present in the FN3K gene.

Variants c.559 C>T (p.R187*) and c.716 A>G (p.Y239C) in the FN3K gene were analyzed also in the familial context. Furthermore, index cases relatives were tested for the presence of the three polymorphisms mostly associated with FN3K activity.

Family 1 (Fig. 51) harboured the missense mutation p.Y239C and consisted of four subjects: the index case and three siblings, one brother and two sisters. All family members presented T2DM, except for the younger sister that has T1DM. The analysis of the FN3K variants has demonstrated the presence of the c.716 A>G mutation in 2 siblings: proband's brother (II:1) and younger sister (II:2). The analysis of the other polymorphisms highlighted that all subjects carrying the mutation presented the same genotype forthe other polymorphisms under investigation (Table 12).

Regarding family 2 (Fig. 52) harbouring truncating mutation p.R187*, only the brother (II:1) of the index case underwent on genetic test. Proband's brother at the

moment of the test didn't present any symptoms of diabetes. The FN3K screening didn't find out the presence of the mutation. Moreover, in this family the two siblings displayed the same genotype for the other polymorphisms under investigation.

These results are hard to evaluate since the role of FN3K variants in the development or progression of diabetes is not yet clear. The lack of an association between some polymorphisms identified in the FN3K gene and DM explains that probably this enzyme cannot by itself account for the entire deglycation story; it alone cannot be responsible for the susceptibility to this pathology or for the development of its complications. However, a better knowledge of FN3K and its aspect could be considered useful since it may be associated with an increased risk on DM onset, reported as a multifactorial disease.

6.4 MicroRNA

Previous literature has demonstrated that endogenous serum and plasma miRNAs are sufficiently stable and quantifiable to serve as clinical biomarkers of diseases including diabetes[142][143][144]. Mitchell et al. [143] found that the abundance of miRNAs in serum and plasma were strongly correlated, which indicates that both serum and plasma samples would be suitable for investigations of miRNAs as blood-based biomarkers.

Mammalian circulating miRNAs are packaged inside lipid or lipid-protein complexes, such as microvesicles, exosomes or apoptotic bodies [95][96], or in associations with proteins [100]. miRNAs secreted inside exosomes may have a different fate compared with those associated with proteins or HDL[98]. Moreover, specific miRNAs tend to be preferentially released by cells and are thus more prone to be found in circulation [145]. Although emerging evidence has established that extracellular miRNAs contained in the body fluids have distinct physiological function [143][144], the main function and the meaning of the existence of circulating miRNAs remain elusive.

The expression of circulating miRNA in plasma samples of patients with T2DM and compensated hypogonadism exposed to soy treatment for twelve weeks has been analyzed. The miRCURY LNA RT-qPCR system were chosen for this analysis because it has been found to be more sensitive and reproducible than microarray systems and showed even more sensitivity and linearity than TaqMan methods at low concentration of plasma miRNA [146].

Three miRNAs, miR-34a-5p, miR-144-3p and miR-19b-3p, were found to be up-regulated and showed a statistically significant difference between pre and post soy

treatment. Among them, there are evidences on miR-144 to be up-regulated by estrogens in cancer cells and cancer-associated fibroblasts [147]. Furthermore, miR-144 was also linked to T2DM. Its expression is up-regulated in T2DM and impairs insulin signaling through down-regulation of insulin receptor substrate 1 (IRS1) [148] in impaired fasting glucose and T2DM patients.

Soy has shown to have beneficial effects on the management of cardiovascular disease and diabetes [149]. In this study, after soy supplementation changes in testosterone levels, in glycemic control and cardiovascular risk markers were observed. Moreover, the number of expressed miRNA seemed to increase after treatment.

A consistent and truly significant effect of soy on the miRNAs plasma profile could not be demonstrated since a validation study is needed to confirm these findings and to proceed with further evaluation. However, these preliminary results are encouraging for future investigation on soy isoflavones potential and for miRNA utilization as non-invasive biomarkers.

7. CONCLUSIONS

It is reported that the human genome of approximately three billion base pairs (www.ornl.gov) contains at least 10 million SNPs. SNPs are defined as genetic variations present in the general population [150]. Human genetic variation occurs as a continuum, ranging from neutral polymorphisms, through functional polymorphisms and disease susceptibility variants to true pathogenic variantswith a high penetrance.

For decades, investigators have worked to unravel the role of genetics in type 2 diabetes through epidemiological studies, studies of candidate genes, and genetic linkage in families. Over the past year, a number of articles have been published based on high-throughput genome-wide association studies (GWAS). These have not only uncovered a number of new genetic loci associated with diabetes and provided new targets for mechanistic investigation, but have also forcing us to reconsider the degree of genetic heterogeneity and possibly even the role of genetics itself in the pathogenesis of diabetes.

It has become increasingly clear that our genomes contain many 'putatively disadvantageous variants' that are probably insufficient on their own to cause disease, but nevertheless still have the potential to contribute to pathogenesis. Moreover, it has also become clear that many genetic disorders are not monogenic as originally supposed, but may instead involve mutations in two or more genes. It could be speculated that different combinations of pathological mutations with low penetrance, functional polymorphisms, disease susceptibility variants and 'putatively disadvantageous variants' may vary quite considerably in terms of their net functional and hence clinical effect[151].

Diabetes management and/or prevention should aim to incorporate genetic and molecular screens to tailor specific treatments (therapeutics as well as lifestyle changes) and optimize benefits for patients.

FN3K appears to be part of a cellular defense and/or repair system to control non-enzymatic glycation of proteins. FN3K would be able to oppose one of the chemical consequences of hyperglycemia, exerting a protective function in the development of diabetic complications, which are reported to differ from individual to individual and to be at least partly genetically determined [152]. Therefore, it could be hypothesized that low activities of this enzyme will increase the glycation level of proteins, and in the case of some specific proteins, participate in the development of complications. This could happen through a direct as well as indirect action. Erythrocyte FN3K enzymatic

activity was found associated with some polymorphisms in the FN3K gene. Therefore the identification of novel SNPs/mutations in this gene is of great interest.

Up to now, this study allowed the identification of 15 genetic variants in the FN3K gene. Seven were novel: 3 missense mutations, 1 synonymous mutation and 3 polymorphisms (1 intronic and 2 in the promoter region). The remaining 8 variants were classified as polymorphisms.

The three SNPs detected by Delpierre et al. founded associated with the erythrocyte FN3K enzymatic activity [71], didn't change FN3K aminoacidic sequence. For this reason, focusing the attention not only on mutation causing an aminoacid change, but also on silent alterations or polymorphisms present in non-coding regions is necessary. Moreover, it has been reported that many causal variants responsible for some common complex disorders are really non-coding[138]: such variants probably don't influence directly the protein function, but produce a quantitative effect on protein expression acting on mRNA stability or on trascriptional capacity of the gene. This study add new useful informations to works concerning the molecular characterization of FN3K gene.

A limit of this work is the lack of the measurement of FN3K enzymatic activity; knowing the correlation between FN3K genotype and enzymatic activity in different diabetic patients, divided according to stage and type of complication, may help to understand the real role of this enzyme and its possible involvement in the pathology. If such involvement really exists, the gene encoding FN3K would be added to the list of all other genes found associated with diabetes through the GWA studies performed till now.

By demonstrating the association between a gene and a disease, the studies uncovered new etiological pathways to be explored, leading to improve prevention and specific treatments for patients. Increasing expression of enzyme against glycation is thought to provide a novel and potentially effective future therapeutic strategy to suppress protein glycation [54].

The lack of association between some polymorphisms identified in FN3K gene and DM suggests that probably the role of FN3K in glycation of biological proteins seems to be more complex than evident at a first glimpse. The variability of FN3K activity may provide another key of explanation in those circumstances in whom HbA1c does not perfectly correlate with the mean glucose level [153]. It has been shown that HbA1c values may not be constant among individuals, despite the presence of similar blood glucose or fructosamine concentration [154]. It would be interesting to expand these studies and to correlate the FN3K activity with the glycation gap and with the

development of diabetic complications. Indeed, probably also other molecular mechanisms impacting gene expression and activity of enzymes in deglycation process may have to be taken into account [132].

Larger studies are necessary for a better understanding of the possible effect of FN3K genetic variants on the progression of the disease and its possible clinical utility in the management of diabetic patients. One of the future aims is to genotype the FN3K gene in other diabetic and control subjects in order to expand the studied cohorts and to make statistical analyses more significant.

Other future perspective is the development of a method for measuring FN3K enzymatic activity in order to find a possible correlation between this parameter and each SNPs/mutation identified in the genetic analysis.

Indeed FN3K represents a piece in the puzzle of diabetes disease.

8. REFERENCES

1. American Diabetes Association. 2. Classification and Diagnosis of Diabetes. Diabetes Care. 2014;38:S8–16.

2. Danaei G, Finucane MM, Lu Y, Singh GM, Cowan MJ, Paciorek CJ, et al. National, regional, and global trends in fasting plasma glucose and diabetes prevalence since 1980: Systematic analysis of health examination surveys and epidemiological studies with 370 country-years and 2·7 million participants. Lancet. Elsevier Ltd; 2011;378:31–40.

3. On the up. Nature. 2012;482:276–276.

4. Shaw JE, Sicree R a., Zimmet PZ. Global estimates of the prevalence of diabetes for 2010 and 2030. Diabetes Res Clin Pract. 2010;87:4–14.

5. Hossain P, Kawar B, El Nahas M. Obesity and diabetes in the developing world a growing challenge. N Engl J Med. 2007;356:213–5.

6. Atkinson MA, Eisenbarth GS MA. Type 1 diabetes. Lancet. 2015;383:69–82.

7. Wild S, Roglic G, Green A, Sicree R KH. Global Prevalence of Diabetes. Estimates for the year 2000 and projections for 2030. Diabetes Care. 2004;27:1047–53.

8. Thomas CC, Philipson LH. Update on Diabetes Classification. Med Clin North Am. Elsevier Inc; 2015;99:1–16.

9. Schwitzgebel VM. Many faces of monogenic diabetes. J Diabetes Investig. 2014;5:121–33.

10. American Diabetes Association. Diagnosis and classification of diabetes mellitus. Diabetes Care. 2014;37:81–90.

11. Groop L, Forsblom C, Lehtovirta M, Tuomi T, Karanko S, Nissén M, et al. Metabolic consequences of a family history of NIDDM (the Botnia study): Evidence for sex-specific parental effects. Diabetes. 1996;45:1585–93.

12. Weijnen CF, Rich SS, Meigs JB, Krolewski a. S, Warram JH. Risk of diabetes in siblings of index cases with Type 2 diabetes: Implications for genetic studies. Diabet Med. 2002;19:41–50.

13. Flegal KM, Ezzati TM, Harris MI, Haynes SG, Juarez RZ, Knowler WC, Perez-Stable EJ SM. Prevalence of diabetes in Mexican Americans, Cubans, and Puerto Ricans from the Hispanic Health and Nutrition Examination Survey, 1982-1984. Diabetes Care. 1991;14:628–38.

14. Grant SF a, Thorleifsson G, Reynisdottir I, Benediktsson R, Manolescu A, Sainz J, et al. Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. Nat Genet. 2006;38:320–3.

15. Brunetti A, Chiefari E, Foti D. Recent advances in the molecular genetics of type 2 diabetes mellitus. World J Diabetes. 2014;5:128–40.

16. McCarthy MI, McCarthy MI, Zeggini E, Zeggini E. Genome-wide association studies in type 2 diabetes. Curr Diab Rep. 2009;9:164–71.

17. Bonnefond A, Froguel P. Rare and Common Genetic Events in Type 2 Diabetes: What Should Biologists Know? Cell Metab. 2015;21:357–68.

18. Morris A, Voight B, Teslovich T. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. Nat Genet. 2012;44:981–90.

19. Doria A, Patti ME KC. The Emerging Genetic Architecture of Type 2 Diabetes. Cell Metab. 2008;8:186–200.

20. Ahlqvist E, Ahluwalia TS, Groop L. Genetics of type 2 diabetes. Clin Chem. 2011;57:241–54.

21. Murray P, Chune GW, Raghavan V a. Legacy effects from DCCT and UKPDS: What they mean and implications for future diabetes trials. Curr Atheroscler Rep. 2010;12:432–9.

22. Dorajoo R, Liu J, Boehm B. Genetics of Type 2 Diabetes and Clinical Utility. Genes (Basel). 2015;6:372–84.

23. American Diabetes Association. Diagnosis and classification of diabetes mellitus. Diabetes Care. 2010;33.

24. World health organization. Definition and Diagnosis of Diabetes Mellitus and Intermediate Hyperglycemia. Who. 2006;50.

25. Nathan DM, Balkau B, Bonora E, Borch-Johnsen K, Buse JB, Colagiuri S, et al. International expert committee report on the role of the A1C assay in the diagnosis of diabetes. Diabetes Care. 2009;32:1327–34.

26. Association AD. Standards of medical care in diabetes - 2013. Diabetes Care. 2013;36.

27. Stratton IM, Adler AI, Neil HA, Matthews DR, Manley SE, Cull CA, et al. Association of glycaemia with macrovascular and microvascular complications of type 2 diabetes (UKPDS 35): prospective observational study. BMJ. 2000;321:405–12.

28. Forbes JM, Cooper ME. Mechanisms of diabetic complications. Physiol Rev. 2013;93:137–88.

29. Group T diabetes control and complication trial research. the effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. N Engl J Med. 1993;329:977–86.

30. Group UKPDS. Intensive Blood-Glucose Control With Sulphonylureas or Insulin Compared With Conventional Treatment and Risk of Complications in Patients With Type 2 Diabetes. Lancet. 1998;352:837–53.

31. Farouque HM, O'Brien RC MI. Diabetes mellitus and coronary heart disease from prevention to intervention: Part I. Aust N Z J Med. 2000;30:351–9.

32. Brownlee M. Biochemicstry and molecular cell biology of diabetic complications. Nature. 2001;414:813–20.

33. Sheetz MJ, King GL. Molecular understanding of hyperglycemia's adverse effects for diabetic complications. JAMA. 2002;288:2579–88.

34. Stitt AW, Jenkins AJ CM. Advanced glycation end products and diabetic complications. Expert Opin Investig Drugs. 2002;11:1205–23.

35. Morgan PE, Dean RT, Davies MJ. Inactivation of cellular enzymes by carbonyls and protein-bound glycation/glycoxidation products. Arch Biochem Biophys. 2002;403:259–69.

36. AJ F. Glycated proteins in diabetes. Br J Biomed Sci. 1997;54:192–200.

37. JM A. The Maillard reaction. Elsevier Appl Sci. 1992;99–153.

38. Ulrich P CA. Protein glycation, diabetes, and aging. Recent Prog Horm Res. 2001;56:1–21.

39. Chilelli NC, Burlina S, Lapolla a. AGEs, rather than hyperglycemia, are responsible for microvascular complications in diabetes: A "glycoxidation-centric" point of view. Nutr Metab Cardiovasc Dis. Elsevier Ltd; 2013;23:913–9.

40. Meerwaldt R, Links T, Zeebregts C, Tio R, Hillebrands J-L, Smit A. The clinical relevance of assessing advanced glycation endproducts accumulation in diabetes. Cardiovasc Diabetol. 2008;7:29.

41. Thornalley PJ, Battah S, Ahmed N, Karachalias N, Agalou S, Babaei-Jadidi R, et al. Quantitative screening of advanced glycation endproducts in cellular and extracellular proteins by tandem mass spectrometry. Biochem J. 2003;375:581–92.

42. Monnier VM, Sell DR, Genuth S. Glycation Products as Markers and Predictors of the Progression of Diabetic Complications. Ann N Y Acad Sci. 2005;1043:567–81.

43. Singh R, Barden a., Mori T, Beilin L. Advanced glycation end-products: A review. Diabetologia. 2001;44:129–46.

44. Monnier VM, Mustata GT, Biemel KL, Reihl O, Lederer MO, Zhenyu D, et al. Cross-Linking of the Extracellular Matrix by the Maillard Reaction in Aging and Diabetes: An Update on "a Puzzle Nearing Resolution." Ann N Y Acad Sci. 2005;1043:533–44.

45. Verzijl N, DeGroot J, Thorpe SR, Bank R a., Shaw JN, Lyons TJ, et al. Effect of Collagen Turnover on the Accumulation of Advanced Glycation End Products. J Biol Chem. 2000;275:39027–31.

46. Schmidt AM. Isolation and characterization of two binding proteins for advanced glycosylation end products from bovine lung which are present on the endothelial cell surface. J Biol Chem. 1992;267:14987–97.

47. Guimarães ELM, Empsen C, Geerts A, van Grunsven LA. Advanced glycation end products induce production of reactive oxygen species via the activation of NADPH oxidase in murine hepatic stellate cells. J Hepatol. 2015;52:389–97.

48. Zhang M. Glycated Proteins Stimulate Reactive Oxygen Species Production in Cardiac Myocytes: Involvement of Nox2 (gp91phox)-Containing NADPH Oxidase. Circulation. 2006;113:1235–43.

49. Bierhaus a, Schiekofer S, Schwaninger M, Andrassy M, Humpert PM, Chen J, et al. Diabetes-associated sustained activation of the transcription factor nuclear factor-kappaB. Diabetes. 2001;50:2792–808.

50. Libermann TA, Baltimore D. Activation of interleukin-6 gene expression through the NF-kappa B transcription factor. Mol Cell Biol. 1990;10:2327–34.

51. Ueda, A.; Ishigatsubo, Y.; Okubo, T.; Yoshimura T. Transcriptional Regulation of the Human monocyte chemoattractant protein-1 gene. Biochemistry. 1997;272:31092–9.

52. Li J. Characterization and Functional Analysis of the Promoter of RAGE, the Receptor for Advanced Glycation End Products. Mol Biol. 1997;272:16498–506.

53. Schmidt AM, Stern DM. RAGE: a new target for the prevention and treatment of the vascular and inflammatory complications of diabetes. Trends Endocrinol Metab. 2000;11:368–75.

54. Ahmed N, Thornalley PJ. Advanced glycation endproducts: What is their relevance to diabetic complications? Diabetes, Obes Metab. 2007;9:233–45.

55. Wiame E, Delpierre G, Collard F, Van Schaftingen E. Identification of a pathway for the utilization of the amadori product fructoselysine in Escherichia coli. J Biol Chem. 2002;277:42523–9.

56. Monnier VM, Wu X. Enzymatic deglycation with amadoriase enzymes from Aspergillus sp. as a potential strategy against the complications of diabetes and aging. Biochem Soc Trans. 2003;31:1349–53.

57. Delpierre G, Vanstapel F, Stroobant V, Van Schaftingen E. Conversion of a synthetic fructosamine into its 3-phospho derivative in human erythrocytes. Biochem J. 2000;352 Pt 3:835–9.

58. Petersen a., Szwergold BS, Kappler F, Weingarten M, Brown TR. Identification of sorbitol 3-phosphate and fructose 3-phosphate in normal and diabetic human erythrocytes. J Biol Chem. 1990;265:17424–7.

59. Szwergold BS, Howell S, Beisswenger PJ. Human Fructosamine-3-Kinase: Purification, sequencing, substrate specificity, and evidence of activity in vivo. Diabetes. 2001;50:2139–47.

60. Delpierre G, Collard F, Fortpied J, Van Schaftingen E. Fructosamine 3-kinase is involved in an intracellular deglycation pathway in human erythrocytes. Biochem J. 2002;365:801–8.

61. Collard F, Delpierre G, Stroobant V, Matthijs G, Van Schaftingen E. A Mammalian Protein Homologous to Fructosamine-3-Kinase Is a Ketosamine-3-Kinase Acting on Psicosamines and Ribulosamines but not on Fructosamines. Diabetes. 2003;52:2888–95.

62. Delplanque J, Delpierre G, Opperdoes FR, Van Schaftingen E. Tissue distribution and evolution of fructosamine 3-kinase and fructosamine 3-kinase-related protein. J Biol Chem. 2004;279:46606–13.

63. Conner JR, Beisswenger PJ, Szwergold BS. Some clues as to the regulation, expression, function, and distribution of fructosamine-3-kinase and fructosamine-3-kinase-related protein. Ann N Y Acad Sci. 2005;1043:824–36.

64. Veiga da-Cunha M, Jacquemin P, Delpierre G, Godfraind C, Théate I, Vertommen D, et al. Increased protein glycation in fructosamine 3-kinase-deficient mice. Biochem J. 2006;399:257–64.

65. Pascal SM a, Veiga-da-Cunha M, Gilon P, Van Schaftingen E, Jonas JC. Effects of fructosamine-3-kinase deficiency on function and survival of mouse pancreatic islets after prolonged culture in high glucose or ribose concentrations. Am J Physiol Endocrinol Metab. 2010;298:E586–96.

66. Delpierre G, Vertommen D, Communi D, Rider MH, Van Schaftingen E. Identification of fructosamine residues deglycated by fructosamine-3-kinase in human hemoglobin. J Biol Chem. 2004;279:27613–20.

67. Van Schaftingen E, Collard F, Wiame E, Veiga-Da-Cunha M. Enzymatic repair of Amadori products. Amino Acids. 2012;42:1143–50.

68. Zoungas S, de Galan BE, Ninomiya T, Grobbee D, Hamet P, Heller S, et al. Combined effects of routine blood pressure lowering and intensive glucose control on macrovascular and microvascular outcomes in patients with type 2 diabetes: New results from the ADVANCE trial. Diabetes Care. 2009;32:2068–74.

69. Action to Control Cardiovascular Risk in Diabetes Study Group, Gerstein HC, Miller ME, Byington RP, Goff DC Jr, Bigger JT, Buse JB, Cushman WC, Genuth S, Ismail-Beigi F, Grimm RH Jr, Probstfield JL, Simons-Morton DG, Friedewald WT.Action to Control Cardio FW. Effects of intensive glucose lowering in type 2 diabetes. N Engl J Med. 2008;358:2545–59.

70. Zoungas S, Patel A, Chalmers J, de Galan BE, Li Q, Billot L, et al. Severe hypoglycemia and risks of vascular events and death. N Engl J Med. 2010;363:1410–8.

71. Delpierre G, Veiga-da-Cunha M, Vertommen D, Buysschaert M, Van Schaftingen E. Variability in erythrocyte fructosamine 3-kinase activity in humans correlates with polymorphisms in the FN3K gene and impacts on haemoglobin glycation at specific sites. Diabetes Metab. 2006;32:31–9.

72. Krause R, Oehme A, Wolf K, Henle T. A convenient HPLC assay for the determination of fructosamine-3-kinase activity in erythrocytes. Anal Bioanal Chem. 2006;386:2019–25.

73. Mohás M, Kisfali P, Baricza E, Mérei A, Maász A, Cseh J, Mikolás E, Szijártó IA, Melegh B WI. A polymorphism within the fructosamine-3-kinase gene is associated with HbA1c Levels and the onset of type 2 diabetes mellitus. Exp Clin Endocrinol Diabetes. 2010;118:209–12.

74. Tanhäuserová V, Kuricóva K, Pácal L, Bartáková V, Řehořová J, Svojanovský J, et al. Genetic variability in enzymes of metabolic pathways conferring protection against non-enzymatic glycation versus diabetes-related morbidity and mortality. Clin Chem Lab Med. 2014;52:77–83.

75. Škrha J, Muravská A, Flekač M, Horová E, Novák J, Novotný A, et al. Fructosamine 3-kinase and glyoxalase I polymorphisms and their association with soluble RAGE and adhesion molecules in diabetes. Physiol Res. 2014;63:S283–91.

76. Mosca L, Penco S, Patrosso MC, Marocchi A, Lapolla A, Sartore G, et al. Genetic variability of the fructosamine 3-kinase gene in diabetic patients. Clin Chem Lab Med. 2011;49:803–8.

77. Alexander RP, Fang G, Rozowsky J, Snyder M, Gerstein MB. Annotating non-coding regions of the genome. Nat Rev Genet. Nature Publishing Group; 2010;11:559–71.

78. Mercer TR, Dinger ME, Mattick JS. Long non-coding RNAs: insights into functions. Nat Rev Genet. 2009;10:155–9.

79. He L, Hannon GJ. MicroRNAs: small RNAs with a big role in gene regulation. Nat Rev Genet. 2004;5:522–31.

80. Mendell JT. MicroRNAs: Critical regulators of development, cellular physiology and malignancy. Cell Cycle. 2005;4:1179–84.

81. Lee RC. The C . elegans Heterochronic Gene lin-4 Encodes Small RNAs with Antisense Complementarity to & II-14. Cell. 1993;75:843–54.

82. Kozomara A, Griffiths-Jones S. MiRBase: Integrating microRNA annotation and deep-sequencing data. Nucleic Acids Res. 2011;39:152–7.

83. Friedman RC, Farh KKH, Burge CB, Bartel DP. Most mammalian mRNAs are conserved targets of microRNAs. Genome Res. 2009;19:92–105.

84. Lee Y, Kim M, Han J, Yeom K-H, Lee S, Baek SH, et al. MicroRNA genes are transcribed by RNA polymerase II. EMBO J. 2004;23:4051–60.

85. Borchert GM, Lanier W, Davidson BL. RNA polymerase III transcribes human microRNAs. Nat Struct Mol Biol. 2006;13:1097–101.

86. Lee Y, Ahn C, Han J, Choi H, Kim J, Yim J, et al. The nuclear RNase III Drosha initiates microRNA processing. Nature. 2003;425:415–9.

87. Lund E, Güttinger S, Calado A, Dahlberg JE, Kutay U. Nuclear export of microRNA precursors. Science. 2004;303:95–8.

88. Pillai RS, Bhattacharyya SN, Filipowicz W. Repression of protein synthesis by miRNAs: how many mechanisms? Trends Cell Biol. 2007;17:118–26.

89. Krol J, Loedige I, Filipowicz W. The widespread regulation of microRNA biogenesis, function and decay. Nat Rev Genet. Nature Publishing Group; 2010;11:597–610.

90. Etheridge A, Lee I, Hood L, Galas D, Wang K. Extracellular microRNA: A new source of biomarkers. Mutat Res - Fundam Mol Mech Mutagen. Elsevier B.V.; 2011;717:85–90.

91. Hwang H-W, Mendell JT. MicroRNAs in cell proliferation, cell death, and tumorigenesis. Br J Cancer. 2006;94:776–80.

92. Esteller M. Non-coding RNAs in human disease. Nat Rev Genet. Nature Publishing Group; 2011;12:861–74.

93. Weber J a., Baxter DH, Zhang S, Huang DY, Huang KH, Lee MJ, et al. The microRNA spectrum in 12 body fluids. Clin Chem. 2010;56:1733–41.

94. Wang K, Zhang S, Weber J, Baxter D, Galas DJ. Export of microRNAs and microRNA-protective protein by mammalian cells. Nucleic Acids Res. 2010;38:7248–59.

95. Valadi H, Ekström K, Bossios A, Sjöstrand M, Lee JJ, Lötvall JO. Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. Nat Cell Biol. 2007;9:654–9.

96. Zernecke A, Bidzhekov K, Noels H, Shagdarsuren E, Gan L, Denecke B, Hristov M, Köppel T, Jahantigh MN, Lutgens E, Wang S, Olson EN, Schober A WC. of microRNA-126 by apoptotic bodies induces CXCL12-dependent vascular protection. Sci Signal. 2009;2:ra81.

97. Arroyo JD, Chevillet JR, Kroh EM, Ruf IK, Pritchard CC, Gibson DF, et al. Argonaute2 complexes carry a population of circulating microRNAs independent of vesicles in human plasma. Proc Natl Acad Sci U S A. 2011;108:5003–8.

98. Vickers KC, Palmisano BT, Shoucri BM, Shamburek RD, Remaley AT. NIH Public Access. 2011;13:423–33.

99. Iguchi H, Kosaka N, Ochiya T. Secretory microRNAs as a versatile communication tool. Commun Integr Biol. 2010;3:478–81.

100. Turchinovich A, Weiz L, Langheinz A, Burwinkel B. Characterization of extracellular circulating microRNA. Nucleic Acids Res. 2011;39:7223–33.

101. Creemers EE, Tijsen AJ, Pinto YM. Circulating MicroRNAs: Novel biomarkers and extracellular communicators in cardiovascular disease? Circ Res. 2012;110:483–95.

102. Lichtenstein a H, Kennedy E, Barrier P, Danford D, Ernst ND, Grundy SM, et al. Dietary fat consumption and health. Nutr Rev. 1998;56:S3–19.

103. Teede HJ, Dalais FS, Kotsopoulos D, Liang Y, Davis S, Mcgrath BP. Placebo-Controlled Study in Men and Postmenopausal Women Dietary Soy Has Both Beneficial and Potentially Adverse Cardiovascular Effects : A Placebo-Controlled Study in Men and Postmenopausal Women *. J Clin Endocrinol Metab. 2001;86:3053–60.

104. Gonzalez S, Jayagopal V, Kilpatrick ES, Chapman T, Atkin SL. Effects of Isoflavone Dietary Supplementation on Cardiovascular Risk Factors in Type 2 Diabetes. Diabetes Care. 2007;30:1871–3.

105. Wu a H, Yu MC, Tseng C-C, Pike MC. Epidemiology of soy exposures and breast cancer risk. Br J Cancer. 2008;98:9–14.

106. Yan L, Spitznagel EL. Soy consumption and prostate cancer risk in men : a revisit of meta analysis. Am J Clin Nutr. 2009;89:1155–63.

107. Zhang X, Shu X-O, Li H, Yang G, Li Q, Gao YT, et al. Prospective cohort study of soy food consumption and risk of bone fracture among postmenopausal women. Arch Intern Med. 2005;165:1890–5.

108. Howes LG, Howes JB, Knight DC. Isoflavone therapy for menopausal flushes: A systematic review and meta-analysis. Maturitas. 2006;55:203–11.

109. Hervouet E, Cartron P-F, Jouvenot M, Delage-Mourroux R. Epigenetic regulation of estrogen signaling in breast cancer. Epigenetics. 2013;8:237–45.

110. Sarkar FH LY. Soy isoflavones and cancer prevention. Cancer Invest. 2003;21:744–57.

111. Evelyne Reiterab, Evelyne Reiterab, Verena Becka SM& AJ. Isoflavones are safe compounds for therapeutical applications—evaluation of in vitro data. Gynecol Endocrinol. 2009;25:554–80.

112. Pan L, Xia X, Feng Y, Jiang C, Huang Y. Exposure to the Phytoestrogen Daidzein Attenuates Apomorphine-Induced Penile Erection Concomitant with Plasma Testosterone Level Reduction in Dose and Time-Related Manner in Adult Rats. Urology. 2007;70:613–7.

113. Faqi AS, Johnson WD, Morrissey RL, McCormick DL. Reproductive toxicity assessment of chronic dietary exposure to soy isoflavones in male rats. Reprod Toxicol. 2004;18:605–11.

114. Chavarro JE, Sadio SM, Toth TL, Hauser R. Soy food and soy isoflavone intake in relation to semen quality parameters. Fertil Steril. 2007;88:S22.

115. Martinez J LJ. An unusual case of gynecomastia associated with soy product consumption. Endocr Pr. 2008;14:415–8.

116. Mitchell JH, Cawood E, Kinniburgh D, Provan a, Collins a R, Irvine DS. Effect of a phytoestrogen food supplement on reproductive health in normal males. Clin Sci (Lond). 2001;100:613–8.

117. Messina M, Messina V. The role of soy in vegetarian diets. Nutrients. 2010;2:855–88.

118. Dhindsa S, Prabhakar S, Sethi M, Bandyopadhyay A, Chaudhuri A, Dandona P. Frequent occurrence of hypogonadotropic hypogonadism in type 2 diabetes. J Clin Endocrinol Metab. 2004;89:5462–8.

119. Chrysohoou C, Panagiotakos D, Pitsavos C, Siasos G, Oikonomou E, Varlas J, et al. Low total testosterone levels are associated with the metabolic syndrome in elderly men: The role of body weight, lip-ids, insulin resistance, and inflammation; the Ikaria study. Rev Diabet Stud. 2013;10:27–38.

120. Kapoor D, Malkin CJ, Channer KS, Jones TH. Androgens, insulin resistance and vascular disease in men. Clin Endocrinol. 2005;63:239–50.

121. Nathan DM, Kuenen J, Borg R, Zheng H, Schoenfeld D, Heine RJ. Translating the A1C assay into estimated average glucose values. Diabetes Care. 2008;31:1473–8.

122. World Medical Association (WMA) declaration of Helsinki. Recommndations guiding physicians in biomedical research involving human subjects. J Am Med Assoc. 1997;277:925–6.

123. Mueller RF YI. Emery's Elements of medical genetics. 9th ed. Livingstone C, editor. Edinburgh; 1995. 317.

124. Kwok P-Y, Chen X. Detection of single nucleotide polymorphisms. Curr Issues Mol Biol. 2003;5:43–60.

125. Frueh FW, Noyer-Weidner M. The use of denaturing high-performance liquid chromatography (DHPLC) for the analysis of genetic variations: impact for diagnostics and pharmacogenetics. Clin Chem Lab Med. 2003;41:452–61.

126. Kosaki K, Udaka T, Okuyama T. DHPLC in clinical molecular diagnostic services. Mol Genet Metab. 2005;86:117–23.

127. Schwarz JM, Rödelsperger C, Schuelke M, Seelow D. MutationTaster evaluates disease-causing potential of sequence alterations. Nat Methods. Nature Publishing Group; 2010;7:575–6.

128. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. Nat Methods. Nature Publishing Group; 2010;7:248–9.

129. Ng PC, Henikoff S. Predicting the Effects of Amino Acid Substitutions on Protein Function. Annu Rev Genomics Hum Genet. 2006;7:61–80.

130. Groop L, Pociot F. Genetics of diabetes - Are we missing the genes or the disease? Mol Cell Endocrinol. Elsevier Ireland Ltd; 2014;382:726–39.

131. Bry L, Chen PC, Sacks DB. Effects of hemoglobin variants and chemically modified derivatives on assays for glycohemoglobin. Clin Chem. 2001;47:153–63.

132. Rabbani N, Thornalley PJ. Glyoxalase in diabetes, obesity and related disorders. Semin Cell Dev Biol. Elsevier Ltd; 2011;22:309–17.

133. Van Schaftingen E, Delpierre G, Collard F, Fortpied J, Gemayel R, Wiame E. Fructosamine 3-kinase and other enzymes involved in protein deglycation. Adv Enzym Regul. 2007;47:261–9.

134. Zhang Q, Ames JM, Smith RD, Baynes JW, Metz TO. A perspective on the Maillard reaction and the analysis of protein glycation by mass spectrometry: probing the pathogenesis of chronic disease. J Proteome Res. 2008;8:754–69.

135. S Roriz-Filho J, Sá-Roriz TM, Rosset I, Camozzato AL, Santos AC, Chaves ML, Moriguti JC R-CM. (Pre)diabetes, brain aging, and cognition. Biochim Biophys Acta. 2009;1792:432–43.

136. Conner JR, Beisswenger PJ, Szwergold BS. The expression of the genes for fructosamine-3-kinase and fructosamine-3-kinase-related protein appears to be constitutive and unaffected by environmental signals. Biochem Biophys Res Commun. 2004;323:932–6.

137. Kim H, Clark D, Dionne RA. Genetic contributions to clinical pain and analgesia: avoiding pitfalls in genetic research. J Pain. 2009;10:663–93.

138. Cordell HJ, Clayton DG. Genetic association studies. Lancet. 2005;366:1121–31.

139. Sacks DB. Global harmonization of hemoglobin A1c. Clin Chem. 2005;51:681–3.

140. Goldstein DE, Little RR, Lorenz R a., Malone JI, Nathan D, Peterson CM, et al. Tests of glycemia in diabetes. Diabetes Care. 2004;27:1761–73.

141. Bunn HF, Haney DN, Kamin S, Gabbay KH, Gallop PM. The biosynthesis of human hemoglobin A1c. Slow glycosylation of hemoglobin in vivo. J Clin Invest. 1976;57:1652–9.

142. Zampetaki A, Kiechl S, Drozdov I, Willeit P, Mayr U, Prokopi M, et al. Plasma MicroRNA profiling reveals loss of endothelial MiR-126 and other MicroRNAs in type 2 diabetes. Circ Res. 2010;107:810–7.

143. Mitchell PS, Parkin RK, Kroh EM, Fritz BR, Wyman SK, Pogosova-Agadjanyan EL, et al. Circulating microRNAs as stable blood-based markers for cancer detection. Proc Natl Acad Sci U S A. 2008;105:10513–8.

144. Chen X, Ba Y, Ma L, Cai X, Yin Y, Wang K, et al. Characterization of microRNAs in serum: a novel class of biomarkers for diagnosis of cancer and other diseases. Cell Res. 2008;18:997–1006.

145. Guduric-Fuchs J, O'Connor A, Camp B, O'Neill CL, Medina RJ, Simpson DA. Selective extracellular vesicle-mediated export of an overlapping set of microRNAs from multiple cell types. BMC Genomics. BMC Genomics; 2012;13:357.

146. Jensen SG, Lamy P, Rasmussen MH, Ostenfeld MS, Dyrskjøt L, et al. Evaluation of two commercial global miRNA expression profiling platforms for detection of less abundant miRNAs. BMC Genomics. BioMed Central Ltd; 2011;12:435.

147. Vivacqua A, Marco P De, Santolla MF, Cirillo F, Pellegrino M, Panno ML, et al. Estrogenic gper signaling regulates mir144 expression in cancer cells and cancer-associated fibroblasts. Oncotarget; 2015;6:16573-87.

148. Karolina DS, Armugam A, Tavintharan S, Wong MTK, Lim SC, Sum CF, et al. MicroRNA 144 impairs insulin signaling by inhibiting the expression of insulin receptor substrate 1 in type 2 diabetes mellitus. PLoS One. 2011;6:e22839

149. Hermansen K, Søndergaard M, Høie L, Carstensen M BB. Beneficial Effects of a Soy-Based Dietary Supplement on Lipid Levels and Cardiovascular Risk Markers in Type 2. Diabetes Care. 2001;24:228–33.

150. Sachidanandam R, Weissman D, Schmidt SC, Kakol JM, Stein LD, Marth G, et al. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. Nature. 2001;409:928–33.

151. Cooper DN, Krawczak M, Polychronakos C, Tyler-Smith C, Kehrer-Sawatzki H. Where genotype is not predictive of phenotype: Towards an understanding of the molecular basis of reduced penetrance in human inherited disease. Human Genetics. 2013;132:1077-130.

152. DW B. Genetics of diabetes complications. Curr Diab Rep. 2002;2:191–200.

153. Leslie RDG, Cohen RM. Biologic variability in plasma glucose, hemoglobin A1c, and advanced glycation end products associated with diabetes complications. J Diabetes Sci Technol. 2009;3:635–43.

154. Cohen RM, Smith EP. Frequency of HbA1c discordance in estimating blood glucose control. Curr Opin Clin Nutr Metab Care. 2008;11(4):512–7.

9. ACKNOWLEDGMENTS

I would like to express my gratitude to all the people contributed in some way to the work described in this thesis.

First I would like to thank my advisor Prof. Andrea Mosca of University of Milan, whose expertise, guidance and patience added considerably to my PhD experience. Thanks also for having provided me the opportunity to enjoy the experience of working abroad through the Erasmus Placement.

I would like to thank Prof. Maurizio Ferrari and Dr Paola Carrera of San Raffale Hospital for insightful comments and encouragements incented me to widen my research from various perspective. Thanks to my fellow labmates and technicians: Mascia, Silvia P, Silvia C, Elena, Marcella, Antonella, Giuseppe, Alessandra, Ozge, Liz, Simona, Veronica, Chiara. In particular, Viviana, Cristina and Armando for all laughings during the last two years.

A very special thanks goes out to Dr Silvana Penco of Niguarda Ca' Granda Hospital, it was under her tutelage that I became interested in molecular biology and genetics. Thanks to my fellow labmates and technicians: Rino, Nico, Anto, Linda, Laura, Lora, Lucia, Ste, Claudia and Mari for all the fun we have had in the last years.

I must also acknowledge Prof. Annunziata Lapolla, Dr. Giovanni Sartore and Dr. Nino Cristiano Chilelli of University of Padova for providing samples of patients enrolled in the study and for the detailed clinical informations.

My sincere thanks goes to Dr Thozhukat Sathyapalan and Prof. Eric Kilpatrick of University of Hull (UK), who provided me an opportunity to join their team and who gave me access to the laboratory and research facilities. Thank you for offering me the opportunity to work on diverse exiting project. Thanks also to Vicky, Huw and Andrew for all the help not just scientific and for the patience in answering to all my questions.

I would also thank my friends, I could not imagine my life without you all!!

Thanks to the "Branko" for all trips, dinners but overall for the peculiar "viciousness". Thanks to Franci, Franceskina and Sharon for the remote support. Thanks to Bea e Sere for our sushiate, aperitifs and schemer discussion. Thanks to Ilaria always with me during this 29 years. A special thanks to Elisa who with an insane patience and dedication has created the fabulous lay out of my thesis. Thanks to Viviana and Claudia for popular pearls of wisdom and all the unreasonable chitchats. Thanks also to little and ridiculous Arcibaldo.

Last but not the least, I would thank my family, aunt Daniela, uncle Gino and Massimo for the support they providing me through my entire life and in particular, I must acknowledge my mother without whose love and encouragement I would not have finished this PhD.