AMERICAN UNIVERSITY OF BEIRUT

EXPLORING THE ASSOCIATION BETWEEN GLUCOSE 6 PHOSPHATE DEHYDROGENASE DEFICIENCY AND ITS PROTECTIVE ROLE IN CANCER

by
FARAH OMAR AL JURDI

A thesis
submitted in partial fulfillment of the requirements
for the degree of Master of Science
to the Department of Biochemistry and Molecular Genetics
of the Faculty of Medicine
at the American University of Beirut

Beirut, Lebanon
September, 2020

AMERICAN UNIVERSITY OF BEIRUT

EXPLORING THE ASSOCIATION BETWEEN GLUCOSE 6 PHOSPHATE DEHYDROGENASE DEFICIENCY AND ITS PROTECTIVE ROLE IN CANCER

by

FARAH OMAR AL JURDI

Approved by:

[Signature]

Dr.Pierre Khoueiry, Professor
Biochemistry and Molecular Genetics — Advisor

[Signature]

Dr. Nadine Darwiche, Professor
Biochemistry and Molecular Genetics — Member of Committee

[Signature]

Dr. Julnar Usta, Professor
Biochemistry and Molecular Genetics — Member of Committee

Date of thesis defense: September 02, 2020

AMERICAN UNIVERSITY OF BEIRUT

THESIS RELEASE FORM

Student Name: _____Al Jurdi_____Farah_____Omar_____
Last First Middle

I authorize the American University of Beirut, to: (a) reproduce hard or electronic copies of my thesis; (b) include such copies in the archives and digital repositories of the University; and (c) make freely available such copies to third parties for research or educational purposes:

☒ As of the date of submission

☐ One year from the date of submission of my thesis.

☐ Two years from the date of submission of my thesis.

☐ Three years from the date of submission of my thesis.

_____September 15, 2020_____

Signature Date

. .

ACKNOWLEDGEMENTS

My recognition and gratitude are addressed to my advisor Dr. Pierre Khoueiry for his knowledge and support all the way, and special thanks for thesis committee Dr. Julnar Osta and Dr. Nadine Darwich. In addition, I would like to acknowledge my colleagues Mr. abed Kurdi and Mrs. Heba fleihan for their excessive and countless help and assistance in the computer work.

ABSTRACT
OF THE THESIS OF

Student's Full Name Farah al Jurdi for Master of Science
Major: Biochemistry

Title: Exploring the association between glucose 6 phosphate dehydrogenase deficiency and its protective role in cancer

Background*:* Theglucose 6 ohosphate dehydrogenase (G6PD) enzyme is the first and rate-limiting enzyme in the oxidative arm of the Pentose Phosphate Pathway (PPP). People with G6PD deficiency are susceptible to hemolytic anemia, an X-linked recessive disorder. Moreover, the PPP is an important pathway in cancer as it supplies cancer cells with their high need for nucleic acid synthesis and the produced NADPH is important for the synthesis of fatty acids and to deal with oxidative stress. Therefore, in cancer cells were NADPH consumption is high, G6PD activity is also high. As such, G6PD deficient cells are not able to produce nucleic acids and to use NADPH for fatty acids synthesis leading to ROS accumulation leaving the cells vulnerable to oxidative and energetic stress. Here we intended to explore anemia related G6PD mutations and their possible protective role in cancer.
Materials and Methods: Anemia related G6PD mutations will be extracted and manually curated from published papers. Databases to use for curation will include the UCSC Genome Browser and ClinVar. Cancer related G6PD mutations and expressions will be extracted from The Cancer Genome Atlas (TCGA). Finally, the R programming language will be used for data analysis.
Results: We extracted and curated 151 G6PD mutations related to favism and anemia in the literature. In cancer, 150 G6PD mutations were identified in different cancer types with very low-frequency rates. In total, 20 mutations are common between both sets.
Conclusion: Our analysis will extend our knowledge of the protective role of G6PD deficiency against cancer.

TABLE OF CONTENTS

ILLUSTRATIONS

Figure

TABLES

ABBREVIATIONS

G6PD	Glucose-6-Phosphate dehydrogenase
PPP	Pentose Phosphate Pathway
HMP	Hexose Monophosphate Pathway
NADPH	Nicotinamide adenine dinucleotide phosphate
TCA	Tri-Carboxylic Acid cycle
CAC	Citric Acid Cycle
RBC	Red Blood cells
HCC	Hepato Cellular Carcinoma
GALT	Galactose-1-phosphate uridylyltransferase
CNSHA	Chronic non Spherocytic Hemolytic Anemia
WHO	World Health Organization
GLUT	Glucose Transporter
GSH	Reduced Glutathione
ROS	Reactive Oxygen Species
TKT	Trans Ketolase
TALDO	Trans Aldolase
Ru5P	Ribulose-5-Phosphate
R5P	Ribose-5-Phosphate
G3P	Glyceraldehyde-3-Phosphate
F6P	Fructose-6-Phosphate
STAT3	Signal Transduction and Activator Of Transcription 3

EMT	Epithelial to Mesenchymal Transition
EGF	Epidermal Growth Factor
PDGF	Platelet-Derived Growth Factor
PI3K	Phosphatidylinositol-3-Kinase
AMPK	5’ AMP-Activated Protein Kinase
cAMP	Cyclic Adenosine Mono Phosphate
TIGAR	TP53-Induced Glycolysis and Apoptosis Regulator
HK	Hexokinase
PTEN	The Phosphatase and Tensin Homologue
mTORC	Mammalian Target of Rapamycin Complex 1
Nrf2	Nuclear Factor Erythroid 2 Related Factor 2
HGVS	Human Genome Variation Society
TCGA	The Cancer Genome Atlas
VEP	Variant Effect Predictor
SIFT	Sorting Intolerant from Tolerant
SNP	Single Nucleotide Variant
CNV	Copy Number Variant

CHAPTER I

INTRODUCTION

Glucose 6 phosphate dehydrogenase (G6PD) is a housekeeping enzyme expressed ubiquitously in all cells, where it has a main role in glucose metabolism and is critical to the functioning and integrity of red blood cells (Pandolfi et al., 1995). G6PD is the first step and the rate-limiting enzyme in the Pentose Phosphate Pathway (PPP), also known as the Hexose Monophosphate Shunt (HMP), which was first discovered by Warburg and Christian in 1931. This pathway is important for the production of ribose for nucleotide synthesis, as well as the production of nicotinamide adenine dinucleotide phosphate (NADPH) molecules that are needed for reductive biosynthetic reactions such as bile acid synthesis, cholesterol biosynthesis, fatty acid synthesis, and steroid hormone biosynthesis (Patra and Hay, 2014). It is a very important pathway in cancer. A wide range of organisms has this gene including protozoa, bacteria, fungi, fish, insects, and mammals.

Red blood cells lack the enzymes for the tricarboxylic acid cycle (TCA), also known as the Krebs cycle or the Citric Acid Cycle (CAC), therefore oxidation of glucose for (1) ATP generation is regulated by anaerobic glycolytic pathway and for (2) NADPH generation is limited through the PPP. Selectively, RBCs are affected by G6PD deficiency because several mutations decrease the stability of the G6PD enzyme, and exogenous oxidizing agents in the blood cause oxidative stress to RBCs. Therefore, when G6PD is deficient, RBCS are incapable to tolerate this stress and increase the risk of hemolysis (Carson et al., 1956).

Many mutations affect the G6PD gene (Luzzatto and Arese, 2018), leading to a decrease in its activity and consequently to G6PD deficient phenotype. Many human populations express this phenotype and numerous fundamental mutant alleles are found at different polymorphic frequencies. Fortunately, deficiency in G6PD in most people is asymptomatic and has no clinical manifestations. This remains true until these asymptomatic people are exposed to hemolytic triggers such as fava beans and antimalarial Primaquine drugs (Luzzatto and Arese, 2018).

For decades, favism, the acute lysis of RBCs after ingestion of fava beans, has been a public health issue in areas where G6PD deficiency is widespread and fava beans ingestion is common. Moreover, antimalarial primaquine drug is a contemporary significant public health concern which is the main drug to control malaria infection from Plasmodium falciparum was found to trigger hemolysis in G6PD deficient people (Arese et al., 1986; Kattamis, 1986).

In cancer, the PPP is a significant pathway for its role in providing NADPH and generating pentose phosphates for nucleic acid synthesis. Evidence has shown that, in cancer cells, neoplastic lesions modulate the PPP mutability either directly or indirectly (Patra and Hay, 2014). In fact, enzymatic machinery synchronize to the tumor's intracellular demands to endorse growth, proliferation, and survival of cancer cells in response to the dynamic microenvironment of a tumor (Au et al., 2000a; Coy et al., 2005; Liu et al., 2010; Sukhatme and Chan, 2012).

Accumulating data have shown that in human cancer the high expression of G6PD is associated with poor prognosis (Batetta et al., 1999; Chen et al., 2018; Nagashio et al., 2019b; Pes et al., 2019). Moreover, high expression and activity of G6PD are detected in

many tumors such as hepatocellular carcinoma (HCC), renal cell carcinoma, leukemia, and gastric cancer (Batetta et al., 1999; Hong et al., 2014b; Langbein et al., 2008; Wang et al., 2012). It is demonstrated that cancer cells maintain a high rate of glycolysis, consume a large amount of glucose, and even in the presence of oxygen, cancer cells convert the majority of glucose into lactic acid (Dang et al., 2011). Moreover, cancer cells in several tumor types exhibit alterations in the PPP, TCA cycle, mitochondrial respiratory chain oxidative phosphorylation, and glutaminolysis (Chen and Russo, 2012).

CHAPTER II

LITERATURE REVIEW

A. History of G6PD deficiency

For the past half-century, three disease-causing enzyme deficiencies have been acknowledged in human red blood cells; these are the galactose-1-phosphate uridyltransferase (GALT) (Isselbacher et al., 1956), catalase, and G6PD (Carson et al., 1956). Out of those three disorders, only deficiency in G6PD leads to hemolytic anemia. In 1973, a study estimated that 300 million people worldwide had a deficiency in G6PD, this number increased now (Luzzatto, 1973).

Deficiency in G6PD was discovered after several episodes of hemolytic anemia following the treatment of malaria. Several convergent events lead to the recognition of G6PD deficiency. First, Warburg, Embden, and Meyerhof carefully unraveled the biochemical pathways through which RBCs metabolize sugar (Beutler, 2008). Second, the improvement of the isotopic methods allowed for an accurate estimation of RBCs survival through Chromium 51 (Cr51) method for erythrocytes labeling which was developed by Sterling and Gray (Necheles et al., 1953; Sterling and Gray, 1950). Lastly, the effect of the Korean War and World War II that created importance in developing new antimalarial drugs (Beutler, 2008).

Table 1: WHO classification for G6PD deficiency

WHO Class	*Level of Deficiency*	*Enzyme Activity*	*Severity of Hemolysis*
1	Severe	< 10%	Chronic non-spherocytic Hemolytic Anemia (CNSHA)
2	Severe	< 10%	Intermittent Hemolysis
3	Moderate	10-60%	Intermittent Hemolysis with stressors
4	Mild to None	60-150%	No Hemolysis
5	None	> 150%	No Hemolysis

Deficiency in G6PD was first revealed in African-American populations and this brought the thought that it had a genetic basis since it was limited to one ethnic group (Browne, 1957). The discovery of the rapid screening test for G6PD made it obvious for researchers that this deficiency was also found in the Middle East and Southern Europe and was not exclusive among the African population. In the 1950s, G6PD deficiency was assumed to be a single disorder since it was not common that a single gene could be hit by several mutations. Nevertheless, Marks and Gross proved that this deficiency in African-Americans was much less severe concerning Mediterranean people (Marks and Gross, 1959). An expert committee organized by the WHO standardized electrophoretic method for studying the kinetic properties and the mobility of enzyme proved that the G6PD enzyme is heterogeneous (1967; Kirkman et al., 1960). A WHO publication then released variants classification with respect to the phenotype severity and sorted them into 5 classes

as shown in table 1, ranking from class 1 for the severe deficiency with less than 10% enzyme activity causing chronic nonspherocytic hemolytic anemia to class 5 with increased enzyme activity with more than twice its normal activity. In the 1970s and early 1980s, around 400 different variants of G6PD enzyme were described after WHO standardization procedures (Yoshida and Beutler, 1983).

B. G6PD and the Pentose Phosphate Pathway

The PPP, also known as the phosphogluconate pathway or the Hexose Monophosphate shunt (HMP), is a critical pathway that is branched at the first committed step from glycolysis. Glucose enters the cell through glucose transport (GLUT), is then phosphorylated by Hexokinase to give Glucose 6 phosphate. This significant pathway caught the attention of researchers almost 100 years ago after the several episodes of hemolytic anemia that was thought to be correlated with decreased levels of reduced glutathione (GSH) (Cordes, 1926). The PPP is essential in erythrocytes since it is the only source of NADPH. The NADPH molecule is needed to produce GSH which is a major scavenger of reactive oxygen species (ROS) thus if the activity of PPP is attenuated, then RBCs become more vulnerable to reagents and oxidants that interfere with it.

The PPP, which occurs in the cytosol, involves two branches, the oxidative and the non-oxidative branch as shown in figure 1. The oxidative phase has three irreversible reactions that produce ribonucleotides and NADPH. G6PD is the control site in the oxidative branch and it is the rate-limiting enzyme. Whereas, transaldolase and transketolase (TKT) are the two main enzymes in the non-oxidative branch.

The first reaction in the oxidative branch shows the dehydrogenation of G6P by G6PD to produce 6-phophogluconolactone and NADPH. The former is hydrolyzed into 6-phosphogluconate which undergoes oxidative decarboxylation and yields a second molecule of NADPH and ribulose-5-phosphate (Ru5P). Ru5P is then converted to ribose 5 phosphate (R5P) in the non-oxidative branch (Patra and Hay, 2014). In the first and third reaction in this branch, NADP+ acts as the electron acceptor. Therefore, the PPP yields two molecules of NADPH for every molecule of glucose.

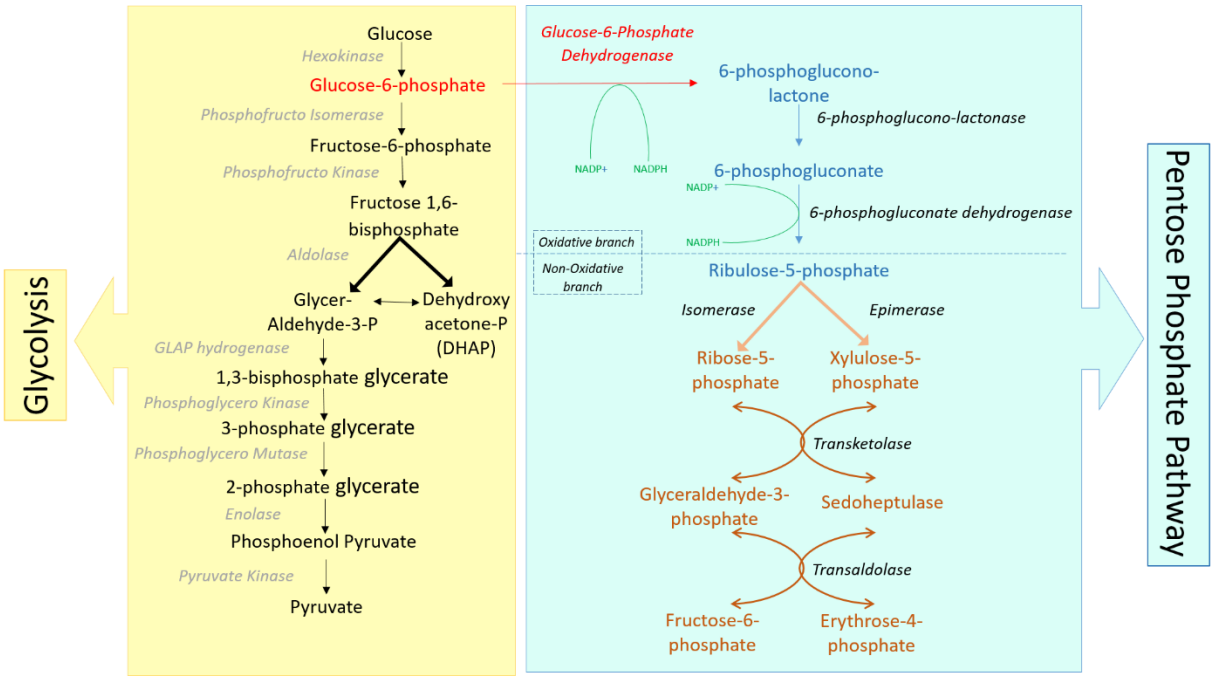

Figure 1: The Pentose Phosphate Pathway branching from Glycolysis. In blue is the oxidative branch and in brown in the non-oxidative branch.

In the non-oxidative branch, a series of reversible reactions take place. The two key enzymes, TKT, and transaldolase, reversibly convert R5P into glyceraldehyde-3-phosphate (G3P) and fructose-6-phosphate (F6P) which are glycolytic intermediates. Therefore, the cellular production of NADPH and R5P is coordinated between the PPP,

glycolysis, and gluconeogenesis. For instance, if the cells need DNA precursors, then it requires R5P more than NADPH. At that point, the glycolytic pathway converts G6P into F6P and G3P and then enters the non-oxidative branch of PPP and converts them to R5P. On the contrary, for reductive biosynthesis or antioxidant defense, the cell needs NADPH more than R5P, R5P is converted into F6P and G3P which through the gluconeogenic pathway G6P is synthesized to enter the oxidative phase of the PPP and thus produce more NADPH (Jiang et al., 2014).

C. Etiology and pathophysiology of G6PD

As mentioned above, deficiency in the G6PD gene was first discovered in the Afro-American population and since it was exclusively found in one ethnic group, it was thought that it had a genetic basis. After the discovery and development of the glutathione stability test, it was proved that transmission was from mother to son (Browne, 1957) and rarely from father to son. However, the father can be affected but can never transmit the genetic defect. Thus, through studies of electrophoretic mobility (Boyer et al., 1962), estimation of enzyme activity (Allison, 1960), and the studies of linkage with color blindness (Adam, 1961) it was confirmed that G6PD deficiency is an X-Linked disorder.

The G6PD gene contains 12 introns and 13 exons; it consists of 515 amino acids with a 59 kDa molecular weight. In 1996, the G6PD 3D model was discovered and published (Naylor et al., 1996) and accordingly the human crystal structure was elucidated (Au et al., 2000b). The G6PD gene shown in figure 2 is found on the long arm of the X chromosome at the telomeric region band Xq28 close to several other genes such as color blindness, hemophilia A, and congenital dyskeratosis (Szabo et al., 1984; Trask et al.,

1991). In a pH-dependent equilibrium, the enzyme is active as a dimer (two monomers) or tetramer (four monomers), with each monomer composed of 515 amino acids. For normal G6PD activity, the stability of the active quaternary structure is crucial (Luzzatto, 2001). The activity of the G6PD enzyme is dependent on the ratio of NADP+/NADPH (1:1ratio), for each molecule of NADP+ consumed results in the production of 1 molecule of NADPH at the same time. Furthermore, stabilizing the proper conformation of the enzyme is reliant on NADP+ (Au et al., 2000b).

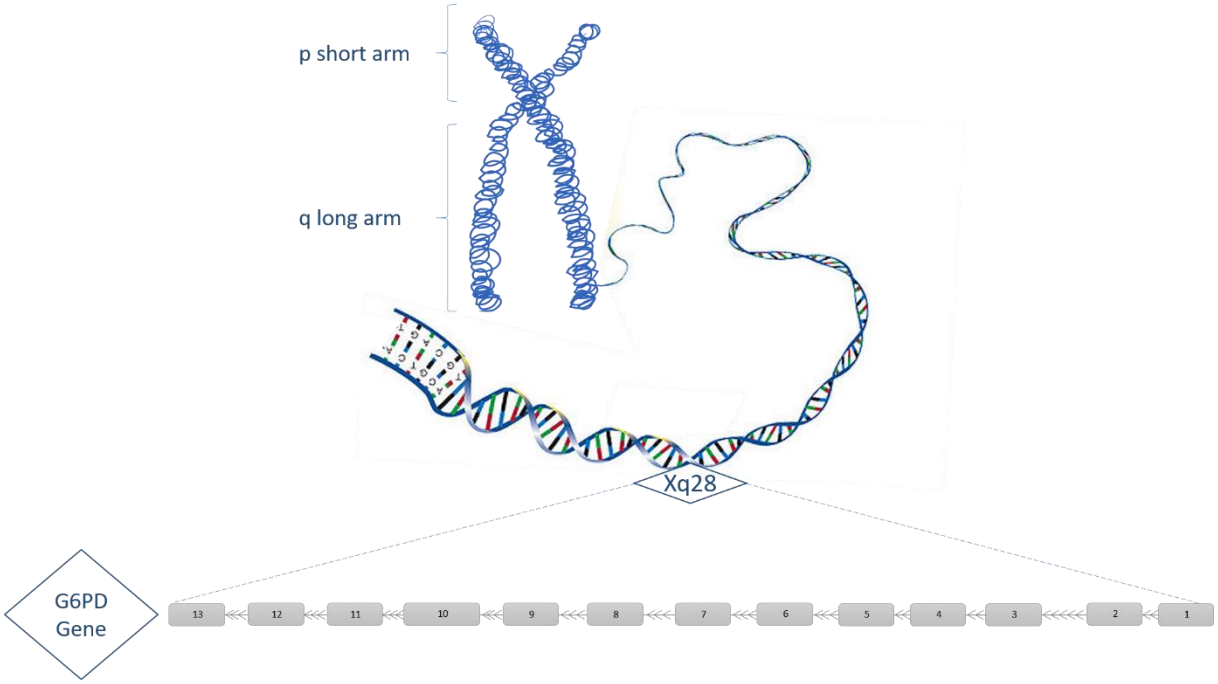

Figure 2: G6PD gene on the Xq28 band long arm of the X chromosome

As previously mentioned, the G6PD enzyme is ubiquitously found in all cells. In healthy individuals with healthy erythrocytes, the enzyme works at 1-2% of its maximal potential. Around 400 different phenotypes and 140 mutations are divided into five classes according to clinical manifestation and enzyme activity. Often the mutated enzyme has

impaired folding, lower stability, or diverse kinetic parameters (Cappellini and Fiorelli, 2008; Mason et al., 2007).

The PPP provides around 85% of pentose for DNA synthesis and this characteristic is essential for rapidly dividing cells and is high in cancer cells (Raïs et al., 1999). Thus, the greater the proliferation of cells, the greater it requires NADPH and ribose 5 phosphate for the synthesis of nucleic acids and then the flux of glucose into the PPP. The physiological high rates of PPP in lipid synthesizing tissues and in the liver is due to the consumption of NADPH as a reducing agent in many synthetic pathways of cholesterol, fatty acid, and many detoxification reactions. Red blood cells also have a high flux of PPP since it is constantly exposed to oxidative stress, and they are the first to be affected by PPP impairment (Riganti et al., 2012). Nevertheless, in areas with endemic malaria, RBCs are protected from being infected with the Plasmodium parasite due to the high persistence of the G6PD deficient phenotype that has deficient NADPH regenerating system (Cappellini and Fiorelli, 2008).

D. G6PD in Favism

In the 5th century B.C, the great mathematician/philosopher, Pythagoras of Samos, may have been first in stating emphatically that fava beans could be dangerous and even fatal to humans (Meletis and Konstantopoulos, 2004; Simoons, 1998). A story often repeated says that when Pythagoras was pursued by his enemies, he stopped at a field of beans and preferred to be taken then to enter the field and that what happened (Arie, 1961).

There are two main actors in favism, the red cell, and the fava bean. Favism occurs when a G6PD deficient person is exposed to substances in the beans which leads to acute

hemolytic anemia. Favism occurs after the ingestion of raw fresh beans, cooked beans (rare cases), or through breastfeeding. All Favism patients are deficient in G6PD, but many G6PD deficient people can eat fava beans (Beutler, 2008). Thus, there are several factors needed to be taken into consideration in favism pathogenesis.

Favism mainly occurs in areas where fava beans are popular and where the frequency of G6PD is high including the Middle East, Southeast Asia, and in Southern Europe. However, this is not true in Northern Germany where G6PD deficiency is rare and fava beans are grown and in West Africa where G6PD deficiency is prevalent but fava beans are not planted (Luzzatto and Arese, 2018). During the past 40 years, several publications showed that favism was reported in 35 countries and more than 3000 cases among children, and this resulted in acute hemolytic anemia which is by most due to G6PD deficiency and favism (Luzzatto and Arese, 2018).

As a rule, G6PD deficient people stay asymptomatic until acute hemolytic anemia appears due to fava beans. It has been suggested that the components of the beans that cause hemolysis are Divicine and Isouramil. As shown in figure 3, these components are transferred to the blood through the intestinal epithelium and thus produce reactive oxygen species such as hydrogen peroxide which in turn oxidize glutathione and NADPH (Albano et al., 1984; Chevion et al., 1982; Winterbourn et al., 1986). In the normal activity of

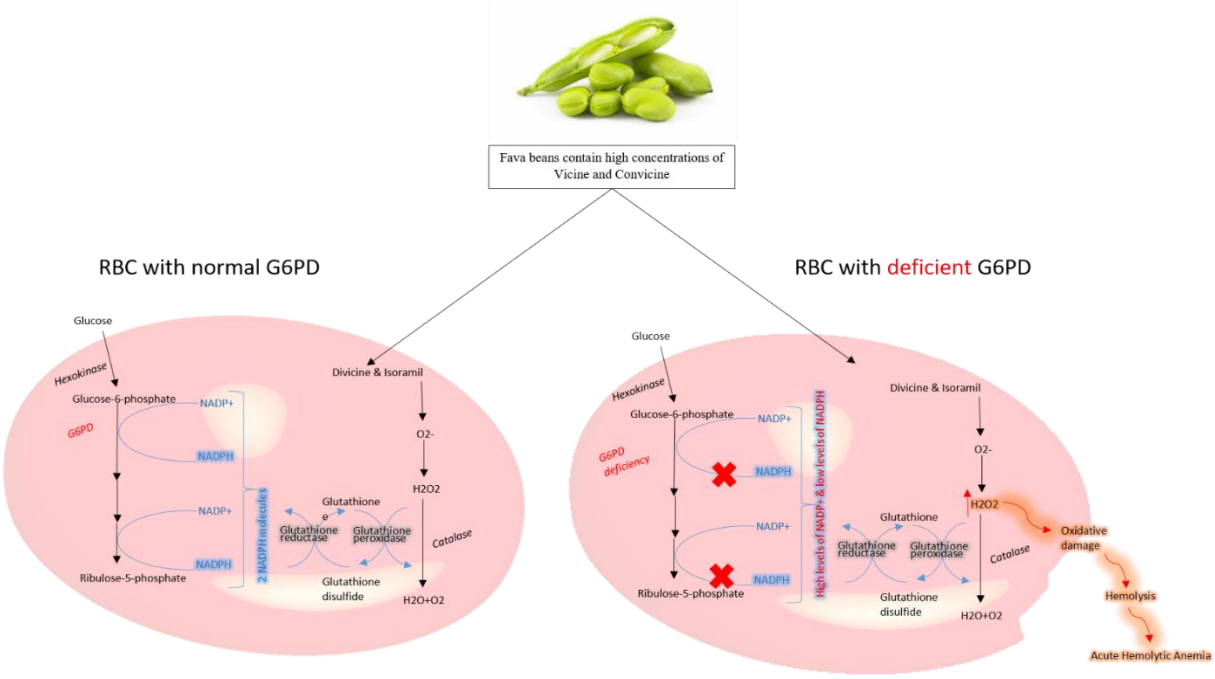

Figure 3: Function of normal and G6PD deficient red blood cells with respect to components of fava beans (Luzzatto and Arese, 2018)

G6PD, hydrogen peroxide is detoxified by glutathione peroxidase and catalase, which are both dependent on NADPH. In G6PD deficient cells, shortage of NADPH supply is reduced and thus the cells are unable to reverse the depletion of glutathione which in turn undergo oxidative damage and result in acute hemolytic anemia as shown in the figure 3 (Chevion et al., 1982; Gaetani et al., 1989; McMillan et al., 2001).

E. G6PD enzyme and its involvement in cancer

In western countries, about one out of three people will have cancer in their lifespan and at least one out of five will die from it. Around 90% of cancer-associated deaths and the ultimate step of tumor progression is due to metastasis (Gupta and Massagué, 2006). Cancer refers to a group of heterogeneous diseases that originate from different tissues and affect different cellular subtypes. In 2000, Hanahan and Weinberg detailed common hallmarks of cancer that simplified the biological understanding of the disease, and it was updated in 2011 (Hanahan and Weinberg, 2000, 2011). The hallmarks included abnormal metabolic activities and mentioned the ability of a tumor to metastasize from a primary tumor to a secondary tumor at a distant site.

One century ago, Otto Warburg described the deregulation of metabolic fluxes of cancer and he reported that cells in cancer converts aerobically glucose to lactate while normal cells use glucose for oxidative metabolism (Wind and Negelein, 1927).

Glycolysis is connected to the PPP metabolic pathway which is a major catabolic pathway that links metabolism of glucose to the synthesis of ribose, nucleotide precursor, and the production of NADPH which in turn is important for both reductive biosynthesis and antioxidant defense. It has been demonstrated that the PPP, along with glycolysis, coordinates the flux of glucose and supports energy production and the cellular biogenesis of macromolecules (Jiang et al., 2014). Glycolysis delivers energy for biogenesis however the uncontrolled proliferation of cancerous cells is maintained from the large amounts of lipids, for the cell membranes construction and energy storage, and nucleotide precursors for continuous DNA replication. Therefore, cancer cells are reprogrammed to metabolically direct glucose flux into the PPP to meet these biosynthetic demands (Jiang et al., 2014).

Research has shown that patients with HCC have high expression of G6PD which is linked with poor prognosis and metastases as well as knockdown of G6PD in HCC cell lines in vitro resulted in the inhibition of proliferation, migration, and invasion (Dore et al., 2018; Lu et al., 2018). Moreover, Lu et al. showed that G6PD is correlated with metastases and invasion in HCC through the activation of signal transduction and activator of transcription 3 (STAT3) pathway to induce epithelial to mesenchymal transition (EMT) (Lu et al., 2018). Another study demonstrated that molecular subtypes of breast cancer are closely associated with G6PD and elevated expression of G6PD in breast cancer is a negative prognostic factor (Dong et al., 2016; Pu et al., 2015). A study done by Benito et al. revealed that the silencing of G6PD in breast cancer cells increases glutamine uptake, increases glycolytic flux, and decreases lipid synthesis (Benito et al., 2017). Likewise, in lung cancer, overexpression of G6PD is associated with a poor survival rate of these patients compared to the normal expression of G6PD with the same type of cancer (Nagashio, 2019).

F. G6PD as a double agent

Glycolysis and PPP are regulated coordinately to support cell survival and cell growth. In cancer cells, glycolysis activation may be linked to an increase in the activity of PPP for biosynthesis. Often, cancer cells go around growth checkpoints through genetic mutations in vital genes. P53 is one of the frequent mutations that occur and leads to the enhancement of both PPP and glycolytic flux (Bensaad et al., 2006; Jiang et al., 2011).

Several signaling pathways and extracellular stimuli modulate the activity and regulate the expression of G6PD through post-translational mechanisms. Studies showed

that growth factors such as epidermal growth factor (EGF) and platelet-derived growth factor (PDGF) induce bound G6PD to be released into a soluble fraction, and thus increase in its activity (Stanton et al., 1991; Tian et al., 1994). Ras and phosphatidylinositol-3-kinase (PI3K) appear to mediate this effect and also the tyrosine kinase Src can phosphorylate G6PD directly and induce its translocation (Pan et al., 2009). Moreover, 5' AMP-activated protein kinase (AMPK) which has a crucial regulatory role in energy homeostasis, regulates G6PD expression (Kohan et al., 2009; Stanton, 2012) and negatively regulates aerobic glycolysis in cancer cells (Faubert et al., 2013). Therefore, in cancer cells, the hyperactivated pro-oncogenic signaling pathways accelerates the PPP by regulation positively G6PD enzyme.

In some type of cells, the activity of G6PD is down-regulated by cyclic AMP (cAMP) directly and indirectly. Protein kinase A (PKA) is activated by cAMP and phosphorylates directly G6PD on threonine and serine residues and thus inhibits the activity of G6PD (Xu et al., 2005). Besides, the transcription of the G6PD gene is inhibited by cAMP through the cAMP response element within the promoter region of the gene (Zhang et al., 2000).

Cancer cells have evolved mechanisms that regulate the PPP to fulfill their need for nucleic acids, fatty acids, and NADPH. As mentioned earlier, pro-oncogenic signaling pathways are hyperactivated and in turn, promote the activation of G6PD by post-translational mechanisms. Nevertheless, there are several other mechanisms by which oncoproteins and tumor suppressor proteins influence the PPP.

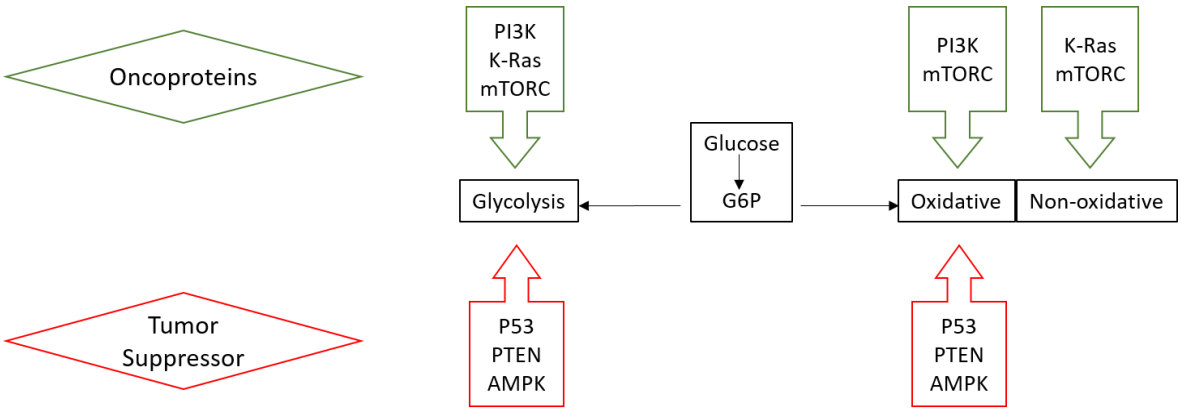

Figure 4: Regulation of oncoproteins and Tumor Suppressor genes on PPP and Glycolysis

a. Tumor suppressor p53

The tumor suppressor p53 is involved in the regulation of PPP as it binds to the promoter region of several genes of the PPP. The expression of GLUT1 and GLUT 4, the glucose transporter genes, is inhibited directly by p53 (Schwartzenberg-Bar-Yoseph et al., 2004). Therefore, glucose uptake is increased when p53 is mutated in cancer cells and high glucose level is shifted into Glycolysis and PPP. On the other hand, p53 indirectly suppresses the expression of phosphoglycerate mutase 1, which converts 3-phosphoglycerate (3PG) to 2-phosphoglycerate (2PG), in glycolysis and thus inhibits the oxidative arm of PPP (Hitosugi et al., 2012; Kondoh et al., 2005). In addition to this, p53 transcriptionally induce TP53-induced glycolysis and apoptosis regulator (TIGAR) and inhibits glycolysis and, therefore, directs the metabolites to the PPP to increase NADPH production and decrease ROS (Bensaad et al., 2006).

b. Oncogenic Ras

Many cancers exhibit activating mutations in K-Ras such as pancreatic, colon, and lung cancers. Studies in a mouse model with pancreatic cancers on the metabolic consequences of K-Ras activation showed that the oxidative branch of PPP is unaffected while the non-oxidative PPP is activated (Ying et al., 2012). Accordingly, these pancreatic cancer cells mainly use the non-oxidative PPP to produce nucleotides via ribulose-5-phosphate isomerase (RPI) and ribulose-5-phosphate epimerase (RPE) enzymes to generate ribose-5-phosphate (R5P) or xylulose-5-phosphate (Xu5P), respectively for nucleic acid biosynthesis (Ying et al., 2012).

Furthermore, oncogenic Ras induces the expression of hexokinase (HK) enzyme required to produce G6P from glucose to facilitate the oxidative and non-oxidative branches of PPP in cancer cells (Patra et al., 2013). This study showed that tumor burden was reduced in the K-Ras induced mouse model with lung cancer after HK genetic ablation, it maintained NADPH production by oxidative PPP while impaired glucose-dependent ribonucleotide synthesis via the non-oxidative PPP (Patra et al., 2013).

c. The phosphatase and tensin homolog (PTEN)

A p53 target gene, PTEN, is a tumor suppressor that is mutated frequently or deleted in cancers (Bonneau and Longy, 2000; Simpson and Parsons, 2001). In 2012, a study showed that an increase in PTEN leads to an increase in mitochondrial oxidative phosphorylation through PI3K- dependent and independent pathway and also decrease in glucose and glutamine uptake (Garcia-Cao et al., 2012). Moreover, PTEN like p53, has a role in suppressing the enzymatic activity of G6PD (Hong et al., 2014a).

d. Mammalian target of rapamycin complex 1 (mTORC1)

In cancer cells, mTORC1 is often activated due to the activation of the PI3K/AKT signaling pathway. Metabolic profiles and gene expression showed that activation of mTORC1 increases the activity of sterol regulatory element-binding protein (SREBP) transcription factor and in turn leads to substantial up-regulation of the oxidative branch of the PPP and transcription of the gene encoding G6PD (Düvel et al., 2010). This activity provides NADPH for the synthesis of fatty acids.

e. Nuclear factor erythroid 2 (NF-E2)-related factor 2 (Nrf2)

Nrf2 activates oxidative and non-oxidative branches of PPP since it promotes the transcription of G6PD, 6PGDH, TKT, TALDO enzymes and thus increases the production NADPH and nucleotides (Mitsuishi et al., 2012b). Elevated expression of Nrf2 is due to several mechanisms such as oncogenic K-Ras, overexpression of Myc, and hyperactivation of PI3K/Akt signaling pathway in cancer (DeNicola et al., 2011; Mitsuishi et al., 2012a). Therefore, in cancer cells, several mechanisms activate Nrf2, and this leads to the activation of PPP branches.

In summary, inactivation of tumor suppressors or the activation of oncoproteins leads to a plethora of mechanisms that regulate the activity and expression of the PPP enzymes in both branches. Consequently, the anabolic demands of cancer cells are met through these mechanisms.

CHAPTER III

HYPOTHESIS AND AIMS

The PPP is a critical pathway for cancer cells because it does not only supply those cells with ribose for their high demand of nucleic acid synthesis but also provides NADPH which is required for the synthesis of fatty acids and cell survival under stress conditions. The key to that pathway is the G6PD. Therefore, if mutated, G6PD will cut the supply for cancer cells that, in turn, will undergo apoptosis. Consequently, the purpose of this work is to explore if patients with G6PD deficiency, in favism and anemia, are protected against cancer. The research aims at mapping and curating all G6PD mutations causing the deficiency, extracting G6PD mutations that have been identified in patients with different cancer types, and identifying G6PD mutations causing Favism/anemia with a putative protective role against cancer.

CHAPTER IV

METHODS OF DATA COLLECTION AND ANALYSIS

A. G6PD Database

The G6PD database is an accessible database made by the Institute of Structural and Molecular Biology at University College London (UCL). It integrates mutational and structural data from various genetics and structural databases such as the Genbank, Protein data bank, etc… and from several publications. Note that the data of this database is old and has not been updated since 2004. This database contains a list of 107 mutations affecting the G6PD gene and display information about the name of the mutation, class of deficiency, location of mutation (DNA and amino acid), secondary structure, and conservation. It also allows us to filter the data by a class of mutation (I, II, III, IV), by residue number, by amino acid, and by mutation name. This database was cross-linked with all other G6PD mutations to remove the redundancy in the information. This database is accessible through (http://www.bioinf.org.uk/mutations/g6pd/db/)

B. UCSC genome browser

The UCSC genome browser is a freely accessible portal that gives access to genomes and its information. It contains different tools and assemblies of model organisms for viewing, analyzing, and downloading data. This database is used to extract all the mutations (Clinvar variants) that are found on the G6PD gene, on introns and exons. Each mutation found on the gene is given with its location, clinical significance, allele ID, dbSNP ID, type of variant, molecular consequence, phenotype, and nucleotide and protein

HGVS. Each mutation is found with a link to the ClinVar database for further information about submitters and references. The browser is accessed through this link (https://genome.ucsc.edu/), from the genomes tab, human GRCh38 is selected then the search for the desired gene happens.

C. ClinVar

ClinVar is a freely accessible, public archive of reports of the relationship among human variations, phenotypes, interpretation, and evidence. It eases access to the relationships proclaimed between human variation and observed health status, in addition to the interpretation's history. The allele submitted in ClinVar is mapped to reference sequences and reported according to the standards of the HGVS. ClinVar is accessed on this link (https://www.ncbi.nlm.nih.gov/clinvar/) then you can search it for gene symbols, HGCV expression, conditions, and more.

D. GDC Data Portal and the Cancer Genome Atlas (TCGA)

The Genomic Data Commons (GDC) Data Portal is a robust data-driven platform full of vigorous information for analysis and download. Also, it allows cancer researchers and bioinformatics to search cancer data from 65 different projects containing 84,031 cases with cancer in 67 different primary sites. The latter includes sites from the brain, breasts, lungs, liver, kidneys, glands, and many others. One of the projects in GDC is The Cancer Genome Atlas (TCGA) which is an overview of 33 different cancer types among 11,315 cases. This data portal is used to extract G6PD mutations found in cancer patients and crosslink them with G6PD mutations causing favism to come up with common mutations

between both. The impact of each mutation will be added these include variant effect predictor (VEP), sorting intolerant from tolerant (SIFT), and polyphen. VEP impact determines the effect of the variants on genes, transcripts, protein sequence, and regulatory regions; Variants such as SNPs, insertions, deletions, CNVs, or structural variants. SIFT is a score that predicts whether an amino acid substitution affects protein function and it ranges from 0.0 (deleterious) to 1.0 (tolerated). PolyPhen is a score that predicts the possible impact of an amino acid substitution on the structure and function of a human protein to know if the substitution is damaging, the score ranges from 0.0 (tolerated) to 1.0 (deleterious). TCGA is retrieved from (https://portal.gdc.cancer.gov/) then G6PD gene is explored.

E. The Human Protein Atlas

In 2003, a Swedish- based program initiated The Human Protein Atlas to map all proteins in cells, tissues, and organs. This mapping used the integration of many omics techniques, including mass spectrometry-based proteomics, antibody-based imaging, system biology, and transcriptomics. This is a freely accessible data portal to ease the access of scientists and researchers to explore the human proteome. This data portal includes six discrete parts, and these are the tissue atlas, the cell atlas, the pathology atlas, the blood atlas, the metabolic atlas, and the brain atlas. The tissue atlas shows the distribution of a protein across all major tissues and organs in the body. While the cell atlas shows, in a single cell, the subcellular localization of proteins. As for the pathology atlas, it shows the protein level impact on cancer patient's survival and this part will be used to compare the survival rates as a function of G6PD expression in different cancer types. This

database is easy in visualizing plots and expression of proteins and is retrieved from (https://www.proteinatlas.org/).

F. R Program for Statistical Computing

R is a free highly extensible software environment for statistical computing and graphics. R provides a wide variety of graphical techniques and statistical techniques such as classical statistical tests, linear and nonlinear modeling, time series analysis, clustering, classification, and much more. This program will be used to compare the expression of a mutation in different cancer patients and come up with a plot to visualize those differences.

CHAPTER V

RESULTS

A. Extraction and manual curation of G6PD mutations Database

We first compiled the list of characterized G6PD mutations from an existing dedicated database (http://www.bioinf.org.uk/mutations/g6pd/db/) and added newly identified mutations characterized in newly published G6PD studies. The final list consisted of a total of 151 mutations (figure 5). Each exon and intron are tagged with different mutations along with the alternative allele. Each SNP is presented with different color with respect to their clinical significance with, for instance, black for uncertain significance, red for pathogenic mutations, and green for benign.

Moreover, since the G6PD database was released in 2004, we first decided to update it by manually curating all entries. For this, we performed literature and databases search for each entry and added new information including among others the genomic location to the most recent genome assembly (hg38), the clinical significance from recent studies, and genome frequency from whole exome or whole genome sequencing from different consortia (Table 2).

The new database listing, shown in table 2, included: chromosome, genomic location start and end, gene assembly, gene location, allele ID, clinVar variation, a reference sequence, clinical significance, type of variant, gene symbol, dbSNP ID, clinVar allele submission, genetic testing registry, phenotypes, name of mutation, other identifiers, data origin, cytogenic status, HGVS transcript and protein ID, nucleotide and protein

HGVS study name and collection method, allele frequency from different studies, and reference/citation for each mutation.

The distribution of mutations in G6PD showed 80% of mutations on exons and 20% on introns. We then checked for the clinical significance of the compiled mutations (Figure 5). Interestingly 37% of G6PD mutations are classified as pathogenic, 32% with uncertain significance, 9% as benign mutations, 4% have conflicting interpretation of pathogenicity, 2% have no interpretation for single variant, and 1% are listed as drug response mutations (A general term for a variant that affects a drug response, not a disease.).

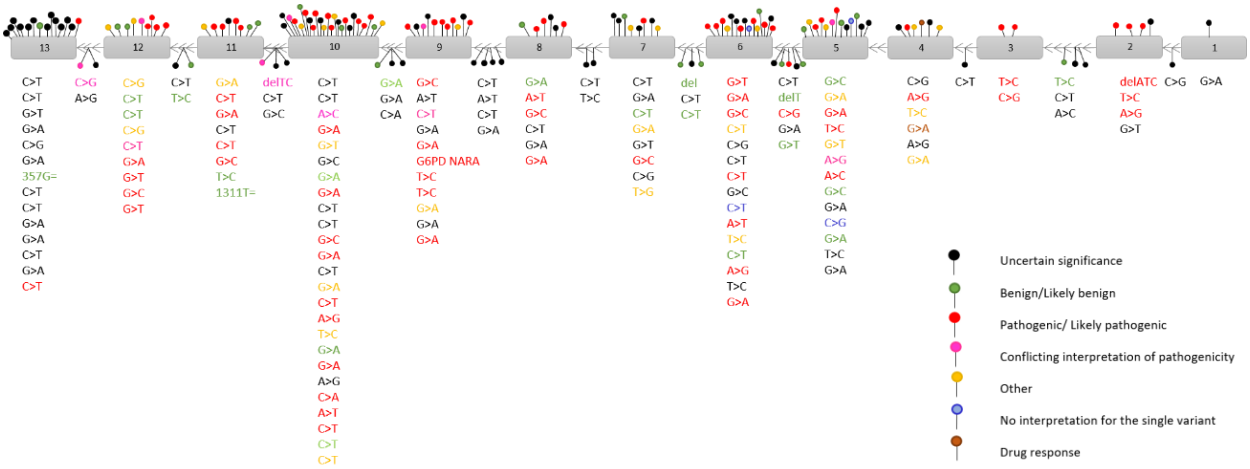

Figure 5: Total of 151 mutations on the introns and exons of the G6PD gene

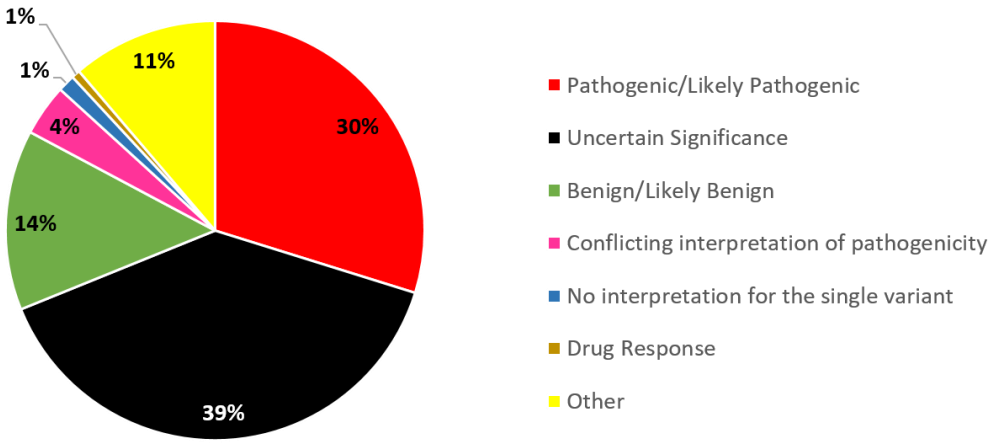

Figure 6: Clinical significance of 151 G6PD mutations

Table 2: G6PD mutations database. Chr: chromosome, Loc: location, Ex: exon, In: intron, CS: clinical significance, US: uncertain significance, Ben: benign, Conflict: conflict interpretation of pathogenicity, Patho: pathogenic, O: other, SNV: single nucleotide variants, Del: deletion, G6PD def: glucose-6-phosphate deficiency, NP: not provided, CNSHA: chronic non-spherocytic hemolytic anemia, GPI: glucose phosphate isomerase, HA: hemolytic anemia

Chr	Start	End	Loc	Nucleotide_HGVS	CS	Type	dbSNP_ID	Pheno	TopMed	gnomAD	gnomAD Exomes	1000 Genomes	ExAc
ChrX	154531401	154531401	Ex 13	c.*599C>T	US	SNV		G6PD def					
ChrX	154531452	154531452	Ex 13	c.*548C>T	US	SNV		G6PD def					
ChrX	154531465	154531465	Ex 13	c.*535G>T	US	SNV	rs1057515820	G6PD def	0.00	0.00			
ChrX	154531480	154531480	Ex 13	c.*520G>A	US	SNV	rs782764609	G6PD def	0.00			0.000	
ChrX	154531534	154531534	Ex 13	c.*466C>G	US	SNV		G6PD def					
ChrX	154531635	154531635	Ex 13	c.*365G>A	US	SNV	rs1034742794	G6PD def	0.00	0.00			
ChrX	154531643	154531643	Ex 13	c.*357G=	Ben	SNV	rs1050757	G6PD def	0.64	0.68		0.59	
ChrX	154531719	154531719	Ex 13	c.*281C>T	US	SNV		G6PD def					
ChrX	154531925	154531925	Ex 13	c.*75C>T	US	SNV		G6PD def					
ChrX	154531950	154531950	Ex 13	c.*50G>A	US	SNV	rs201294737	G6PD def	0.00023	0.0003	0.00011	0.001	0.0003
ChrX	154531953	154531953	Ex 13	c.*47G>A	US	SNV	rs398123543	G6PD def	0.00006	0.0001	0.00007		0.0001
ChrX	154531975	154531975	Ex 13	c.*25C>T	US	SNV	rs781866772	G6PD def	0.00015	0.0001	0.0001		0.0001
ChrX	154532068	154532068	Ex 13	c.1480G>A	US	SNV		G6PD def					
ChrX	154532083	154532083	Ex 13	c.1465C>T	Patho	SNV		NP					
ChrX	154532103	154532103	In 12	c.1458-13C>G	Conflict	SNV	rs371772243	NP	0.00002	0.0009	0.00033	0.001	
ChrX	154532154	154532154	In 12	c.1457+34A>G	US	SNV	rs398123548	NP					0
ChrX	154532203	154532203	Ex 12	c.1532C>G	O	SNV	rs137852348	Other					
ChrX	154532214	154532214	Ex 12	c.1521C>T	Ben	SNV	rs77214077	CNSHA,NP	0.02571	0.0212	0.0065	0.025	0.0094
ChrX	154532240	154532240	Ex 12	c.1405C>T	Ben	SNV	rs369482861	NP	0.00002	0.0002	0.00006		
ChrX	154532245	154532245	Ex 12	c.1490C>G	O	SNV	rs137852344	Other					0
ChrX	154532247	154532247	Ex 12	c.1398C>T	Conflict	SNV	rs398123547	NP	0.00002		0.0014	0.001	0.0014
ChrX	154532257	154532257	Ex 12	c.1478G>A	Patho	SNV	rs72554664	CNSHA	0.00022	0.0003	0.00051	0.001	0.0004
ChrX	154532267	154532267	Ex 12	c.1378G>T	Patho	SNV	rs1603411214	CNSHA					
ChrX	154532269	154532269	Ex 12	c.1466G>C	Patho	SNV	rs72554665	NP	0.00001	0.0005	0.00001	0.002	
ChrX	154532269	154532269	Ex 12	c.1466G>T	Patho	SNV	rs72554665	CNSHA,G6PD def	0.00045	0.0005	0.00001	0.002	
ChrX	154532285	154532285	In 11	c.1455-5C>T	US	SNV	rs886044905	NP					
ChrX	154532293	154532293	In 11	c.1455-13C=	Ben/ Ben	SNV	rs2071429	G6PD def	0.6355	0.6702	0.76831	0.585	0.7456
ChrX	154532389	154532389	Ex 11	c.1451G>A	O	SNV	rs137852324	Other					
ChrX	154532390	154532390	Ex 11	c.1360C>T	Patho	SNV	rs398123546	CNSHA,G6PD def	0.00015	0.0001	0.00015	0.001	0.0001
ChrX	154532411	154532411	Ex 11	c.1339G>A	Patho	SNV	rs137852317	CNSHA					

Chr	Start	End	Loc	Nucleotide_HGVS	CS	Type	dbSNP_ID	Pheno	TopMed	gnomAD	gnomAD Exomes	1000 Genomes	ExAc
ChrX	154532421	154532421	Ex 11	c.1329C>T	US	SNV	rs138919671	NP	0.00006	0		0	0
ChrX	154532432	154532432	Ex 11	c.1408C>T	Patho	SNV	rs1557229599	NP					
ChrX	154532434	154532434	Ex 11	c.1316G>C	Patho	SNV	rs137852337	CNSHA					0
ChrX	154532439	154532439	Ex 11	c.1311T>C	Ben	SNV	rs2230037	CNSHA	0.83433	0.8451	0.83522	0.784	0.8328
ChrX	154532439	154532439	Ex 11	c.1311T=	Ben	SNV	rs2230037	G6PD def	0.83433	0.8451	0.83522	0.784	0.8328
ChrX	154532471	154532472	In 10	c.1378-10_1378-9de	Conflict	Del	rs199586268	CNSHA,G6PD def	0.00931	0.0074	0.00227	0.008	0.003
ChrX	154532547	154532547	In 10	c.1287+20C>T	US	SNV	rs1358637530	NP	0.00001				
ChrX	154532564	154532564	In 10	c.1287+3G>C	US	SNV		G6PD def					
ChrX	154532609	154532609	Ex 10	c.1245C>T	US	SNV	rs147131392	NP	0.00087	0.0005	0.00015	0	0.0002
ChrX	154532611	154532611	Ex 10	c.1243C>T	US	SNV	rs1603411292	CNSHA					
ChrX	154532614	154532614	Ex 10	c.1240A>C	Conflict	SNV	rs201794043	CNSHA,G6PD def	0.00		0.00		0.00
ChrX	154532625	154532625	Ex 10	c.1319G>A	Patho, O	SNV	rs137852336	CNSHA					
ChrX	154532626	154532626	Ex 10	c.1318G>T	O	SNV	rs137852323	Other					
ChrX	154532645	154532645	Ex 10	c.1209G>A	Ben	SNV	rs1355739430	CNSHA	0.00				
ChrX	154532662	154532662	Ex 10	c.1192G>A	Patho	SNV	rs137852325	CNSHA					
ChrX	154532667	154532667	Ex 10	c.1277C>T	US	SNV	rs1557229683	CNSHA					
ChrX	154532674	154532674	Ex 10	c.1180G>C	Patho, O	SNV	rs137852335	CNSHA					
ChrX	154532676	154532676	Ex 10	c.1268G>A	Patho	SNV	rs137852316	CNSHA					
ChrX	154532690	154532690	Ex 10	c.1254C>T	US	SNV	rs782623392	G6PD def			0.00003		0
ChrX	154532694	154532694	Ex 10	c.1160G>A	O	SNV	rs137852321	Other					
ChrX	154532695	154532695	Ex 10	c.1249C>T	Patho, O	SNV	rs137852334	CNSHA					
ChrX	154532698	154532698	Ex 10	c.1156A>G	Patho	SNV	rs137852320	NP					
ChrX	154532701	154532701	Ex 10	c.1243T>C	O	SNV	rs137852322	Other					
ChrX	154532702	154532702	Ex 10	c.1152G>C	US	SNV	rs368832453	NP	0.00	0.00	0.00		0.00
ChrX	154532738	154532738	Ex 10	c.1206G>A	Ben	SNV	rs2230036	G6PD def	0.03593	0.0277	0.00862	0.033	0.012
ChrX	154532752	154532752	Ex 10	c.1102G>A	Patho	SNV	rs387906468	CNSHA					
ChrX	154532753	154532753	Ex 10	c.1101C>T	US	SNV		G6PD def					
ChrX	154532758	154532758	Ex 10	c.1096A>G	US	SNV	rs1057518975	Low GPI activity,HA					
ChrX	154532765	154532765	Ex 10	c.1179C>A	Patho, O	SNV	rs137852329	CNSHA					
ChrX	154532766	154532766	Ex 10	c.1088A>T	Patho	SNV	rs1557229736	Inborn genetic dx					

Chr	Start	End	Loc	Nucleotide_HGVS	CS	Type	dbSNP_ID	Pheno	TopMed	gnomAD	gnomAD Exomes	1000 Genomes	ExAc
ChrX	154532772	154532772	Ex 10	c.1082C>T	Patho, O	SNV	rs137852345	CNSHA					
ChrX	154532789	154532789	Ex 10	c.1065C>T	Ben	SNV	rs1557229753	NP			0.00		
ChrX	154532797	154532797	Ex 10	c.1057C>T	O	SNV	rs137852333	Other					
ChrX	154532806	154532806	In 9	c.1052-4G>A	Ben	SNV	rs372124193	CNSHA	0.00	0.00	0.00		0.00
ChrX	154532926	154532926	In 9	c.1051+16G>A	US	SNV	rs782637386	NP	0.00006		0.00008		6E-05
ChrX	154532927	154532927	In 9	c.1051+15C>A	US	SNV		G6PD def					
ChrX	154532945	154532945	Ex 9	c.1048G>C	Patho	SNV	rs34193178	CNSHA	0.0013	0.001	0.00001	0.001	
ChrX	154532956	154532956	Ex 9	c.1127A>T	US	SNV	rs398123544	NP					
ChrX	154532969	154532969	Ex 9	c.1024C>T	Conflict, O	SNV	rs137852342	NP	0.00006	0.0001	0.00012	0.001	0.0002
ChrX	154532972	154532972	Ex 9	c.1021G>A	US	SNV	rs782174983	NP	0.00	0.00	0.00		0.00
ChrX	154532990	154532990	Ex 9	c.1003G>A	Patho	SNV	rs5030869	Inborn genetic dx	0.00002		0.0002		0.0002
ChrX	154533013	154533036	Ex 9	c.957_980del	Patho, O	Del	rs587776730	CNSHA					
ChrX	154533025	154533025	Ex 9	c.1058T>C	Patho	SNV	rs76723693	CNSHA,G6PD def	0.0013	0.0014	0.00051	0.003	0.0007
ChrX	154533029	154533029	Ex 9	c.964T>C	Patho, O	SNV	rs137852347	CNSHA					
ChrX	154533044	154533044	Ex 9	c.949G>A	O	SNV	rs137852339	Other	0.00002	0	0.00128	0.002	0.0014
ChrX	154533104	154533104	Ex 9	c.889G>A	US	SNV	rs781975796	NP	0.00		0.00		0.00
ChrX	154533122	154533122	Ex 9	c.871G>A	Patho	SNV	rs137852327	CNSHA, G6PD def	0.00011	0	0.00022	0.002	0.0003
ChrX	154533144	154533144	In 8	c.865-16C>T	US	SNV	rs199970830	NP	0.00003		0.00001	0.001	0
ChrX	154533559	154533559	In 8	c.864+17A>T	US	SNV	rs377041776	NP	0.00009		0.0002		0.0002
ChrX	154533562	154533562	In 8	c.864+14C>T	US	SNV	rs782416820	G6PD def	0.00002		0.00001		0
ChrX	154533571	154533571	In 8	c.864+5G>A	US	SNV	rs372876649	CNSHA	0.00		0.00		0.00
ChrX	154533586	154533586	Ex 8	c.854G>A	Ben	SNV	rs74575103	G6PD def					
ChrX	154533592	154533592	Ex 8	c.848A>T	Patho	SNV	rs1557230040	NP					
ChrX	154533596	154533596	Ex 8	c.934G>C	Patho	SNV	rs137852318	CNSHA	0.00001	0.0006	0.00001	0	
ChrX	154533625	154533625	Ex 8	c.815C>T	US	SNV	rs1603411458	NP					
ChrX	154533627	154533627	Ex 8	c.813G>A	US	SNV		CNSHA					
ChrX	154533634	154533634	Ex 8	c.806G>A	Patho, O	SNV	rs137852346	CNSHA					
ChrX	154533672	154533672	In 7	c.861-3C>T	US	SNV	rs398123551	NP	0.00001				
ChrX	154533680	154533680	In 7	c.861-11T>C	US	SNV	rs782622284	G6PD def	0.00001		0.00002		0
ChrX	154534055	154534055	Ex7	c.750C>T	US	SNV		G6PD def					

Chr	Start	End	Loc	Nucleotide_HGVS	CS	Type	dbSNP_ID	Pheno	TopMed	gnomAD	gnomAD Exomes	1000 Genomes	ExAc
ChrX	154534108	154534108	Ex 7	c.697G>A	US	SNV		CNSHA					
ChrX	154534115	154534115	Ex 7	c.690C>T	Ben	SNV	rs781917123	CNSHA	0.00	0.00	0.00	0.00	0.00
ChrX	154534125	154534125	Ex 7	c.680G>A	O	SNV	rs137852328	Other			0.00005		0
ChrX	154534125	154534125	Ex 7	c.680G>T	Patho	SNV	rs137852328	G6PD def			0.00005		0
ChrX	154534130	154534130	Ex 7	c.675G>C	US	SNV	rs398123550	NP					
ChrX	154534145	154534145	Ex 7	c.660C>G	US	SNV	rs782771682	CNSHA	0.00		0.00		0.00
ChrX	154534157	154534157	Ex 7	c.648T>G	O	SNV	rs137852319	Other			0.00001		
ChrX	154534165	154534165	In 6	c.645-8_645-5del	Ben	Del	rs782160396	NP					
ChrX	154534172	154534172	In 6	c.645-12C>T	US	SNV		G6PD def					
ChrX	154534177	154534177	In 6	c.735-17C>T	Ben	SNV	rs5986875	NP	0.00956	0.0084	0.00001	0.01	
ChrX	154534345	154534345	Ex 6	c.727G>T	Patho	SNV	rs137852326	CNSHA	0.00001		0.00001		0
ChrX	154534389	154534389	Ex 6	c.593G>A	Patho	SNV	rs137852332	Other					
ChrX	154534389	154534389	Ex 6	c.683G>C	Patho, O	SNV	rs137852332	CNSHA					
ChrX	154534390	154534390	Ex 6	c.592C>T	O	SNV	rs137852330	Chronic granuloma	0.00001		0.00003	0	0
ChrX	154534400	154534400	Ex 6	c.582C>G	US	SNV	rs145247580	G6PD def	0.00058	0.0005	0.00047		0.0004
ChrX	154534408	154534408	Ex 6	c.664C>T	US	SNV	rs1557230370	CNSHA			0.00001		
ChrX	154534419	154534419	Ex 6	c.563C>T	Patho/ Path	SNV	rs5030868	CNSHA,G6PD def	0.00039	0.0002	0.00263		0.0027
ChrX	154534437	154534437	Ex 6	c.635G>C	US	SNV	rs98123549	NP	0.00002		0.00002		0
ChrX	154534438	154534438	Ex 6	c.634C>T	No interpre	SNV	rs267606836	Chronic granuloma			0		0
ChrX	154534440	154534440	Ex 6	c.542A>T	Patho	SNV	rs5030872	CNSHA	0.00039	0.0004	0.00013		0.0001
ChrX	154534463	154534463	Ex 6	c.519C>T	Ben	SNV	rs200111236	CNSHA	0.00	0.00		0.00	0.00
ChrX	154534465	154534465	Ex 6	c.607T>C	O	SNV	rs137852343	Other			0.00001		0
ChrX	154534489	154534489	Ex 6	c.493A>G	Patho/ Path	SNV	rs137852331	CNSHA	0.0001	0.0001	0.00005		0.0001
ChrX	154534492	154534492	Ex 6	c.490T>C	US	SNV	rs1603411632	NP					
ChrX	154534495	154534495	Ex 6	c.487G>A	Patho	SNV	rs137852314	CNSHA	0.00002	0	0.00005		0
ChrX	154534510	154534510	In 5	c.486-14C>T	US	SNV	rs200833520	NP	0.00022	0	0.00014		0.0001
ChrX	154534525	154534525	In 5	c.486-29G>T	Ben	SNV	rs370403856	NP	0.00	0.00	0.00	0.00	0.00
ChrX	154534530	154534530	In 5	c.576-34delT	Ben	Del	rs3216174	NP	0.00698	0.0059			0.0089
ChrX	154534556	154534556	In 5	c.576-60C>G	Patho	SNV	rs2515904	G6PD def	0.04753	0.0422		0.048	
ChrX	154535155	154535155	In 5	c.575+13G>A	US	SNV	rs781898381	G6PD def	0.00002	0	0.00002		0

Chr	Start	End	Loc	Nucleotide_HGVS	CS	Type	dbSNP_ID	Pheno	TopMed	gnomAD	gnomAD Exomes	1000 Genomes	ExAc
ChrX	154535176	154535176	Ex 5	c.477G>C	Ben	SNV	rs370918918	CNSHA	0.00	0.00	0.00	0.00	0.00
ChrX	154535187	154535187	Ex 5	c.466G>A	O	SNV	rs137852313	G6PD Ilesha	0.00034	0.0001	0.00026	0	0.0003
ChrX	154535205	154535205	Ex 5	c.448G>A	Patho	SNV	rs1557230573	Congenital HA					
ChrX	154535249	154535249	Ex 5	c.404A>C	Patho	SNV		CNSHA					
ChrX	154535261	154535261	Ex 5	c.392G>T	O	SNV	rs137852341	Other	0.00006	0	0.00001	0	
ChrX	154535270	154535270	Ex 5	c.383T>C	Patho	SNV	rs78365220	CNSHA			1E-06		
ChrX	154535277	154535277	Ex 5	c.466A>G	Conflict,O	SNV	rs1050829	CNSHA,G6PD def	0.00007	0.0873	0.02448	0.095	
ChrX	154535278	154535278	Ex 5	c.375G>C	Ben	SNV	rs782130334	NP	0.00	0.00	0.00		0.00
ChrX	154535316	154535316	Ex 5	c.337G>A	US	SNV	rs5030870	NP	0.00002	0	0.00003		0
ChrX	154535336	154535336	Ex 5	c.317C>G	No interpre	SNV	rs267606835	Chronic granuloma		0	0.00001		0
ChrX	154535342	154535342	Ex 5	c.311G>A	Ben	SNV	rs181277621	CNSHA	0.00	0.00	0.00	0.00	0.00
ChrX	154535348	154535348	Ex 5	c.305T>C	US	SNV	rs886044847	NP			0.00001		
ChrX	154535356	154535356	Ex 5	c.297G>A	US	SNV		G6PD def					
ChrX	154535984	154535984	Ex 4	c.310C>G	US	SNV	rs781848254	NP	0.00004		0.00004		0
ChrX	154535995	154535995	Ex 4	c.299A>G	Patho	SNV	rs782090947	NP	0.00001		0.00002		0
ChrX	154535996	154535996	Ex 4	c.298T>C	O	SNV	rs137852349	Other			0.00002		
ChrX	154536002	154536002	Ex 4	c.292G>A	DR	SNV	rs1050828	CNSHA,G6PD def	0.03704	0.0328	0.00869	0.038	
ChrX	154536011	154536011	Ex 4	c.193A>G	US	SNV		CNSHA					
ChrX	154536032	154536032	Ex 4	c.262G>A	O	SNV	rs137852315	Other					
ChrX	154536128	154536128	In 3	c.248+13C>T	US	SNV	rs990745079	NP	0.00008		0.00001		
ChrX	154536156	154536156	Ex 3	c.233T>C	Patho	SNV	rs76645461	CNSHA	0.00002		0.00001	0	0
ChrX	154536168	154536168	Ex 3	c.131C>G	Patho	SNV	rs78478128	CNSHA	0.00001		0.00018	0	0.0002
ChrX	154542419	154542419	In 2	c.156A>G	Ben	SNV	rs368857323	NP	0.00	0.00	0.00	0.00	0.00
ChrX	154546027	154546027	In 2	c.210+9C>T	US	SNV	rs886044853	NP					
ChrX	154546029	154546029	In 2	c.120+7A>C	US	SNV	rs369904290	G6PD def	0.0001	0.0002		0.007	
ChrX	154546046	154546048	Ex 2	c.103_105delATC	Patho, O	Del	rs137852338	CNSHA					
ChrX	154546068	154546068	Ex 2	c.-16+3681C>A	US	SNV		G6PD def					
ChrX	154546058	154546058	Ex 2	c.98T>C	Patho	SNV	rs398123552	CNSHA			0.00001		
ChrX	154546061	154546061	Ex 2	c.185A>G	O	SNV	rs137852340	Other	0.00016	0.0001	0.00014		
ChrX	154547454	154547454	In 1	c.-15-4534G>C	US	SNV		G6PD def					

B. Extracting cancer-associated G6PD SNPs

To check if G6PD deficiency, in favism and anemia, is protective against cancer, we needed to extract G6PD mutations related to cancer. For this, we compiled all G6PD mutations identified in 33 different cancer types from The Cancer Genome Atlas (TCGA) database (Table 3). The table includes the location of the mutation, the reference and alternative nucleotide, the mutation consequence, the amino acid mutated, number of affected individuals as well as the impact of the mutation (VEP, SIFT, PolyPhen). We then checked for the consequences of the compiled mutations (Figure 7). Overall, 58% of G6PD mutations consequence are classified as missense, 22% with synonymous consequence, 12% as 3'UTR G6PD, 2.67% have frame shift and stop gained mutation consequence, and 0.667% mutation as intron, splice acceptor, splice region, and start lost consequence.

Table 3: G6PD mutation in cancer patients from TCGA. Chr: chromosome, Loc: location, Ref: reference, Alt: alternative, AA: amino acid

Chr	Position	Ref	Alt	Consequence	Mutation	#_cases	Impact	SIFT	PolyPhen
ChrX	154546087	G	A	Synonymous	G53G	4/888,0.45%	Low		
ChrX	154534055	G	A	Synonymous	F280F	4/888,0.45%	Low		
ChrX	154534385	G	A	Synonymous	I229I	3/888,0.34%	Low		
ChrX	154534408	G	A	Missense	R222C	2/888,0.23%	Moderate	Deleterious	Benign
ChrX	154532601	G	A	Missense	S448L	2/888,0.23%	Moderate	Deleterious	Benign
ChrX	154532785	G	A	Missense	R387C	2/888,0.23%	Moderate	Deleterious	Probably_damaging
ChrX	154532585	G	A	Synonymous	T453T	2/888,0.23%	Low		
ChrX	154532761	G	A	Missense	R395C	2/888,0.23%	Moderate	Deleterious	Probably_damaging
ChrX	154535237	T	C	Missense	Y169C	2/888,0.23%	Moderate	Deleterious	Probably_damaging
ChrX	154532204	G	A	Missense	P511S	2/888,0.23%	Moderate	Tolerated	Benign
ChrX	154534108	C	T	Missense	V263I	2/888,0.23%	Moderate	Tolerated	Benign
ChrX	154534151	G	A	Synonymous	N248N	2/888,0.23%	Low		
ChrX	154535180	C	A	Missense	C188F	1/888,0.11%	Moderate	Deleterious	Probably_damaging
ChrX	154533057	G	A	Synonymous	P342P	1/888,0.11%	Low		
ChrX	154546100	delTCTTCCCG		Frameshift	R47Afs*23	1/888,0.11%	High		
ChrX	154533629	C	T	Missense	V301M	1/888,0.11%	Moderate	Deleterious	Possibly_damaging
ChrX	154534069	G	A	Missense	R276C	1/888,0.11%	Moderate	Deleterious	Probably_damaging
ChrX	154531544	G	T	3' UTR G6PD		1/888,0.11%	Modifier		
ChrX	154532605	C	A	Stop Gained	E447*	1/888,0.11%	High		
ChrX	154531464	G	A	3' UTR G6PD		1/888,0.11%	Modifier		
ChrX	154534075	C	G	Missense	E274Q	1/888,0.11%	Moderate	Tolerated	Benign
ChrX	154536016	T	G	Missense	E93A	1/888,0.11%	Moderate	Tolerated	Benign
ChrX	154532772	G	A	Missense	A391V	1/888,0.11%	Moderate	Deleterious	Probably_damaging
ChrX	154535363	T	C	Missense	K127R	1/888,0.11%	Moderate	Tolerated	Benign
ChrX	154531966	G	T	3' UTR G6PD		1/888,0.11%	Modifier		
ChrX	154532013	G	A	Missense	P542L	1/888,0.11%	Moderate	Deleterious	Possibly_damaging
ChrX	154535948	G	A	Missense	P116S	1/888,0.11%	Moderate	Tolerated	Probably_damaging
Chr	Position	Ref	Alt	Consequence	Mutation	#_cases	Impact	SIFT	PolyPhen
ChrX	154531479	A	G	3' UTR G6PD		1/888,0.11%	Modifier		
ChrX	154531501	G	T	3' UTR G6PD		1/888,0.11%	Modifier		
ChrX	154534437	C	A	Missense	R212L	1/888,0.11%	Moderate	Deleterious	Benign
ChrX	154535242	G	T	Synonymous	L167L	1/888,0.11%	Low		
ChrX	154532674	C	T	Missense	V424M	1/888,0.11%	Moderate	Deleterious	Possibly_damaging
ChrX	154532435	G	A	Missense	R469C	1/888,0.11%	Moderate	Deleterious	Probably_damaging
ChrX	154534062	C	T	Missense	G278D	1/888,0.11%	Moderate	Deleterious	Probably_damaging
ChrX	154531890	G	A	3' UTR G6PD		1/888,0.11%	Modifier		
ChrX	154546042	A	C	Synonymous	G68G	1/888,0.11%	Low		
ChrX	154534102	G	T	Missense	L265I	1/888,0.11%	Moderate	Tolerated	Benign
ChrX	154532752	C	T	Missense	E398K	1/888,0.11%	Moderate	Deleterious	Probably_damaging
ChrX	154533069	G	A	Synonymous	Y338Y	1/888,0.11%	Low		
ChrX	154532191	C	A	Missense	G515V	1/888,0.11%	Moderate	Deleterious	Probably_damaging
ChrX	154535336	G	A	Missense	S136F	1/888,0.11%	Moderate	Tolerated	Benign
ChrX	154534462	C	T	Missense	G204R	1/888,0.11%	Moderate	Deleterious	Probably_damaging
ChrX	154534430	delGGACAGCC		Frameshift	R212Qfs*9	1/888,0.11%	High		
ChrX	154534059	T	C	Missense	Y279C	1/888,0.11%	Moderate	Deleterious	Probably_damaging
ChrX	154532018	C	A	Synonymous	V540V	1/888,0.11%	Low		
ChrX	154534133	G	T	Synonymous	I254I	1/888,0.11%	Low		
ChrX	154531975	delG		3' UTR G6PD		1/888,0.11%	Modifier		
ChrX	154534463	G	A	Synonymous	F203F	1/888,0.11%	Low		
ChrX	154546134	T	A	Missense	S38C	1/888,0.11%	Moderate	Deleterious_l	Probably_damaging
ChrX	154532440	T	A	Missense	Y467F	1/888,0.11%	Moderate	Deleterious	Probably_damaging
ChrX	154532221	T	C	Missense	E505G	1/888,0.11%	Moderate	Tolerated	Benign
ChrX	154531918	T	A	3' UTR G6PD		1/888,0.11%	Modifier		
ChrX	154534119	T	C	Missense	N259S	1/888,0.11%	Moderate	Tolerated	Benign
ChrX	154532950	C	A	Missense	R378M	1/888,0.11%	Moderate	Deleterious	Probably_damaging

Chr	Position	Ref	Alt	Consequence	Mutation	#_cases	Impact	SIFT	PolyPhen
ChrX	154532662	C	A	Stop Gained	E428*	1/888,0.11%	High		
ChrX	154533111	G	A	Synonymous	C324C	1/888,0.11%	Low		
ChrX	154536170	C	A	Synonymous	L73L	1/888,0.11%	Low		
ChrX	154534126	G	A	Missense	R257W	1/888,0.11%	Moderate	Deleterious	Probably_damaging
ChrX	154532753	G	A	Synonymous	A397A	1/888,0.11%	Low		
ChrX	154532662	C	T	Missense	E428K	1/888,0.11%	Moderate	Deleterious	Probably_damaging
ChrX	154535940	G	A	Synonymous	F118F	1/888,0.11%	Low		
ChrX	154534051	C	G	Missense	E282Q	1/888,0.11%	Moderate	Tolerated	Benign
ChrX	154532216	G	T	Missense	P507T	1/888,0.11%	Moderate	Tolerated	Benign
ChrX	154532390	G	A	Missense	R484C	1/888,0.11%	Moderate	Deleterious	Probably_damaging
ChrX	154546067	T	C	Missense	D60G	1/888,0.11%	Moderate	Tolerated	Benign
ChrX	154532676	C	A	Missense	R423L	1/888,0.11%	Moderate	Deleterious	Probably_damaging
ChrX	154534357	G	A	Stop Gained	Q239*	1/888,0.11%	High		
ChrX	154533112	C	T	Missense	C324Y	1/888,0.11%	Moderate	Deleterious	Possibly_damaging
ChrX	154533016	G	A	Missense	P356L	1/888,0.11%	Moderate	Deleterious	Possibly_damaging
ChrX	154532711	G	A	Synonymous	F411F	1/888,0.11%	Low		
ChrX	154533646	A	G	Missense	L295P	1/888,0.11%	Moderate	Deleterious	Probably_damaging
ChrX	154534397	C	A	Missense	Q225H	1/888,0.11%	Moderate	Deleterious	Probably_damaging
ChrX	154535335	G	T	Synonymous	S136S	1/888,0.11%	Low		
ChrX	154531447	C	T	3' UTR G6PD		1/888,0.11%	Modifier		
ChrX	154532705	C	A	Missense	Q413H	1/888,0.11%	Moderate	Tolerated	Benign
ChrX	154531481	G	A	3' UTR G6PD		1/888,0.11%	Modifier		
ChrX	154546869	A	G	Start Lost	M1?	1/888,0.11%	High	Deleterious	Benign
ChrX	154534443	G	A	Missense	S210F	1/888,0.11%	Moderate	Tolerated	Probably_damaging
ChrX	154535961	G	T	Synonymous	R111R	1/888,0.11%	Low		
ChrX	154532210	G	T	Missense	P509T	1/888,0.11%	Moderate	Deleterious	Probably_damaging
ChrX	154532456	T	C	Missense	K462E	1/888,0.11%	Moderate	Tolerated	Probably_damaging

Chr	Position	Ref	Alt	Consequence	Mutation	#_cases	Impact	SIFT	PolyPhen
ChrX	154532975	C	T	Missense	V370I	1/888,0.11%	Moderate	Tolerated	Benign
ChrX	154534351	G	A	Missense	L241F	1/888,0.11%	Moderate	Deleterious	Probably_damaging
ChrX	154535176	C	T	Missense	M189I	1/888,0.11%	Moderate	Tolerated	Benign
ChrX	154531704	G	A	3' UTR G6PD		1/888,0.11%	Modifier		
ChrX	154531588	G	A	3' UTR G6PD		1/888,0.11%	Modifier		
ChrX	154546159	G	A	Synonymous	S29S	1/888,0.11%	Low		
ChrX	154531542	G	A	3' UTR G6PD		1/888,0.11%	Modifier		
ChrX	154531703	G	T	3' UTR G6PD		1/888,0.11%	Modifier		
ChrX	154531732	C	A	3' UTR G6PD		1/888,0.11%	Modifier		
ChrX	154532981	insA		Frameshift	A368Cfs*10	1/888,0.11%	High		
ChrX	154532456	T	G	Missense	K462Q	1/888,0.11%	Moderate	Tolerated	Probably_damaging
ChrX	154546130	C	T	Missense	R39Q	1/888,0.11%	Moderate	Tolerated	Benign
ChrX	154532675	G	A	Synonymous	R423R	1/888,0.11%	Low		
ChrX	154535278	C	A	Missense	M155I	1/888,0.11%	Moderate	Tolerated	Benign
ChrX	154534462	C	G	Missense	G204R	1/888,0.11%	Moderate	Deleterious	Probably_damaging
ChrX	154531710	C	T	3' UTR G6PD		1/888,0.11%	Modifier		
ChrX	154534456	C	T	Missense	D206N	1/888,0.11%	Moderate	Deleterious	Probably_damaging
ChrX	154533094	G	A	Missense	A330V	1/888,0.11%	Moderate	Tolerated	Benign
ChrX	154532460	G	A	Splice Region	N460N	1/888,0.11%	Low		
ChrX	154532462	delTC		Splice Acceptor	X460_splice	1/888,0.11%	High		
ChrX	154546138	G	T	Synonymous	A36A	1/888,0.11%	Low		
ChrX	154535175	T	G	Missense	S190R	1/888,0.11%	Moderate	Deleterious	Probably_damaging
ChrX	154532057	C	T	Synonymous	K527K	1/888,0.11%	Low		
ChrX	154536017	C	T	Missense	E93K	1/888,0.11%	Moderate	Tolerated	Benign
ChrX	154532695	G	A	Missense	R417C	1/888,0.11%	Moderate	Deleterious	Probably_damaging
ChrX	154535343	G	A	Missense	R134C	1/888,0.11%	Moderate	Tolerated	Benign
ChrX	154533590	C	T	Missense	V314I	1/888,0.11%	Moderate	Tolerated	Benign

Chr	Position	Ref	Alt	Consequence	Mutation	#_cases	Impact	SIFT	PolyPhen
ChrX	154532996	T	C	Missense	T363A	1/888,0.11%	Moderate	Tolerated	Benign
ChrX	154532722	C	T	Missense	G408S	1/888,0.11%	Moderate	Tolerated	Probably_damaging
ChrX	154534414	G	T	Missense	L220M	1/888,0.11%	Moderate	Deleterious	Probably_damaging
ChrX	154532967	G	T	Synonymous	L372L	1/888,0.11%	Low		
ChrX	154532574	C	G	Missense	R457T	1/888,0.11%	Moderate	Deleterious	Probably_damaging
ChrX	154532021	C	T	Stop Gained	W539*	1/888,0.11%	High		
ChrX	154533646	A	T	Missense	L295Q	1/888,0.11%	Moderate	Deleterious	Probably_damaging
ChrX	154533003	G	A	Synonymous	R360R	1/888,0.11%	Low		
ChrX	154535378	delG		Frameshift	P122Qfs*34	1/888,0.11%	High		
ChrX	154533665	C	T	Missense	V289M	1/888,0.11%	Moderate	Deleterious	Probably_damaging
ChrX	154533041	C	T	Missense	A348T	1/888,0.11%	Moderate	Tolerated	Possibly_damaging
ChrX	154534080	C	T	Missense	G272D	1/888,0.11%	Moderate	Deleterious	Probably_damaging
ChrX	154534088	C	G	Missense	E269D	1/888,0.11%	Moderate	Deleterious	Probably_damaging
ChrX	154532606	C	A	Missense	E446D	1/888,0.11%	Moderate	Tolerated	Benign
ChrX	154546086	C	T	Missense	D54N	1/888,0.11%	Moderate	Tolerated	Benign
ChrX	154532639	C	T	Missense	M435I	1/888,0.11%	Moderate	Tolerated	Benign
ChrX	154532280	G	T	Missense	S485R	1/888,0.11%	Moderate	Tolerated	Possibly_damaging
ChrX	154532228	C	T	Missense	E503K	1/888,0.11%	Moderate	Deleterious	Benign
ChrX	154532202	G	C	Synonymous	P511P	1/888,0.11%	Low		
ChrX	154532206	A	G	Missense	I510T	1/888,0.11%	Moderate	Tolerated	Benign
ChrX	154546078	G	T	Missense	F56L	1/888,0.11%	Moderate	Tolerated	Benign
ChrX	154532792	G	A	Synonymous	F384F	1/888,0.11%	Low		
ChrX	154534114	C	T	Missense	A261T	1/888,0.11%	Moderate	Tolerated	Possibly_damaging
ChrX	154532434	C	T	Missense	R469H	1/888,0.11%	Moderate	Deleterious	Possibly_damaging
ChrX	154533056	C	T	Missense	D343N	1/888,0.11%	Moderate	Tolerated	Benign
ChrX	154532700	C	T	Missense	C415Y	1/888,0.11%	Moderate	Tolerated	Probably_damaging
ChrX	154534054	C	T	Missense	D281N	1/888,0.11%	Moderate	Deleterious	Benign

Chr	Position	Ref	Alt	Consequence	Mutation	#_cases	Impact	SIFT	PolyPhen
ChrX	154535207	G	A	Missense	A179V	1/888,0.11%	Moderate	Tolerated	Benign
ChrX	154532954	C	G	Missense	E377Q	1/888,0.11%	Moderate	Deleterious	Probably_damaging
ChrX	154535114	G	A	Intron		1/888,0.11%	Modifier		
ChrX	154535964	G	T	Synonymous	I110I	1/888,0.11%	Low		
ChrX	154532720	G	A	Synonymous	G408G	1/888,0.11%	Low		
ChrX	154533045	G	A	Synonymous	G346G	1/888,0.11%	Low		
ChrX	154533648	T	C	Synonymous	L294L	1/888,0.11%	Low		
ChrX	154533611	C	T	Missense	A307T	1/888,0.11%	Moderate	Tolerated	Benign
ChrX	154534142	G	A	Synonymous	F251F	1/888,0.11%	Low		
ChrX	154535350	G	A	Synonymous	F131F	1/888,0.11%	Low		
ChrX	154534419	G	A	Missense	S218F	1/888,0.11%	Moderate	Deleterious	Benign
ChrX	154535943	G	T	Missense	F117L	1/888,0.11%	Moderate	Deleterious	Benign
ChrX	154531719	G	A	3' UTR G6PD		1/888,0.11%	Modifier		
ChrX	154531448	G	A	3' UTR G6PD		1/888,0.11%	Modifier		
ChrX	154534384	C	T	Missense	D230N	1/888,0.11%	Moderate	Deleterious	Probably_damaging

These consequences have the following impacts: Vep impact (that determines the effect of the variants on genes, transcripts, protein sequence, and regulatory regions), SIFT (predicts whether an amino acid substitution affects protein function), and PolyPhen (predicts the possible impact of an amino acid substitution on the structure and function of a human protein). VEP impact displays 58% as moderate impact which is non-disruptive variant that might change protein effectiveness, 22.7% as low impact which is harmless, unlikely to change protein behavior, and 12.7% are modifier which are usually non-coding variants, and 6.67% have a high impact which is a disruptive impact in the protein, protein truncation or loss of function. Moreover, SIFT shows 41.3% of the mutations as unknown, 31.3% as deleterious, and 25.3% as tolerated. Whereas, PolyPhen also have 41.3% of the mutations as unknown, 27.3% are considered probably damaging, 25.3% are benign, and 6% of the mutations are considered possibly damaging.

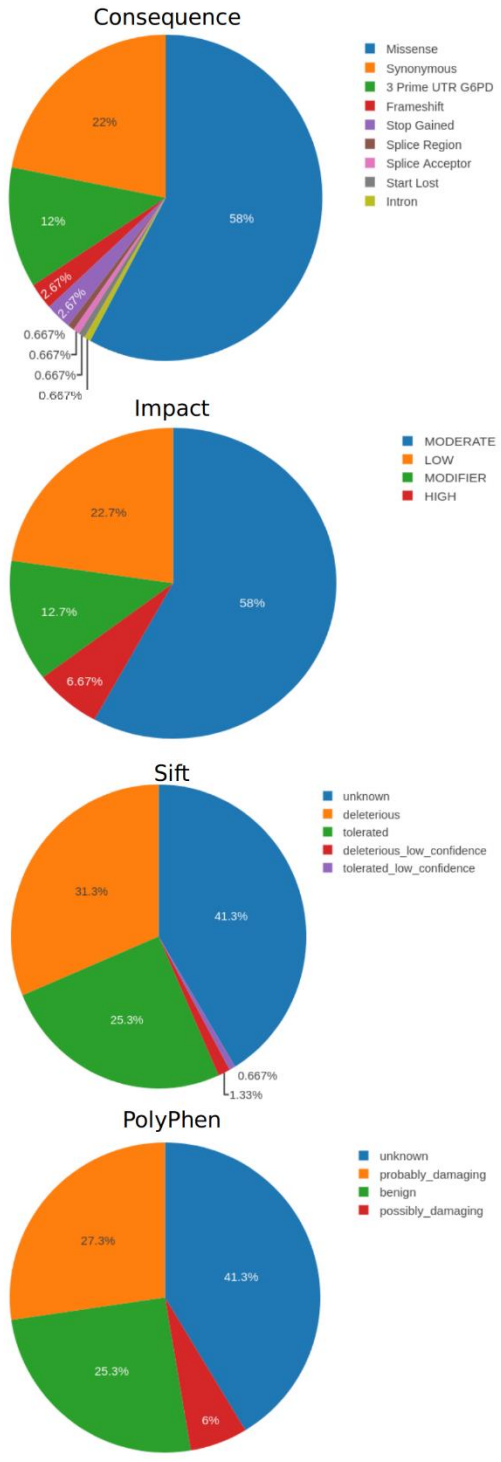

Figure 7: Consequence and the impact (VEP, SIFT, PolyPhen) of 150 G6PD mutations in cancer

These results show that most of the mutations are due to a mistake in the DNA that results in a wrong

amino acid assembly. However, most of them have a moderate or an unknown impact and further functional studies will be needed to characterize their function and impact in-vivo.

C. SNPs affecting G6PD and existing in cancer

After compiling the lists of G6PD mutations causing Favism/anemia (151 mutations; Table 2) and cancer (150 mutations; Table 3), we intersected both lists and identified only 20 common mutations considered as putative protective G6PD mutations against cancer (Figure 8 and Table 4). We, then, extended the table by adding from TCGA the mutation type, mutation consequence, number of cases affected, the mutations impact from the Ensemble database (VEP), the impact according to SIFT and PolyPhen, the project and case ID, gender, project name, disease type, and site.

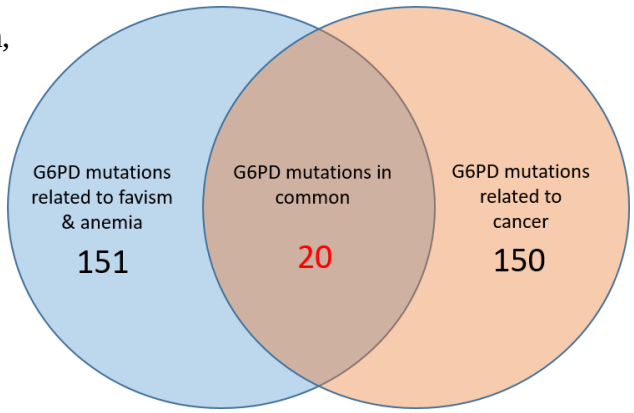

Figure 8: Result of 20 mutations after crosslinking

Table 4: Variant intersections of 20 common mutations between G6PD and cancer. Loc: location, Ex: exon. In: intron, Sub: substitution, Del: deletion, Dele: deleterious, Tole: tolerated, Prob: probably damaging, Possi: possibly damaging, Ben: benign, F: female, M: male

Chr	Position	dbSNP_ID	Type	Consequence	Cases	VEP	SIFT	PolyPhen	ID	Gender	TCGA Project Name
ChrX	154531719		Sub	3' UTR	1/888	Modifier			UCEC	F	Uterine Corpus Endometrial Carcinoma
ChrX	154531975	rs781866772	Del	3' UTR		Modifier			ESCA	M	Esophageal Carcinoma
ChrX	154532390	rs398123546	Sub	Missense	1/888	Moderate	Dele	Prob	CESC	F	Cervical Squamous Cell Carcinoma and Endocervical Adenocarcinoma
ChrX	154532434	rs137852337	Sub	Missense		Moderate	Dele	Possi	LUAD	F	Lung Adenocarcinoma
ChrX	154532662	rs137852325	Sub	Missense	1/888	Moderate	Dele	Prob	HNSC	M	Head and Neck Squamous Cell Carcinoma
ChrX	154532674	rs137852335	Sub	Missense		Moderate	Dele	Possi	COAD	F	Colon Adenocarcinoma
ChrX	154532676	rs137852316	Sub	Missense		Moderate	Dele	Prob	OV	F	Ovarian Serous Cystadenocarcinoma
ChrX	154532695	rs137852334	Sub	Missense	1/888	Moderate	Dele	Prob	GBM	F	Glioblastoma Multiforme
ChrX	154532752	rs387906468	Sub	Missense	1/888	Moderate	Dele	Prob	UCEC	F	Uterine Corpus Endometrial Carcinoma
ChrX	154532753		Sub	Synonymous	1/888	Low			UCEC	F	Uterine Corpus Endometrial Carcinoma
ChrX	154532772	rs137852345	Sub	Missense	1/888	Moderate	Dele	Prob	BRCA	F	Breast Invasive Carcinoma
ChrX	154534055		Sub	Synonymous	4/888	Low			GBM	F	Glioblastoma Multiforme
ChrX	154534055		Sub	Synonymous	4/888	Low			COAD	M	Colon Adenocarcinoma
ChrX	154534055		Sub	Synonymous	4/888	Low			UCEC	F	Uterine Corpus Endometrial Carcinoma
ChrX	154534055		Sub	Synonymous	4/888	Low			BRCA	F	Breast Invasive Carcinoma
ChrX	154534108		Sub	Missense	2/888	Moderate	Tole	Ben	UCEC	F	Uterine Corpus Endometrial Carcinoma
ChrX	154534408	rs1557230370	Sub	Missense	2/888	Moderate	Dele	Ben	SKCM	M	Skin Cutaneous Melanoma
ChrX	154534408	rs1557230370	Sub	Missense	2/888	Moderate	Dele	Ben	UCEC	F	Uterine Corpus Endometrial Carcinoma
ChrX	154534419	rs5030868	Sub	Missense	1/888	Moderate	Dele	Ben	LGG	F	Brain Lower Grade Glioma
ChrX	154534437	rs98123549	Sub	Missense		Moderate	Dele	Ben	LUSC	M	Lung Squamous Cell Carcinoma
ChrX	154534463	rs200111236	Sub	Synonymous	1/888	Low			HNSC	F	Head and Neck Squamous Cell Carcinoma
ChrX	154535176	rs370918918	Sub	Missense	1/888	Moderate	Tole	Ben	SKCM	F	Skin Cutaneous Melanoma
ChrX	154535278	rs782130334	Sub	Missense	1/888	Moderate	Tole	Ben	UCEC	F	Uterine Corpus Endometrial Carcinoma
ChrX	154535336	rs267606835	Sub	Missense		Moderate	Tole	Ben	SKCM	M	Skin Cutaneous Melanoma

D. Expression of mutations in different cancer

Last, we asked if the expression of G6PD in cancer patients with favism related G6PD mutations (Table 4) is different from other patients in the same cancer type. For this, we downloaded RNA expression data of G6PD in the 12 different cancer types from Table 4 and compared the expression of G6PD of the patient that show the mutation in G6PD to all remaining patients (Figure 9). The box plot shows the expression levels of G6PD mutations in all patients of each of the 12 cancer types. The red point on each bar shows the cancer patient that has one of the 20 G6PD mutations. The x-axis shows the cancer types that the 20 mutations are found in.

The plot indicates if patients with mutations in G6PD have higher or lower expression than other patients in the same cancer type. Four G6PD mutations in cervical squamous cell carcinoma (CESC), colon adenocarcinoma (COAD), head and neck squamous cell carcinoma (HNSC), and uterine corpus endometrial carcinoma (UCEC) show relatively high expression of G6PD in the patient with the mutation when compared to the median (horizontal line in the boxplot; Figure 9). A study showed that deficiency in G6PD may be correlated with the decreased ability of cancer cells to proliferate and migrate as a result of the abnormal reorganization of cell cytoskeleton and abnormal biomechanical properties caused by the increased ROS (Fang et al., 2016). According to the Human Protein Atlas database, high expression of G6PD in CESC and COAD is linked to a higher survival rate of the patients (Figure 10). Whereas, in HNSC and UCEC different patients show different expressions of the mutation. This may be due to other unknown factors such as age, gender, and most importantly the cancer stage and differentiation level. On the contrary, G6PD expression is less than the average in breast cancer (BRCA) and

lung adenocarcinoma (LUAD). Human protein atlas shows that low expression of G6PD is correlated with a higher survival rate in BRCA. But this cannot be considered true for

LUAD. However, a study on G6PD expression showed that it may be a novel predictive prognostic marker for lung adenocarcinoma (Nagashio et al., 2019a).

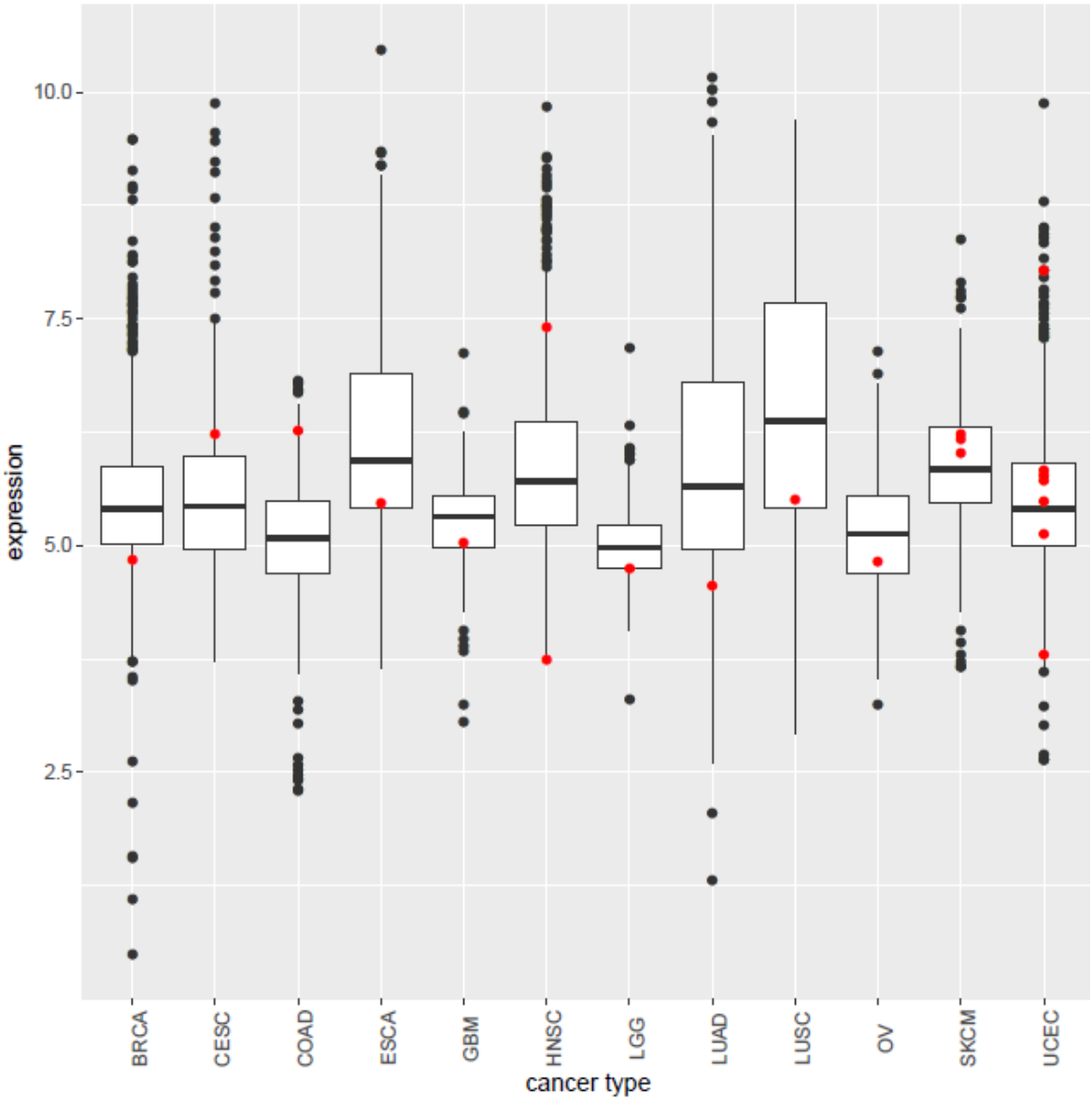

Figure 9: Box plot that shows the expression of G6PD in cancer. BRCA: breast cancer, CESC: cervical squamous cell carcinoma, COAD: colon adenocarcinoma, ESCA: esophageal carcinoma, GBM: glioblastoma multiform, HNSC: head and neck squamous cell carcinoma, LGG: low-grade glioma, LUAD: lung adenocarcinoma, LUSC: lung squamous cell carcinoma, OV: ovarian cancer, SKCM: skin cutaneous melanoma, UCEC: uterine corpus endometrial carcinoma.

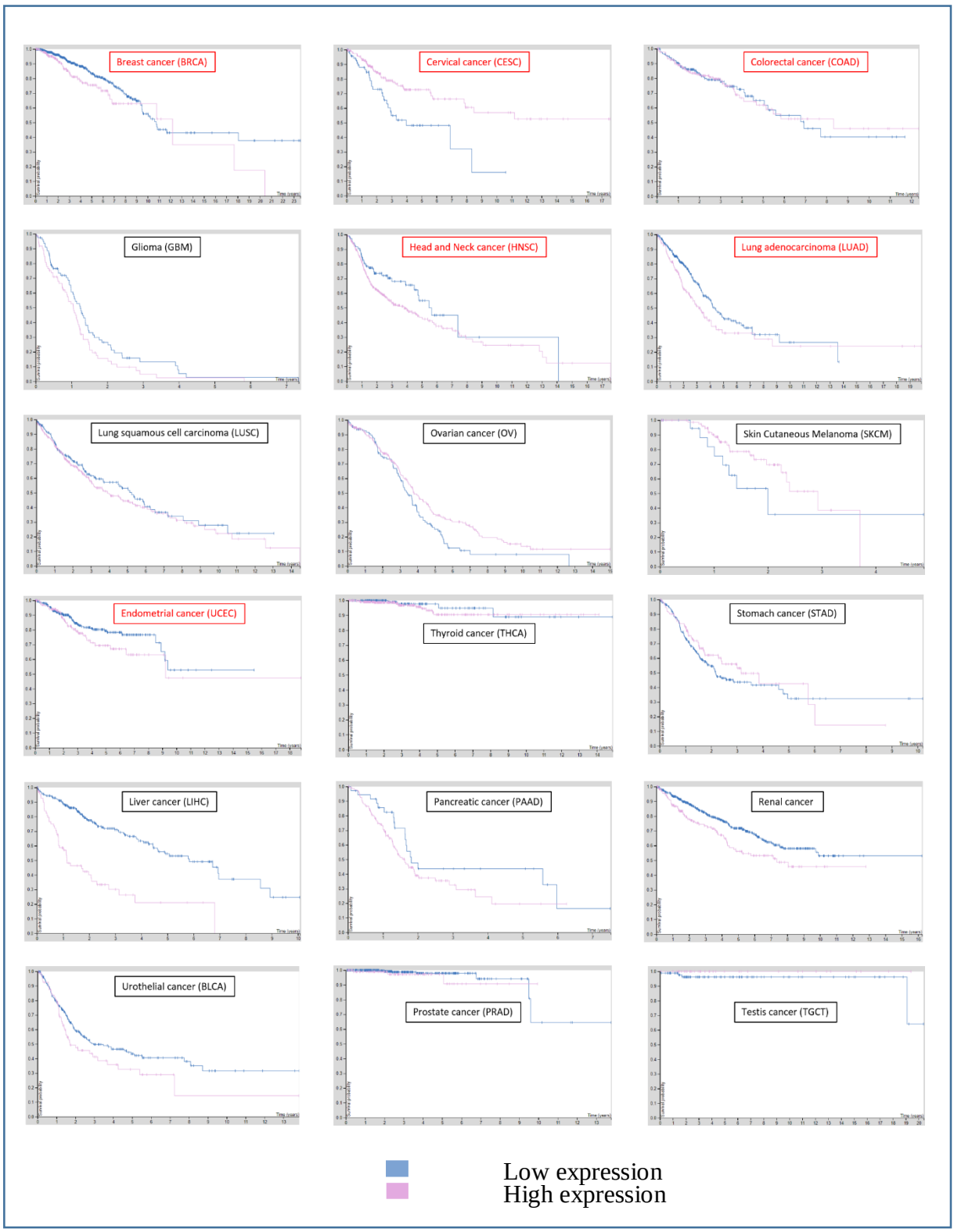

Figure 10: Survival rate of cancer patients with high and low G6PD expression

CHAPTER VI

DISCUSSION AND CONCLUSION

In human genetics, G6PD deficiency is a sample of conditional mutant, whereas hemolytic anemia is expressed after exposure to exogenous agents such as drugsand fava beans. G6PD is essential in growth and development due to its major role in the production of NADPH for ROS elimination and the synthesis of ribose for nucleotide synthesis. Consequently, targeting G6PD in the PPP might be a promising way to hinder tumor growth due to its dependence on NADPH and ribose from this pathway. As mentioned earlier, increased G6PD activity or PPP flux has been found in cancer cell lines and is linked to key oncogenic mutations in human cancers.

From this perspective, a high PPP flux could be considered a marker of tumor aggressiveness and bad prognosis. Studies have shown for several cancers, including cervical squamous cell carcinoma and colon adenocarcinoma, a correlation of diagnosis and prognosis with the overexpression of G6PD.

The further investigation remains essential to have a clear view on how PPP is regulated and what advantages does it have for cancer cells if it was altered. It is important to perform functional assays on the 20 identified mutations to prove if they cause a decrease in G6PD expression and if this decrease is protective against cancer. Moreover, discover additional strategies and mechanisms that science can take advantage of to have therapeutic value to fight against cancer.

The key enzyme of the PPP that provides NADPH and ribose for nucleic acid synthesis for cancer cells is the G6PD, so, in theory, if G6PD was deficient, cancer growth

and proliferation will be hindered. Many studies and research are still needed to prove if G6PD deficiency is protective against cancer.

CHAPTER VII

FUTURE PERSPECTIVE

It has been now almost 65 years since the discovery of G6PD deficiency and its cause of hemolytic anemia and sensitivity to primaquine. Thousands of papers documented its clinical significance, molecular biology, population distribution, and biochemical characteristics have been published. The natural occurrence of many mutations in the G6PD gene opened for further understanding of structure function relationships of enzymes. Moreover, research have widened up the importance of G6PD and its relationship with cancer. Many questions remain unanswered and need further investigations such as if it is possible to inhibit the G6PD at the organismal level to selectively inhibit tumorigenesis without significant adverse physiological consequences? Is the depletion of NADPH by the inhibition of the oxidative PPP through inhibiting G6PD elicits compensatory alternative mechanisms that generate NADPH? Is inhibition of the oxidative PPP through inhibiting G6PD could contribute to selective eradication of cancer cells due to ROS build up? Further investigation remains essential to uncover the role of G6PD mutations with cancer cells.

REFERENCES

(1967). Standardization of procedures for the study of glucose-6-phosphate dehydrogenase. Report of a WHO Scientific Group. World Health Organ Tech Rep Ser 366, 1-53.

Adam, A. (1961). Linkage between deficiency of glucose-6-phosphate dehydrogenase and colour-blindness. Nature 189, 686-686.

Albano, E., Tomasi, A., Mannuzzu, L., and Arese, P. (1984). Detection of a free radical intermediate from divicine of Vicia faba. Biochemical pharmacology 33, 1701-1704.

Allison, A. (1960). Glucose-6-phosphate dehydrogenase deficiency in red blood cells of East Africans. Nature 186, 531-532.

Arese, P., Mannuzzu, L., Turrini, F., Galiano, S., and Gaetani, G.F. (1986). ETIOLOGICAL ASPECTS OF FAVISM11Supported by CNR Grants PF 84.00864.51 to P.A. and PF 60/84. 0087.51 to G.F., and by the Thalassemia Research institute (CNR Cagliari). In Glucose-6-phosphate Dehydrogenase, A. Yoshida, and E. Beutler, eds. (Academic Press), pp. 45-75.

Arie, T. (1961). Pythagoras and the Beans. British Medical Journal 2, 709.

Au, S.W., Gover, S., Lam, V.M., and Adams, M.J. (2000a). Human glucose-6-phosphate dehydrogenase: the crystal structure reveals a structural NADP(+) molecule and provides insights into enzyme deficiency. Structure (London) 8, 293.

Au, S.W., Gover, S., Lam, V.M., and Adams, M.J. (2000b). Human glucose-6-phosphate dehydrogenase: the crystal structure reveals a structural NADP+ molecule and provides insights into enzyme deficiency. Structure 8, 293-303.

Batetta, B., Pulisci, D., Bonatesta, R.R., Sanna, F., Piras, S., Mulas, M.F., Spano, O., Putzolu, M., Broccia, G., and Dessì, S. (1999). G6PD activity and gene expression in leukemic cells from G6PD-deficient subjects. Cancer Letters 140, 53-58.

Benito, A., Polat, I.H., Noé, V., Ciudad, C.J., Marin, S., and Cascante, M. (2017). Glucose-6-phosphate dehydrogenase and transketolase modulate breast cancer cell metabolic reprogramming and correlate with poor patient outcome. Oncotarget 8, 106693-106706.

Bensaad, K., Tsuruta, A., Selak, M.A., Vidal, M.N.C., Nakano, K., Bartrons, R., Gottlieb, E., and Vousden, K.H. (2006). TIGAR, a p53-inducible regulator of glycolysis and apoptosis. Cell 126, 107-120.

Beutler, E. (2008). Glucose-6-phosphate dehydrogenase deficiency: a historical perspective. Blood 111, 16-24.

Bonneau, D., and Longy, M. (2000). Mutations of the human PTEN gene. Human mutation 16, 109-122.

Boyer, S.H., Porter, I.H., and Weilbacher, R.G. (1962). Electrophoretic heterogeneity of glucose-6-phosphate dehydrogen ase and its relationship to enzyme deficiency in man. Proceedings of the National Academy of Sciences of the United States of America 48, 1868.

Browne, E. (1957). The inheritance of an intrinsic abnormality of the red blood cell predisposing to drug induced hemolytic anemia. Paper presented at: Bulletin of the Johns Hopkins Hospital (JOHNS HOPKINS UNIV PRESS JOURNALS PUBLISHING DIVISION, 2715 NORTH CHARLES ST …).

Cappellini, M.D., and Fiorelli, G. (2008). Glucose-6-phosphate dehydrogenase deficiency. The lancet 371, 64-74.

Carson, P.E., Flanagan, C.L., Ickes, C.E., and Alving, A.S. (1956). Enzymatic Deficiency in Primaquine-Sensitive Erythrocytes. Science 124, 484-485.

Chen, J.-Q., and Russo, J. (2012). Dysregulation of glucose transport, glycolysis, TCA cycle and glutaminolysis by oncogenes and tumor suppressors in cancer cells. Biochimica et biophysica acta 1826, 370.

Chen, X., Xu, Z., Zhu, Z., Chen, A., Fu, G., Wang, Y., and Pan, H. (2018). Modulation of G6PD affects bladder cancer via ROS accumulation and the AKT pathway in vitro. International Journal of Oncology 53, 1703.

Chevion, M., Navok, T., Glaser, G., and Mager, J. (1982). The chemistry of favism-inducing compounds: the properties of isouramil and divicine and their reaction with glutathione. European Journal of Biochemistry 127, 405-409.

Cordes, W. (1926). Experiences with plasmochin in malaria. United Fruit Co. Med Dept) 15th Annual Report, Boston, 66-71.

Coy, J.F., Dressler, D., Wilde, J., and Schubert, P. (2005). Mutations in the transketolase-like gene TKTL1: clinical implications for neurodegenerative diseases, diabetes and cancer. Clinical laboratory (Heidelberg) 51, 257.

Dang, C.V., Koppenol, W.H., and Bounds, P.L. (2011). Otto Warburg's contributions to current concepts of cancer metabolism. Nature Reviews Cancer 11, 618-618.

DeNicola, G.M., Karreth, F.A., Humpton, T.J., Gopinathan, A., Wei, C., Frese, K., Mangal, D., Kenneth, H.Y., Yeo, C.J., and Calhoun, E.S. (2011). Oncogene-induced Nrf2 transcription promotes ROS detoxification and tumorigenesis. Nature 475, 106-109.

Dong, T., Kang, X., Liu, Z., Zhao, S., Ma, W., Xuan, Q., Liu, H., Wang, Z., and Zhang, Q. (2016). Altered glycometabolism affects both clinical features and prognosis of triple-negative and neoadjuvant chemotherapy-treated breast cancer. Tumor Biology 37, 8159-8168.

Dore, M.P., Vidili, G., Marras, G., Assy, S., and Pes, G.M. (2018). Inverse Association between Glucose6Phosphate Dehydrogenase Deficiency and Hepatocellular Carcinoma. Asian Pac J Cancer Prev 19, 1069-1073.

Düvel, K., Yecies, J.L., Menon, S., Raman, P., Lipovsky, A.I., Souza, A.L., Triantafellow, E., Ma, Q., Gorski, R., and Cleaver, S. (2010). Activation of a metabolic gene regulatory network downstream of mTOR complex 1. Molecular cell 39, 171-183.

Fang, Z., Jiang, C., Feng, Y., Chen, R., Lin, X., Zhang, Z., Han, L., Chen, X., Li, H., Guo, Y., et al. (2016). Effects of G6PD activity inhibition on the viability, ROS generation and mechanical properties of cervical cancer cells. Biochim Biophys Acta 1863, 2245-2254.

Faubert, B., Boily, G., Izreig, S., Griss, T., Samborska, B., Dong, Z., Dupuy, F., Chambers, C., Fuerth, B.J., and Viollet, B. (2013). AMPK is a negative regulator of the Warburg effect and suppresses tumor growth in vivo. Cell metabolism 17, 113-124.

Gaetani, G.F., Galiano, S., Canepa, L., Ferraris, A.M., and Kirkman, H.N. (1989). Catalase and glutathione peroxidase are equally active in detoxification of hydrogen peroxide in human erythrocytes.

Garcia-Cao, I., Song, M.S., Hobbs, R.M., Laurent, G., Giorgi, C., De Boer, V.C., Anastasiou, D., Ito, K., Sasaki, A.T., and Rameh, L. (2012). Systemic elevation of PTEN induces a tumor-suppressive metabolic state. Cell 149, 49-62.

Gupta, G.P., and Massagué, J. (2006). Cancer metastasis: building a framework. Cell 127, 679-695.

Hanahan, D., and Weinberg, R.A. (2000). The hallmarks of cancer. cell 100, 57-70.

Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. cell 144, 646-674.

Hitosugi, T., Zhou, L., Elf, S., Fan, J., Kang, H.-B., Seo, J.H., Shan, C., Dai, Q., Zhang, L., and Xie, J. (2012). Phosphoglycerate mutase 1 coordinates glycolysis and biosynthesis to promote tumor growth. Cancer cell 22, 585-600.

Hong, X., Song, R., Song, H., Zheng, T., Wang, J., Liang, Y., Qi, S., Lu, Z., Song, X., and Jiang, H. (2014a). PTEN antagonises Tcl1/hnRNPK-mediated G6PD pre-mRNA splicing which contributes to hepatocarcinogenesis. Gut 63, 1635-1647.

Hong, X., Song, R., Song, H., Zheng, T., Wang, J., Liang, Y., Qi, S., Lu, Z., Song, X., Jiang, H., et al. (2014b). PTEN antagonises Tcl1/hnRNPK-mediated G6PD pre-mRNA splicing which contributes to hepatocarcinogenesis. Gut 63, 1635-1647.

Isselbacher, K.J., Anderson, E.P., Kurahashi, K., and Kalckar, H.M. (1956). Congenital Galactosemia, a Single Enzymatic Block in Galactose Metabolism. Science 123, 635-636.

Jiang, P., Du, W., Wang, X., Mancuso, A., Gao, X., Wu, M., and Yang, X. (2011). p53 regulates biosynthesis through direct inactivation of glucose-6-phosphate dehydrogenase. Nature cell biology 13, 310-316.

Jiang, P., Du, W., and Wu, M. (2014). Regulation of the pentose phosphate pathway in cancer. Protein & cell 5, 592-602.

Kattamis, C. (1986). FAVISM: EPIDEMIOLOGICAL AND CLINICAL ASPECTS. In Glucose-6-phosphate Dehydrogenase, A. Yoshida, and E. Beutler, eds. (Academic Press), pp. 25-43.

Kirkman, H.N., Riley, H.D., and Crowell, B.B. (1960). DIFFERENT ENZYMIC EXPRESSIONS OF MUTANTS OF HUMAN GLUCOSE-6-PHOSPHATE DEHYDROGENASE. Proc Natl Acad Sci U S A 46, 938-944.

Kohan, A.B., Talukdar, I., Walsh, C.M., and Salati, L.M. (2009). A role for AMPK in the inhibition of glucose-6-phosphate dehydrogenase by polyunsaturated fatty acids. Biochemical and biophysical research communications 388, 117-121.

Kondoh, H., Lleonart, M.E., Gil, J., Wang, J., Degan, P., Peters, G., Martinez, D., Carnero, A., and Beach, D. (2005). Glycolytic enzymes can modulate cellular life span. Cancer research 65, 177-185.

Langbein, S., Frederiks, W.M., zur Hansen, A., Popa, J., Lehmann, J., Weiss, C., Alken, P., and Coy, J.F. (2008). Metastasis is promoted by a bioenergetic switch: New targets for progressive renal cell cancer. International journal of cancer Journal international du cancer 122, 2422-2428.

Liu, H., Huang, D., McArthur, D.L., Boros, L.G., Nissen, N., and Heaney, A.P. (2010). Fructose induces transketolase flux to promote pancreatic cancer growth. Cancer research (Chicago, Ill) 70, 6368-6376.

Lu, M., Lu, L., Dong, Q., Yu, G., Chen, J., Qin, L., Wang, L., Zhu, W., and Jia, H. (2018). Elevated G6PD expression contributes to migration and invasion of hepatocellular carcinoma cells by inducing epithelial-mesenchymal transition. Acta Biochimica et Biophysica Sinica 50, 370-380.

Luzzatto, L. (1973). New developments in glucose-6-phosphate dehydrogenase deficiency. Isr J Med Sci 9, 1484-1498.

Luzzatto, L. (2001). Glucose 6-phosphate dehydrogenase deficiency. The metabolic and molecular bases of inherited disease, 4517-4553.

Luzzatto, L., and Arese, P. (2018). Favism and Glucose-6-Phosphate Dehydrogenase Deficiency. New England Journal of Medicine 378, 60-71.

Marks, P.A., and Gross, R.T. (1959). Erythrocyte glucose-6-phosphate dehydrogenase deficiency: evidence of differences between Negroes and Caucasians with respect to this genetically determined trait. J Clin Invest 38, 2253-2262.

Mason, P.J., Bautista, J.M., and Gilsanz, F. (2007). G6PD deficiency: the genotype-phenotype association. Blood reviews 21, 267-283.

McMillan, D.C., Bolchoz, L.J., and Jollow, D.J. (2001). Favism: effect of divicine on rat erythrocyte sulfhydryl status, hexose monophosphate shunt activity, morphology, and membrane skeletal proteins. Toxicological Sciences 62, 353-359.

Meletis, J., and Konstantopoulos, K. (2004). Favism-from the “avoid fava beans” of Pythagoras to the present. Haema 7, 17-21.

Mitsuishi, Y., Motohashi, H., and Yamamoto, M. (2012a). The Keap1–Nrf2 system in cancers: stress response and anabolic metabolism. Frontiers in oncology 2, 200.

Mitsuishi, Y., Taguchi, K., Kawatani, Y., Shibata, T., Nukiwa, T., Aburatani, H., Yamamoto, M., and Motohashi, H. (2012b). Nrf2 redirects glucose and glutamine into anabolic pathways in metabolic reprogramming. Cancer cell 22, 66-79.

Nagashio, R., Oikawa, S., Yanagita, K., Hagiuda, D., Kuchitsu, Y., Igawa, S., Naoki, K., Satoh, Y., Ichinoe, M., and Murakumo, Y. (2019a). Prognostic significance of G6PD expression and localization in lung adenocarcinoma. Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics 1867, 38-46.

Nagashio, R., Oikawa, S., Yanagita, K., Hagiuda, D., Kuchitsu, Y., Igawa, S., Naoki, K., Satoh, Y., Ichinoe, M., Murakumo, Y., et al. (2019b). Prognostic significance of G6PD expression and localization in lung adenocarcinoma. Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics 1867, 38-46.

Nagashio, R., Oikawa, S., Yanagita, K., Hagiuda, D., Kuchitsu, Y., Igawa, S., ... & Saegusa, M. (2019). Prognostic significance of G6PD expression and localization in lung adenocarcinoma. Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics 1, 38-46.

Naylor, C., Rowland, P., Basak, A., Gover, S., Mason, P., Bautista, J., Vulliamy, T., Luzzatto, L., and Adams, M. (1996). Glucose 6-phosphate dehydrogenase mutations causing enzyme deficiency in a model of the tertiary structure of the human enzyme.

Necheles, T.F., Weinstein, I.M., and Leroy, G.V. (1953). Radioactive sodium chromate for the study of survival of red blood cells. I. The effect of radioactive sodium chromate on red cells. J Lab Clin Med 42, 358-367.

Pan, S., World, C.J., Kovacs, C.J., and Berk, B.C. (2009). Glucose 6-phosphate dehydrogenase is regulated through c-Src–mediated tyrosine phosphorylation in endothelial cells. Arteriosclerosis, thrombosis, and vascular biology 29, 895-901.

Pandolfi, P.P., Sonati, F., Rivi, R., Mason, P., Grosveld, F., and Luzzatto, L. (1995). Targeted disruption of the housekeeping gene encoding glucose 6-phosphate dehydrogenase (G6PD-null): G6PD is dispensable for pentose synthesis but essential for defense against oxidative stress. EMBO Journal 14, 5209-5215.

Patra, K.C., and Hay, N. (2014). The pentose phosphate pathway and cancer. Trends in biochemical sciences 39, 347-354.

Patra, K.C., Wang, Q., Bhaskar, P.T., Miller, L., Wang, Z., Wheaton, W., Chandel, N., Laakso, M., Muller, W.J., and Allen, E.L. (2013). Hexokinase 2 is required for tumor initiation and maintenance and its systemic deletion is therapeutic in mouse models of cancer. Cancer cell 24, 213-228.

Pes, G.M., Errigo, A., Soro, S., Longo, N.P., and Dore, M.P. (2019). Glucose-6-phosphate dehydrogenase deficiency reduces susceptibility to cancer of endodermal origin. Acta Oncologica 58, 1205-1211.

Pu, H., Zhang, Q., Zhao, C., Shi, L., Wang, Y., Wang, J., and Zhang, M. (2015). Overexpression of G6PD is associated with high risks of recurrent metastasis and poor progression-free survival in primary breast carcinoma. World journal of surgical oncology 13, 323.

Raïs, B., Comin, B., Puigjaner, J., Brandes, J.L., Creppy, E., Saboureau, D., Ennamany, R., Lee, W.-N.P., Boros, L.G., and Cascante, M. (1999). Oxythiamine and dehydroepiandrosterone induce a G1 phase cycle arrest in Ehrlich's tumor cells through inhibition of the pentose cycle. FEBS letters 456, 113-118.

Riganti, C., Gazzano, E., Polimeni, M., Aldieri, E., and Ghigo, D. (2012). The pentose phosphate pathway: an antioxidant defense and a crossroad in tumor cell fate. Free Radical Biology and Medicine 53, 421-436.

Schwartzenberg-Bar-Yoseph, F., Armoni, M., and Karnieli, E. (2004). The tumor suppressor p53 down-regulates glucose transporters GLUT1 and GLUT4 gene expression. Cancer research 64, 2627-2633.

Simoons, F.J. (1998). Plants of life, plants of death (Univ of Wisconsin Press).

Simpson, L., and Parsons, R. (2001). PTEN: life as a tumor suppressor. Experimental cell research 264, 29-41.

Stanton, R., Seifter, J., Boxer, D.C., Zimmerman, E., and Cantley, L. (1991). Rapid release of bound glucose-6-phosphate dehydrogenase by growth factors. Correlation with increased enzymatic activity. Journal of Biological Chemistry 266, 12442-12448.

Stanton, R.C. (2012). Glucose-6-phosphate dehydrogenase, NADPH, and cell survival. IUBMB life 64, 362-369.

Sterling, K., and Gray, S.J. (1950). Determination of the circulating red cell volume in man by radioactive chromium. J Clin Invest 29, 1614-1619.

Sukhatme, V.P., and Chan, B. (2012). Glycolytic cancer cells lacking 6-phosphogluconate dehydrogenase metabolize glucose to induce senescence. FEBS Letters 586, 2389-2395.

Szabo, P., Purrello, M., Rocchi, M., Archidiacono, N., Alhadeff, B., Filippi, G., Toniolo, D., Martini, G., Luzzatto, L., and Siniscalco, M. (1984). Cytological mapping of

the human glucose-6-phosphate dehydrogenase gene distal to the fragile-X site suggests a high rate of meiotic recombination across this site. Proceedings of the National Academy of Sciences 81, 7855-7859.

Tian, W.-N., Pignatare, J.N., and Stanton, R.C. (1994). Signal transduction proteins that associate with the platelet-derived growth factor (PDGF) receptor mediate the PDGF-induced release of glucose-6-phosphate dehydrogenase from permeabilized cells. Journal of Biological Chemistry 269, 14798-14805.

Trask, B.J., Massa, H., Kenwrick, S., and Gitschier, J. (1991). Mapping of human chromosome Xq28 by two-color fluorescence in situ hybridization of DNA sequences to interphase cell nuclei. American journal of human genetics 48, 1.

Wang, J., Wang, J., Yuan, W., Yuan, W., Chen, Z., Chen, Z., Wu, S., Wu, S., Chen, J., Chen, J., et al. (2012). Overexpression of G6PD is associated with poor clinical outcome in gastric cancer. Tumor Biology 33, 95-101.

Wind, F., and Negelein, E. (1927). The metabolism of tumors in the body. J Gen Physiol 8, 519-530.

Winterbourn, C.C., Benatti, U., and De Flora, A. (1986). Contributions of superoxide, hydrogen peroxide, and transition metal ions to auto-oxidation of the favism-inducing pyrimidine aglycone, divicine, and its reactions with haemoglobin. Biochemical pharmacology 35, 2009-2015.

Xu, Y., Osborne, B.W., and Stanton, R.C. (2005). Diabetes causes inhibition of glucose-6-phosphate dehydrogenase via activation of PKA, which contributes to oxidative stress in rat kidney cortex. American Journal of Physiology-Renal Physiology 289, F1040-F1047.

Ying, H., Kimmelman, A.C., Lyssiotis, C.A., Hua, S., Chu, G.C., Fletcher-Sananikone, E., Locasale, J.W., Son, J., Zhang, H., and Coloff, J.L. (2012). Oncogenic Kras maintains pancreatic tumors through regulation of anabolic glucose metabolism. Cell 149, 656-670.

Yoshida, A., and Beutler, E. (1983). G-6-PD variants: another up-date. Annals of Human Genetics 47, 25-38.

Zhang, Z., Apse, K., Pang, J., and Stanton, R.C. (2000). High glucose inhibits glucose-6-phosphate dehydrogenase via cAMP in aortic endothelial cells. Journal of Biological Chemistry 275, 40042-40047.