Title:

Biochemical and *In silico* Approaches to Study Glucose-6-Phosphate Dehydrogenase (G6PDH) in Potato (*Solanum tuberosum* L.)

A
DISSERTATION
SUBMITTED
FOR
PARTIAL FULFILMENT FOR THE AWARD OF DEGREE
MASTER OF SCIENCE
in
BIOTECHNOLOGY
BY

Eakdeep

(Regd. No. 302101008)

THAPAR INSTITUTE
OF ENGINEERING & TECHNOLOGY
(Deemed to be University)

Under the Supervision of
Dr. Niranjan Das
Professor

DEPARTMENT OF BIOTECHNOLOGY
Thapar Institute of Engineering and Technology, Patiala
July 2023

CANDIDATE'S DECLARATION

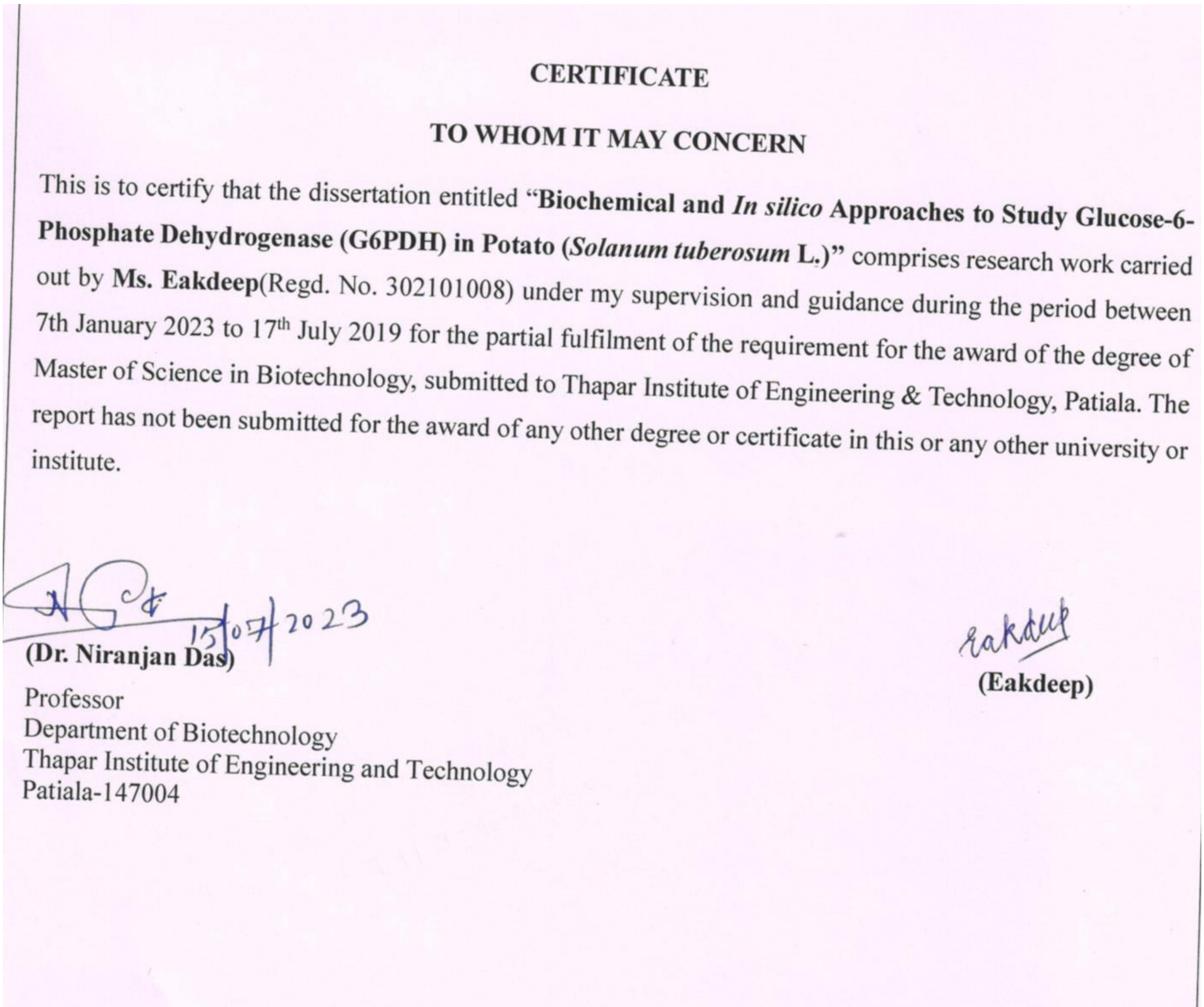
I hereby certify that the research presented in the thesis **"Biochemical and *In silico* Approaches to Study Glucose-6-Phosphate Dehydrogenase (G6PDH) in Potato (*Solanum tuberosum* L.)"** submitted in partial fulfilment of the requirements for the award of the degree of Master of Science in Biotechnology from Thapar Institute of Engineering & Technology, Patiala, is an authentic account of the work I conducted under the guidance and supervision of Dr. Niranjan Das The dissertation's material has never been given to another university or institute in order to grant a different degree.

Eakdeep

(Eakdeep)

Date: 15th July, 2023

Place: TIET, Patiala



CERTIFICATE

TO WHOM IT MAY CONCERN

This is to certify that the dissertation entitled "**Biochemical and *In silico* Approaches to Study Glucose-6-Phosphate Dehydrogenase (G6PDH) in Potato (*Solanum tuberosum* L.)**" comprises research work carried out by **Ms. Eakdeep**(Regd. No. 302101008) under my supervision and guidance during the period between 7th January 2023 to 17th July 2019 for the partial fulfilment of the requirement for the award of the degree of Master of Science in Biotechnology, submitted to Thapar Institute of Engineering & Technology, Patiala. The report has not been submitted for the award of any other degree or certificate in this or any other university or institute.

15/07/2023

(Dr. Niranjan Das)
Professor
Department of Biotechnology
Thapar Institute of Engineering and Technology
Patiala-147004

Eakdeep
(Eakdeep)

ACKNOWLEDGEMENT

I express an abundance of gratitude to every person who has assisted me in making this effort a success. First and foremost, I want to thank God for providing me the courage and support I needed to accomplish the job without incident.

I am honoured to thank my mentor **Dr. N. Das**, Professor of the Department of Biotechnology at the Thapar Institute of Engineering & Technology in Patiala, Punjab, for his advice, courtesy, inspiration, and excellent oversight during my work. I want to thank him from the bottom of my heart for his kind assistance and wise counsel. I was able to learn a lot from him throughout this time because of his continual support, helpful criticism, and capacity for dealing with challenges.

I want to give **Dr. M.S. Reddy**, Head of the Department of Biotechnology at the Thapar Institute of Engineering & Technology in Patiala, my deepest gratitude for allowing me the chance to work on this project.

Thank you to all the professors for their assistance and helpful advice during the project. I owe my gratitude to **Ms. Yadveer Kaur** consistent efforts and support, I would like to thank **Ms. Shreya Sharma** for her support and motivation. I owe her a debt of appreciation for helping me finish the project. I want to thank my lab buddy **Chestha Mittal** for all her support with the project. I appreciate the assistance provided by the lab attendants and non-teaching personnel.

In particular, I deserve my gratitude to **Mr. Ram Newal Yadav, Mr. Lallan Yadav, Mr. Mohinder Singh and Mr. Surinder pal Kumar.**

I especially want to express my heartfelt gratitude to my parents, **Mr. Pardeep Singh** and **Mrs. Iqbal Kaur**, for their love and faith. Additionally, I want to thank my sister **Navdeep Kaur** for her inspiration.

I thank my close friend **Amarjeet Singh** for supporting me through difficult moments. The entire credit belongs to everyone whose unwavering trust in me has always motivated me.

Eakdeep

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

Abbreviations	**Name**
6PG	6-phosphogluconate
Aa	Amino acid
ABA	Abscisic acid
AGPase	ADP-Glucose pyrophosphorylase
BLASTn	Basic alignment search tool for nucleotide
BLASTp	Basic alignment search tool for protein
Bp	Base pair
BSA	Bovine serum albumin
cDNA	Complementary DNA
CDS	Coding sequence
ChG6PDH	Chloroplastic Glucose-6-Phosphate Dehydrogenase
CK	Cytokinin
CPRI	Central research potato institution
CyG6PDH	Cytoplasmic Glucose-6-Phosphate Dehydrogenase
DE	Desiree
DNA	Deoxyribose nucleic acid
EDTA	Ethylene diamene tetraacetic acid
EDTA	Ethylenediaminetetraacetic acid
G6P	Glucose-6-phosphate
G6P	Glucose-6-phosphate
G6PDH	Glucose-6-Phosphate Dehydrogenase
GA	Gibberllic acid
GPX	Glutathione peroxidases
GRAVY	Grand Average of Hydropathicity
GSH	Glutathione
H_2O	Water
HSP	Heat Shock Proteins
I TASSER	Iterative threading assembly refinement
KC1	Kufri Chipsona-1
KC2	Kufri Chipsona-2
kDa	Kilo Dalton
KJ	Kufri Jyoti

LOX	Lipoxygenase
$MgCl_2$	Magnesium chloride
mM	Millimolar
MS	Murashiage and Skoog Medium
MSA	Multiple sequence alignment
MW	Molecular weight
N_2	Nitrogen
NAD	Nicotinamide adenine dinucleotide
NADPH	nicotinamide adenine dinucleotide phosphate
NCBI	National center for biotechnology information
OD	Optical density
OPPP	Oxidative Pentose Phosphate Pathway
Pi	Isoelectric point
PMSF	Phenylmethylsulfonyl flouride
POTH1	Potato MADS-box 1
POTM1	Potato Homobox 1
PPP	Pentose Phosphate Pathway
PR	Kufri Pukhraj
RNA	Ribose Nucleic Acid
ROS	Reactive Oxygen Species
StBEL	*Solanum tubersoum*
TALE	Three amino acid loop extension
TF	Transcription factors
TPS	True potato seeds
uL	Micro litre
UV	Ultra violet
ZWF1	Zwischenferment

ABSTRACT

The potato, a *Solanaceae* member, is a popular non-grain food crop. India ranks third worldwide in terms of potato consumption. It is a short-day crop. Tuberization is a crucial survival strategy for potato plants. It is a sensitive phase that could be easily affected by extreme environmental conditions. G6PDH is an important enzyme interconnecting glycolysis and pentose phosphate pathway (PPP). This enzyme also plays an important role in stress mitigation. In this study, we adopted both biochemical and *in silico* approaches to characterize G6PDH enzyme. A total of 4 genes encoding 14 isoforms were identified. This family contains both stable and unstable hydrophilic proteins. Protein-protein interactions revealed association with various other carbohydrate metabolism enzymes. *In silico* expression pattern analysis in the form of Heatmap revealed differential expression in various potato organs. Multiple sequence alignment revealed significant divergence among G6PDH of various plant species of *Solanaceae* members. Phylogenetic analysis revealed evolutionary relationship between the G6PDH belonging to both *Solanaceae* and non-*Solanaceae* members. 3-D structures were also predicted. At biochemical level, enzyme assays of five Indian potato cultivars were performed. Potato organs at various stages of development showed variation in G6PDH enzyme activities. Mature leaves and stems showed maximum enzyme activity. Among tuberization stages, maximum G6PDH was found in young tubers indicating the crucial role of this enzyme in tuber development. The results of this research will be very helpful in understanding the biochemical properties, catalytic effectiveness and evolutionary dynamics of the G6PDH enzyme in the *Solanaceae* family.

Keywords: Potato (*Solanum tuberosum L.*); Glucose-6-Phosphate Dehydrogenase (G6PDH); Molecular studies by *in silico* analysis; Organ extracts; G6PDH activities

Chapter 1: INTRODUCTION

1.1 About Potato (*Solanum tuberosum* L.)

Potato *(Solanum tuberosum L.)* has gained global popularity as a food item, partly due to its extremely high yield per unit area in comparison to many other food crops such as wheat and rice (Dutt et al. 2017). It belongs to the *Solanaceae* family also known as Nightshades family. It is an annual, herbaceous, dicotyledonous and vegetatively propagated plant (Sarkar 2008). Potato is a seasonal crop grown in temperate, tropical, and subtropical regions (Reddy et al. 2018). It can be vegetatively propagated through true potato seeds (TPS). Plants grown using tubers are the clones of parent plant, however, plants grown using seeds produce different varieties (Sarkar 2008). It is flexible, inexpensive, nutritionally rich staple food popular among developed countries. A medium-sized (150 g) potato has around 30 g carbohydrates, 3 g fibre, and 3 g of protein. Every year, billions of dollars are exchanged in the worldwide potato trade, with numerous nations importing and exporting potatoes and processed potato products. This trade helps to support the economies of different countries and provide opportunities for small-scale farmers to access international markets (Abelenda et al. 2011). Around 5000 varieties of potato along with 200 wild species and sub species have been reported till date. Tobacco (*Nicotiana tabacum* and *Nicotiana rustica*), belladonna (*Atropa belladonna*), and toxic jimsonweed (*Datura stramonium*) are also members of the Solanaceae family (Reddy et al. 2018). Out of the 1500 species of the *Solanum* L. genus, 150 are known to produce tubers (Abelenda et al. 2011 and Reddy et al. 2018). *Solanum tuberosum*, is a tetraploid with 48 chromosomes. It is reported to be the world's most developed species. It is a cross between the diploid species *Solanum stentotomum* and the diploid weed S*olanum sparsipilum*. Ploidy levels range from four diploid species—*S. stenotomum, S. goniocalyx,* and *S. Ajanhuiri* with ($2n = 2x = 24$) chromosomes is hexaploid species with ($24 = 6x = 72$) chromosomes. Additionally, *S. chaucha* and *S. juzepczukii* are two triploid species ($2n = 3x = 36$), while *S. curtilobum* is one of two farmed tetraploid species ($2n = 5x = 60$). The tetraploid cultivated potatoes are not diploid, so that there are four interchangeable genes at each locus. Hence, the plant is heterozygous and show polyploidy.

1.2 History of Potato

Potatoes were first cultivated around 10,000 years ago in the Andes Mountains of Southern America, and were derived from the wild species of *Solanum breviculae* complex (Abelenda et al. 2011). In the 16th century, Spanish explorers who discovered potatoes in South America brought them to Europe, where they became a popular temperate crop (Reddy et al. 2018). Potatoes were first introduced to India in 1610 and were initially grown in the gardens of Surat and Karnataka in 1675. Commercial cultivation of potatoes began in the Nilgiris hills in 1822. The British East India Company introduced potatoes to the northern hills of India and Sri Lanka, thinking that it would be a strong competitor to the abundant rice grown in India (Singh and Rana 2014). However, potatoes were readily accepted in India. In 1945, the Government of India's agricultural advisor at a time developed Central Potato Research Institute (CPRI) to conduct a national potato research and development program. The CPRI's headquarters were moved to Shimla in 1956. CPRI was first given permission to conduct fundamental and strategic research projects that were intended to improve potato production and consumption in the nation by strengthening varieties and renewable technology. India is the second-largest producer of potatoes in the world, with a production of 51.3 million tonnes from an area of around 2 million hectares in 2017-18. As potato is a cold-weather crop, the harvest season for potatoes is from March to April. They typically take 90 to 100 days to reach full maturity. Temperature below 10°C (50°F) and over 30°C (86°F) disturb tuber development. The increase in population rate, there is a direct need for its improvement in the term of quality and quantity.

1.3 Morphology and Anatomy

Potato is a perennial and herbaceous plant that can grow up to 24 inches in height. It is composed of two to four pairs of primary leaflets that are arranged on the mid-rib with a terminal leaflet present on the end. Each node on the potato plant typically has one major leaf which is green in colour. The flowers come in various colours such as purple, white, and pink. The tuber develops from underground stem that enlarges to become a storage organ. This tuber is the consumable part of the plant (Struik 2007). The tubers are present in different shapes such as compressed, round, or oval with skin that ranges from yellow to purple colour. The tubers bear auxiliary buds on their surface, which develop into the stem when planted to form the next vegetative generation. Fig. 1.1 shows the various organs of the potato plant.

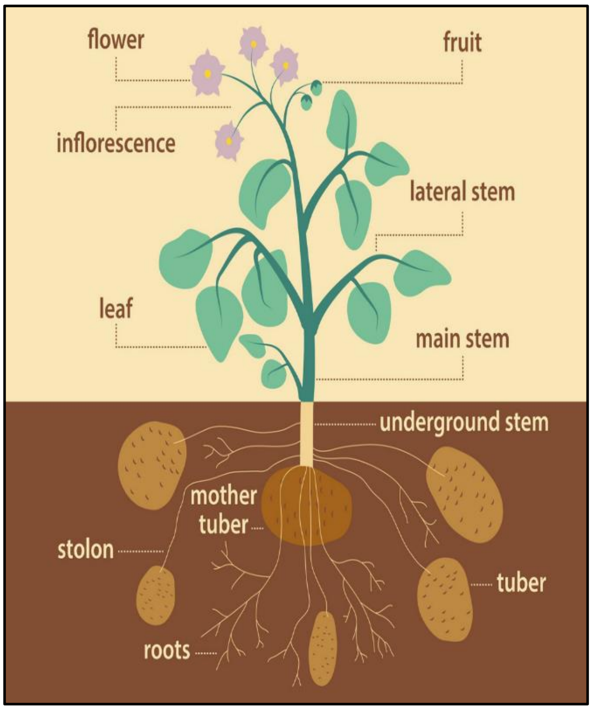


Fig. 1.1 Potato plant (Ref: https://cdn1.vectorstock.com)

1.3 Nutritional Aspects of Potato

The primary component in potato is carbohydrate. The United States Department of Agriculture reports that a raw potato has 79% water, 17% carbs, 2% protein, and a negligible amount of fat. Vitamins C and B6 are abundant in potatoes, along with other minerals including potassium, magnesium, and iron. The primary vitamins included in potatoes are folate, thiamin, and riboflavin (Beals 2019). Methionine and cysteine levels in potato tubers are higher than in cereals (Abelenda et al. 2011). Carotenoids and phenolic acids are the two phytonutrients which are also found in potatoes (Brown et al. 2005; Liu 2013; McGill et al. 2013). The minerals and nutritional elements included in potatoes have a key role in the promotion of health. it reduces the chance of chronic diseases (Beals 2019). Although eating potatoes is a healthy way to get energy and other nutrients, it may also sometimes lead to unpleasant health concerns including indigestion or other allergic responses (Zaheer and Akhtar 2016). Potato organs are known to contain several harmful substances, such as glycoalkaloids. According to reports these substances are essential for the plant's defence against predators (Friedman 2006). Additionally, potatoes contain antioxidants and dietary fibre that protect against illnesses linked to ageing (Singh et al. 2020). Fig. 1.2 Shows the nutritional value of potato.

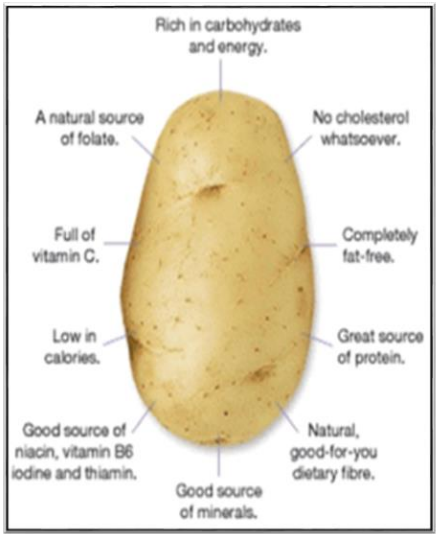


Fig. 1.2 Picture showing nutritional information of Potato (Ref: http://cremeroyale.com.)

1.4 Tuberization: An Important Aspect of Potato Life Cycle

Tuberization is a coordinated morphophysiological process that is a crucial survival strategy for potato plant (Van den Berg et al. 1996). The two main storage organs in higher plants are tubers and seeds. In terms of structure and development, seeds and tubers are fundamentally different. Tubers (somatic storage tissues) are generated by altering somatic structures of the plant, but seeds have a closely controlled developmental route that begins with fertilisation and ends in the seed's planned dehydration (Cutter 1982). Potato tubers are structurally modified subterranean stems that have progressively grown larger via a process of constrained cell division and cell growth. There are environmental factors cause the potato plant to begin tuberizing (Cutter 1982). These are cold night time temperatures and short-day photoperiods (SD). These ideal growing conditions for potatoes are only present in the subtropical Indo-Gangetic plains during the winter. This explains the growth conditions for potatoes in temperate regions of North America and Europe are quite different from those in India. Extrinsic and intrinsic genetic, physiological, and environmental elements interact in this intricately expanding process. Two different mechanisms are connected to the morphological and biochemical levels of potato tuber development. Morphological processes include stolon formation and subsequent tuber induction at the stolon tip. Contrarily, tuber induction is characterised by changes in the direction of cell development in the subapical area of the stolon tip as well as cell division and cell expansion at the cellular level (Vreugdenhil et al. 1999). Starch synthesis and storage protein buildup are two examples of biochemical processes. Both morphological and metabolic processes must be regulated by certain gene expression patterns (Taylor et al. 1998 and Verhees et al. 2002). The induction, initiation, and development of an underground shoot (the stolon), the end of the stolon's longitudinal growth, and the subsequent induction, initiation, and growth of a storage organ (the tuber) are all crucial phases in the tuberization process. When photoperiodic sensing is mediated by phytochrome B and GA in leaves, RNA then serves as a systemic signal in the long-distance signalling pathway to start tuberization in the subsurface stolon. Fig.1.3 Shows the prominent stolon hook.

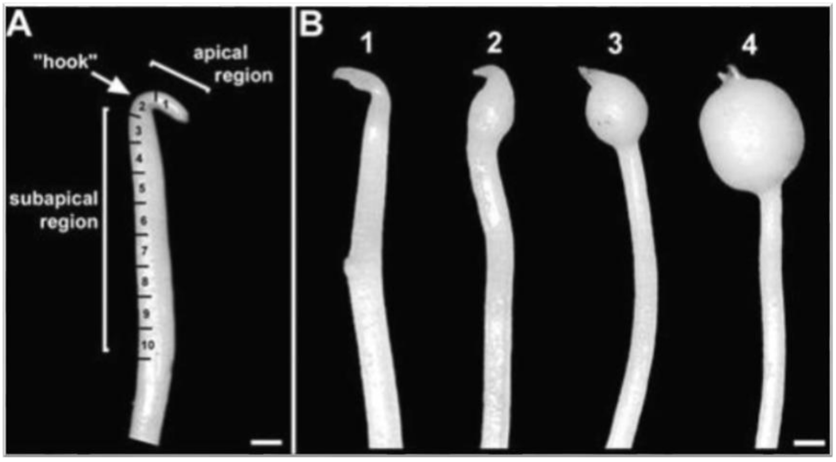


Fig. 1.3 A. Prominent stolon hook B. Different stages of tuberization: 1. Hooked stolon tip with no swelling, 2. Slight swelling below apex that results in straightening of apical hook, 3. Stolon tip swells until the hook is completely open and develops into tuber, 4. Apical bud is incorporated into growing tuber. (Viola et al. 2001)

1.5 Stages of Tuberization

There is total 8 stages which come under tuberization, In the 1, 2 and 3 stage the longitudinal growth of the tuber starts to appear.

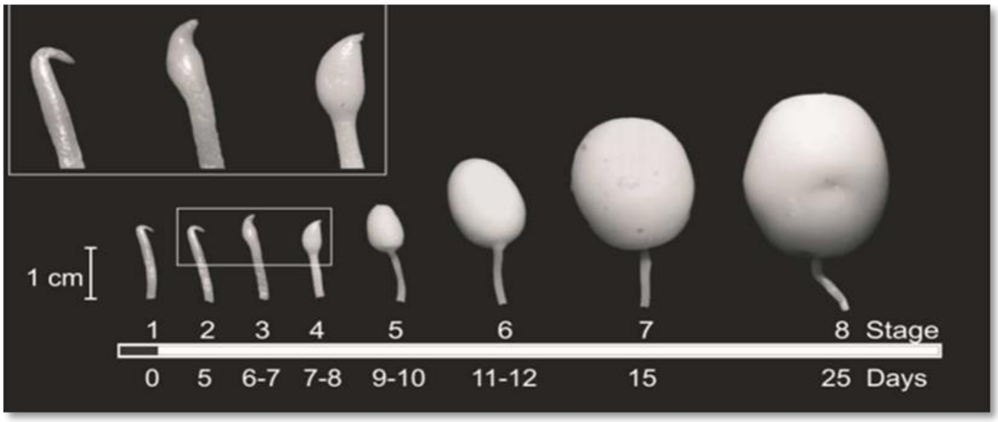


Fig. 1.4 Picture showing different tuberization stages of potato. (Ref: https://ars.els-cdn.com)

In 1st stage – Hooked stolon with no apparent swelling, 2nd stage – slight swelling on the hook starts which leads to the hook begin to straighten, 3rd stage – stolon terminal continues to swell until the hook is completely open, 4th stage – Tuber continues to develop (The developing tuber is approximately the twice the diameter of the stolon. In 5,6,7 and 8 stage the tuber increases in weight. 5th stage 0.6 -1.5 g tuber,6th stage 1.5-2.5g tuber,7th stage 2.5-5.0g tuber 8th stage 5.0-10.0g tuber as shown in (Fig. 1.4).

1.6 Factors Affecting Tuberization

1.6.1 External Factors

There are several external factors that have a significant impact on tuber formation, including high irradiance, photoperiod, carbohydrate and nitrogen supply, temperature, and phytohormones. A decrease in temperature during the night to 14-20 ºC can sharply reduce the number and weight of growing tubers, while an increase in temperature to 25-27 ºC during the night can lead to the same outcome.

1.6.2 Phytohormones

Plant growth regulators plays an important role during tuber growth and development. Gibberellins (GAs) are known to be the dominant regulators of stolon extension during the process of tuberization. GA levels has been reported to be high in non-induced stolon as promote longitudinal cell expansion (Xu et al. 1998). However, decline in GAs were observed after tuber induction. Cytokinin's (CKs) are known to be involved in the regulation of cell cycle and sprouting. Abscisic acid (ABA) plays a crucial role in tuberization by counteracting the inhibitory effects of GAs. Levels of Jasmonic acid (JA) were found to be high during the transition of stolon into tubers indicating its important role during early stages of tuberization (Nam et al. 2007).

1.6.3 Transcription Factors

Proteins are the functional form of gene. Various proteins known as transcription factors control gene activity that indirectly influence hormone levels. Numerous DNA-binding proteins are involved in controlling the growth of plants, including the development of tubers and meristem (Hannapel et al. 2004). Such processes are molecularly controlled or regulated by transcription factors (TF). Protein-nature TFs operate as developmental switches by binding to DNA and modulating the expression of the targeted genes. For example, the proteins POTM1 (potato MADS-box), POTH1 (potato homeobox), StSP6A, StBEL5, StPHYB, StCONSTANS, and sucrose transporters like StSUT4 and StSP5G are all implicated in the regulation of tuber development (Dutt et al. 2017). POTM1 is a member of the MADS-box gene subfamily SQUAMOSA. POTM1 is known to be expressed in various vegetative organs such as juvenile tubers, leaves, stolon, and roots (Kang and Hannapel, 1995, 1996). Kang and Hannapal (1995) reported the isolation of POTM1 from growing tuber indicating important role of this TF during early stages of tuberization. POTH1 is a member of *TALE* superclass. It regulates the tuber

development in conjunction with another member of this family namely StBEL5 by modulating the GAs levels in the stolon tip (Rosin et al. 2003 and Chen et al. 2004).

1.6.4 ROS Metabolism Influences Tuberization

Reactive oxygen species (ROS) are harmful byproducts of oxygen-containing metabolic processes including photosynthesis and respiration. Over the last several years, ROS plays a significant signalling function in plant growth, development, and reaction to biotic and abiotic stress situations. They are very dangerous and cause damage to DNA, proteins, and lipids, which affects how cells operate and may even cause cell death. Various enzymatic and non-enzymatic antioxidants play crucial role in detoxification ROS byproducts. Among enzymatic machinery various enzymes namely ascorbate peroxidase (APX), catalase (CAT) and superoxide dismutase (SOD) are involved in ROS detoxification in a cell. Tuberization is a complex process that involves the continual oxidation of biomolecules. Various enzymatic antioxidants have been reported to be highly active during early stages of tuber development (Aggarwal 2008; Yu et al. 2012).

1.7 Enzymes Involved in Tuberization

Enzymes play an important role in the process of tuberization. These enzymes are implicated in various metabolic pathway that contribute to the initiation and growth of potato tubers. Lipoxygenase (LOX), a member of dioxygenase family, is found to be active in the growing tubers indicating a positive correlation between LOX activity and tuber development (Kolomiets et al. 2001). GA20oxidase1 (GA20OX1) is a biosynthetic enzyme which is repressed during tuberization. However, this enzyme is known to be actively expressed during stolon elongation. Apart from this, carbohydrate metabolism plays major role during tuber growth and development. Various enzymes such as hexokinase, sucrose synthase, fructokinase, UGPase and AGPase plays an important role during tuberization. AGPase is involved in starch biosynthesis and its activity was found to be high during potato tuber dormancy and sprouting as revealed by Hannapel et al. (2017). Others enzymes such as Phosphofructokinase, Aldolase, Trio phosphate isomerase, Glyceraldehyde3-phosphate dehydrogenase, Phophoglycerate kinase, Enolase, Pyruvate kinase, 6-phosphogluconate dehydrogenase, Transketolase are involved in the conversion of sucrose to hexose-phosphate that is the crucial during sucrose synthesis-an important process during tuber development (Appeldoorn et al. 2002). Oxidative pentose phosphate pathway (OPPP) is involved in the energy production that is essential for the growth and development of potato tubers (Scharte et al. 2009). An important enzyme of this pathway namely Glucose 6 phosphate dehydrogenase is known to be expressed in young tubers. Differential regulation of G6PDH isoenzyme activities has been observed in potato (Hauschild et al. 2003). Keeping this view, we have focussed on the G6PDH. Salient features and recent advances of this enzyme are discussed in the following sections.

Chapter 2: REVIEW OF LITERATURE

2.1 History of Glucose-6-phosphate dehydrogenase (G6PDH)

Glucose-6-Phosphate Dehydrogenase (G6PDH) was discovered first in yeast in 1930's, during the enzymatic oxidation of Glucose-6-phosphate (G6P) to 6-phosphogluconate (6PG). In the first instance it was called Zwischenferment (ZWF1) (Masi et al. 2015). In plants, G6PDH-cDNA was first isolated from potato (Graeve et al. 1994). Till date, It has been identified in various monocot and dicot crops (Hauschild et al, 2013; Landi et al. 2016). This enzyme plays crucial role during plant growth and stress resistance.

2.2 Introduction to G6PDH

G6PDH (EC 1.1.1.49) enzyme is a key player in the metabolic and physiological processes of plants. It is found in many different organisms across the living world, including bacteria, fungi, plants, and animals (William et al. 1980). In plants, It is a metabolic hub (Fig. 2.1) between two important cycles namely glycolysis and pentose phosphate pathway (PPP) (Hauschild et al. 2003). The main role of the enzyme is to generate the NADPH for reductive biosynthesis which is done by sequential action of two dehydrogenase in the first part of the glycolysis pathway. In the second fully reversible step of the route, sugar phosphate such as erythrose-4-phosphate is produced which serves as a precursor for the synthesis of secondary plant products (Schaewen et al. 1995). It provides NADPH for anabolic metabolism, which includes the production of fatty acids, mevalonate, and amino acids (Graeve et al. 1994). Two G6PDH isoforms namely cytosolic and chloroplastic are found in plants (Graeve et al. 1994). This enzyme is involved in the production of NADPH which is essential for scavenging reactive oxygen species (ROS) during stress conditions (Jiang et al. 2022). Mutations in the G6PDH gene in plants have been shown to affect growth and development along with their ability to tolerate stress. Therefore, understanding the regulation and function of G6PDH in plants is important for improving plant growth, development, and stress tolerance (Landi et al. 2021).

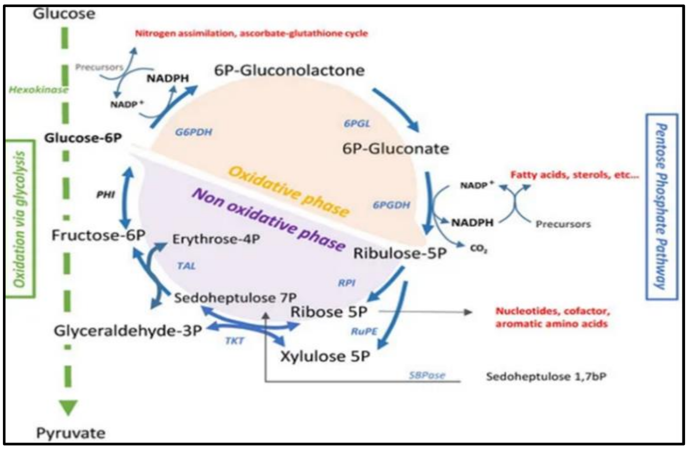


Fig. 2.1 Picture showing the reactions of the pentose phosphate pathway (PPP) and its connection to glycolysis. Ref: (Jhamg et al. 2022).

2.3 Classification of G6PDH Enzyme

It is a crucial enzyme in plants because it controls the carbon travel through the pentose phosphate pathway (Jiang et al. 2022). Genome-wide analysis of G6PDH *in Arabidopsis* indicated 6 genes encoding 4 plastidial and 2 cytosolic isoforms. Zhao et al. (2020) reported 6 G6PDH gene. In strawberry, 19 G6PDH has been reported (Lei et al. 2022). In higher plants, G6PDH has been reported to be present in two cellular compartments such as cytoplasm and plastids. Further, these isoforms are distinctly divided into subgroups. Plastidial G6PDH include three functional isoforms belonging to two distinct groups [P1 (G6PD1), P2 (G6PD2, G6PD3)] and a non-functional one (G6PD4) belonging to the P0 group. Cytoplasmic G6PDH is subdivided into two groups [P5 (G6PDH5) and P6 (G6PDH6)], providing 60–80% of the total activity. These subgroups are thoroughly discussed in the following section:

2.3.1 Cytosolic G6PDH (Cy-G6PDH)

This isoform is mainly present in cytosol of G6PDH. G6PDH5 and G6PDH6 is commonly found in many plant species and is primarily localized in the cytosol, it expressed differently and were initially purified from roots which is mainly shown in *Arabidopsis*. The NADPH/NADP+ ratio controls and NADPH inhibits energy fluctuations which is the key role of NADPH (Wakao et al. 2008).

2.3.2 Chloroplastic G6PDH (Ch-G6PDH)

The chloroplastic G6PDH is a P1-G6PDH and that its post-transcriptional inhibition is mostly caused by high redox status. P1-G6PDH is mainly obsereves in spinach. (Lendzian et al. 1975; Anderson et al. 1976). This activates the Calvin cycle and a number of stromal target enzymes, including fructose-1,6-bisphosphatase, NADP-malate dehydrogenase, phosphoribulokinase, and others. (Lendzian et al. 1980).

2.3.3 Plastidial G6PDH (P2-G6PDH)

The plastid P2-G6PDH is mainly seen in all plant organs, including developing tissues, photosynthetic tissues, and non-photosynthetic tissues. It is strongly correlated with the rate of the OPPP (Wakao et al. 2005) The regulation of P2-G6PDH activity is more likely to be similar to that of P1-G6PDH, but their sensitivity to NADPH is different. (Lendzian et al. 1980).

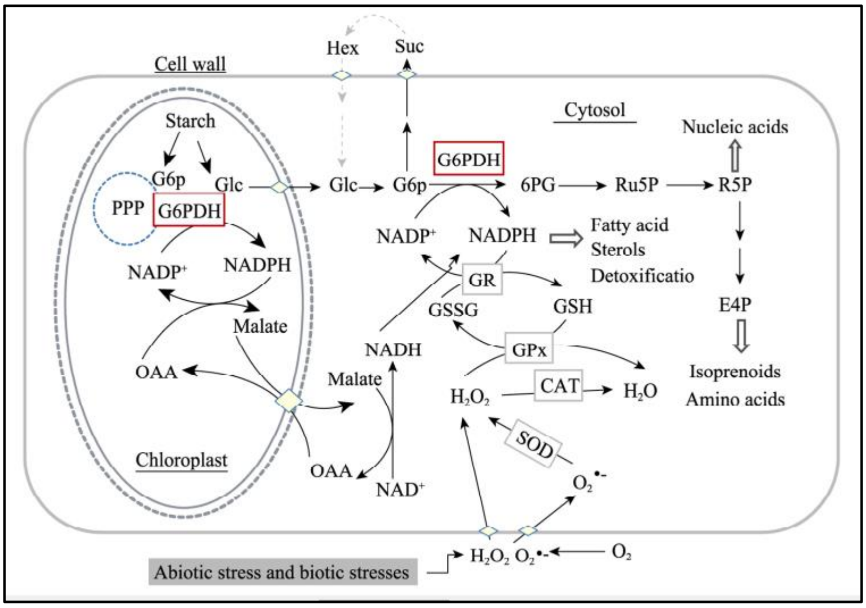


Fig. 2.2 Sites of production of G6PDH isoform Ref:(https://www.creative-enzymes.com)

2.4 Role of G6PDH in Plants

In plants, the biochemical activity of G6PDH is crucial for maintaining the balance of NADPH and NADP+ in the cell. NADPH produced by G6PDH in the pentose phosphate pathway is used in several essential plant processes, including fatty acid biosynthesis, nucleotide biosynthesis, and the reduction of oxidized glutathione. Additionally, NADPH act as an electron donor in the photosynthetic electron transport chain (Jiang et al. 20220).

2.4.1 G6PDH in Potato

In potato, three G6PDH isoforms are reported to be present in different cellular compartments. The first isoform is located in the cytosol and is involved in the production of NADPH for fatty acid biosynthesis. The second isoform is found in the plastids and is involved in the biosynthesis of starch. The third isoform is located in the mitochondria and is involved in the production of NADPH for respiration. The expression of the different isoforms of G6PDH in potato can be influenced by a range of factors, including developmental stage, environmental conditions, and stress. G6PDH plays different function which is as follow section.

2.4.2 G6PDH as Stress Regulator

G6PDH in plants plays an important role for protecting against oxidative stress caused by ROS. These are by products of metabolic processes that can damage plant cells and tissues, but NADPH produced by G6PDH can be used to neutralize ROS through the reduction of glutathione, which is an important antioxidant in plants. CyG6PDH controls germination by maintaining steady state level of ROS. The involvement of different isoforms of g6pdh i.e. cytosolic as well as plastid ones in dormancy release was investigated in pea (Barba et al. 2012).

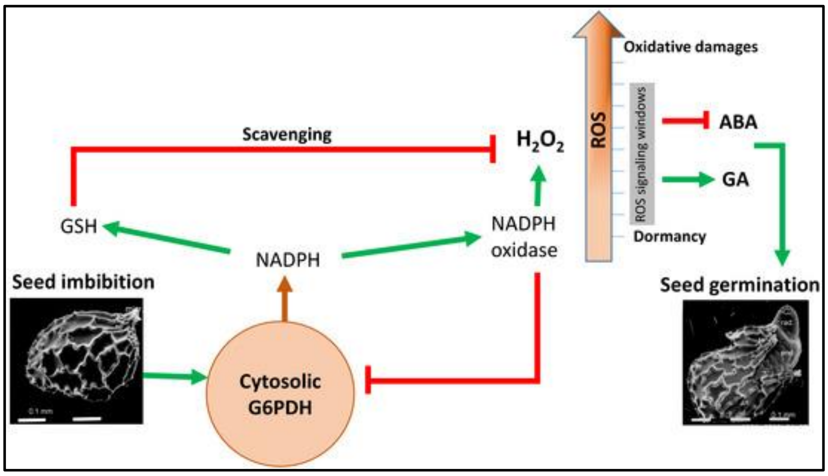


Fig. 2.3 Cytosolic G6PDH in radicle of imbibed seed modulates ROS homeostasis and hormonal signalling in the control of seed germination in pea plant (Barba et al. 2012).

2.4.3 G6PDH role in Nitrogen Assimilation

The most important factors for plant growth and productivity are Nitrogen. Pentose phosphate pathway plays an important role in non-photosynthetic tissue i.e., root system Plastid G6PDH plays a great role in nitrogen assimilation. Recent data shows NiR activity in the barley roots and in *Arabidopsis* G6PDH shows a great role in nitrogen assimilation. (De Jong et al. 2014; Lejay et al 2008).

2.4.4 G6PDH role in Sugar Signalling

Plants have the complex mechanistic system that includes hexose-, disaccharide-, and the pentose phosphate signalling pathways. Studies shows that the rose bud have a positive response to sugar signalling (Porchar et al. 2022). Signalling pathway is reported sink organs. However, OPPP-dependent sugar signalling might be more complex and operate through different pathways.

2.5 Origin of the Problem

A staple meal for many communities, the potato (*Solanum tuberosum* L.) is an economically significant crop grown all over the globe. Tubers has been used as a primary nutrient and a source of carbohydrates in diet. It is also consumed in the form of processed food items such as French fries and chips. Tuberization is a developmentally regulated process. It involves the interaction between various extrinsic and intrinsic factors. In last few decades researchers have made considerable progress in understanding this complex process. Various environmental factors determine the yield of potato. Climate change has a drastic impact on plant growth and development. Tuberization is the sensitive stage of potato life cycle. There is a dire need to develop varieties that can withstand the negative effects of global warming without affecting the yield. Carbohydrate metabolism is the dominant regulator of tuber development. Various enzymes are known to be dominantly expressed throughout tuberization stages. G6PDH is a crucial enzyme interconnecting Pentose phosphate pathway and glycolysis. Pentose phosphate pathway provides energy in the form NADPH that is used in multiple metabolic pathways taking place in tuber development. Keeping this in view, this study is focussed on G6PDHs in potato.

2.6 Objectives

This thesis work was based on the following objectives:

- Sequence analyses, multiple sequence alignment, motif search and predicting structural features of the of Glucose-6-Phosphate Dehydrogenase (G6PDH) sequences from different plant species including potato by *in silico* approaches

- Determination of G6PDH activities in different potato organs including developing tubers

Chapter 3: MATERIALS AND METHODS

3.1 *In-silico* Analysis

Nucleotide sequence of the cDNA from Indian potato cultivar was obtained.
↓
BLASTn at NCBI(http://www.ncbi.nlm.nih.gov)
↓
BLASTp at NCBI(http://www.ncbi.nlm.nih.gov)
↓
The sequence were retrieved and selected for further analysis
↓
Sequence analysis, prediction of motifs and 3-dimensional structure of the G6PDH

3.1.2 Genome Wide Identification

The genome-wide analysis of the G6PDH gene was conducted using the Ensembl plants (https://plants.ensembl.org/index.html). It is a vertebrate genome browser that supports research in comparative genomics, evolution, sequence variation, and transcriptional regulation. Gene annotation is carried out via Ensembl. The Swiss Institute of Bioinformatics' (SIB) ProtParam tool from the Expert Protein Analysis System ExPASy was used (SIB; https://www.expasy.org/) to predict various physico-chemical properties such as molecular weight, isoelectric point (pI), aliphatic acid and GRAVY index.

3.1.3 Multiple Sequence Alignment and Phylogenetic Analysis

NCBI tools were used to explore various isoforms of G6PDH family from different plant species namely potato, tomato, tobacco and capsicum. Multiple sequence alignment was done to know the convergence/divergence in crucial motifs/domains of G6PDH in aforementioned plants. The analysis was performed using MultAlin server (http://multalin.toulouse.inra.fr/multalin/). This server was established by Florence Corpet in 1988. It is used to study the closely related genes or proteins in order to find the evolutionary relationship.

For phylogenetic analysis, various G6PDH gene family members from plants such as potato, tomato and *Arabidopsis* were identified from genome centric portal namely Ensemble Plants. The MEGA-X software was used to construct the phylogenetic tree by using maximum likelihood method (Kumar et al. 2018).

3.1.4 Identification of the Conserved Domains

Prediction of motifs was done by uploading the amino acid sequence of the G6PDH to MY HITS server (https://myhits.sib.swiss/cgi-bin/motif_scan).

3.1.5 Expression Atlas

Expression Atlas (https://www.ebi.ac.uk/gxa/home) was explored for *in silico* analysis of G6PDH gene. The European Bioinformatics Institute is responsible for maintaining this database. It is helpful to analyse the expression of genes and proteins. It gives easily available information on the prevalence and distribution of RNA (and proteins) are present in many cell types, organs and other biological situations across all species. Gene expression information was retrieved in the form of heatmap

3.1.6 Three-Dimensional Structure Prediction

It is critical to understand a protein's structural information in order to predict its biological activity. For the prediction of protein 3-D structure, I-TASSER server was used (https://zhanggroup.org/I-TASSER/). This server is a web-based platform that uses I-TASSER algorithms to predict protein structure and function.

3.1.7 Ramachandran Plot

The Ramachandran plot was developed in 1963 by G. N. Ramachandran, C. Ramakrishnan, and V. Sasisekharan. The plot shows the distribution of the angles φ and ψ for each amino acid residue in a protein. The Ramachandran plots were made using PROCHECK server (https://saves.mbi.ucla.edu/). The plot is widely used in protein structure analysis and prediction, as well as in the study of protein folding and stability. The 3-D structure was subjected for Ramachandran plot analysis.

3.1.8 STRING database

String database (https://string-db.org/) is known and predicted database of protein-protein interactions. It aims to integrate all known and predicted associations between proteins, including both physical interactions as well as indirect (functional) associations. The amino acid sequence of G6PDH was uploaded on STRING database in order to know the functional associations within the cell.

3.2 Experimental Analysis

3.2.1 Plant Materials

Pathogen free germplasm of Indian potato cultivars namely Kufri-Chipsona-1 (KC-1), Kufri-Chipsona-2 (KC-2), Kufri Pukhraj (PR), Kufri Jyoti (KJ), and Desiree (DE) procured from Central Potato Research Institute (CPRI) are routinely maintained on MS media under proper laboratory conditions (25–27 °C, ~ 70% relative humidity under 16-h photoperiod with a light intensity of 40 − 42 µmol/m2/s spectral flux photon of photosynthetically active i.e. 460 − 700 nm radiations with 2.5% sucrose at the interval of 4–5 weeks). These micropropagated plants were then hardened, acclimatized and transferred to the fields for a period of 15-16 weeks (mid of November-March). Different potato organs were harvested at various stages of development (Fig 3.1). The organs were then frozen in liquid nitrogen and stored at − 80 °C for further biochemical studies. Chemicals and enzymes were procured from Sigma-Aldrich Pvt. Ltd, Genei, SRL and HiMedia Laboratories, India.

3.2.2 Biochemical Analysis

a) Preparation of the Extraction Buffer

Extraction buffers are essential tools in biological and biochemical research for isolating specific biomolecules or components from complex samples In order to ensure the release of the required molecules, they contain substances like detergents, chaotropic agents, or enzymes that disintegrate cellular structures and break down cell membranes. Extraction buffers also assist in solubilizing the target molecules, stabilising them to stop them from degrading, and removing undesirable impurities. The physiological conditions required for the stability and functioning of the molecules being extracted can also be maintained by them. To carry out G6PDH isolation the enzyme buffer was prepared by using the reagents mentioned in Table 1.

Table 1 Composition of extraction buffer

Reagents	Amount
Tris HCL	40 mM
MgCl2	3 mM
EDTA	1 mM
PMSF	0. 1 mM
Benzamidine	1 mM
Beta- mercaptoethanol	14 mM
NADP	24 µmol

b) Preparation of Enzyme Assay Buffer

Determination of G6PDH was done by the method described elsewhere (Jammer et al. 2015). The composition of the assay buffer is mentioned in Table 2. Increase in absorbance due to the conversion of NADP to NADPH at 340 nm was recorded.

Table 2 Composition of kinetic buffer assay

Component	**Composition**
$MgCl_2$	5 mM
G6P	1 mM
NADP	0.4 mM
TRIS-HCl	100 mM
Extract	100 µL

c) Preparation of Standard Curve for Protein Estimation

The Lowry method, also known as the Lowry protein assay, is a widely used colorimetric method for quantifying protein concentrations in biological samples. It was first introduced by Oliver H. Lowry and his colleagues in 1951.This method was employed to estimate the total protein concentration of various potato plant organs such as leaves, stem and tubers at different stages of development. In a protein, phenolic compound in amino acids namely, tyrosine and tryptophan produce a blue purple complex, which absorbs at 750 nm. Boven Serum Albumin (BSA) was used as a standard (Table 3).

Table 3 Chemical reagents used in protein estimation.

No.	Chemical name
Solution I	0.1 M alkaline Sodium Carbonate (2% Sodium carbonate in 0.1M NaOH)
Solution II	Freshly prepared copper sulphate-sodium potassium tartrate solution.
Solution A	Solution I: Solution II (50:1)
Solution III	Folin-Ciocalteu reagent: distilled water (1:1)

d) Estimation of Total Protein

100 µL of protein sample was taken and volume was maintained upto 1 mL using distilled water. Then 5 mL of reagent I was added and after thorough mixing, sample was incubated for 10 min. 500 uL of Follin reagent was added to each sample and the test tubes were left in the dark for roughly 30 minutes and absorbance was taken at 750 nm. Using the standard curve equation, the amount of protein was calculated in various plant organs

Chapter 4: RESULTS AND DISSCUSSION

4.1 *In Silico* Analysis

In this study, we obtained two G6PDH cDNA encoding cytoplasmic and chloroplastic isoforms. This is basically an enzyme that participates in the glycolysis and pentose phosphate metabolic pathways. Chloroplastic isoform consists of 577 amino acids encoded by 2058 bp cDNA sequence (Acc: NM_001288109). In cytoplasmic isoform contains 571 amino acids encoded by 1718 bp gene sequence (Acc: NM_001288468). Various *in silico* tools were used to analyse and characterize the G6PDH isoforms from potato as discussed in the following sections:

4.2. BLAST Analysis

4.2.1a Sequence Analysis of *StCyG6PDH* by BLASTn

To compare the nucleotide sequence of *ChG6PDH* and *CyG6PDH* gene with another plant species BLASTn was performed. The result showed significant *ChG6PDH* and *CyG6PDH* sequence similarity with various plant species belonging to both *Solanaceae* and non-*Solanaceae* family members. From all the sequences, we selected 9 nucleotide sequences belonging to *Solanaceae* family for further analysis.

BLASTn analysis between both *StG6PDH* (*ChG6PDH* and *CyG6PDH*) showed no significant similarity among them. However, BLASTn analysis of *StChG6PDH* showed significant sequence similarity with various other plant species. For e.g., 85% query coverage and 97.61% sequence identity with *Solanum lycopersicum (*XM_004243519), 98%query coverage and 97.59% sequence identity with *Capsicum annum (*XM_016724979), 98% query coverage and 92.55% sequence identity with *Nicotiana tabacum (*XM_009783930), 88% query coverage and 85.04% sequence identity with *Ipomoea triloba* (XM_031240175.1) as shown in Table 4.

Table 4 Details of some homologous *StCyG6PDH* gene sequence as available in database

Species name	Accession no	Max score	Identity	E value	QC%	Bp
Solanum lycopersicum	XM_010317635	2905	97.59%	0.0	98%	1879
Solanum stenotomum	XM_049529069	3101	99.24%	0.0	100%	1851

Solanum verrucosum	XM_049490162	3101	99.24%	0.0	100%	1771
Solanum pennellii	XM_015211311	2900	97.53%	0.0	98%	1879
Capsicum annum	XM_016706198	2423	92.55%	0.0	98%	1901
Nicotiana tabacum	XM_016659313	2425	94.69%	0.0	90%	1808
Nicotiana tomentosiformis	XM_033656392	2420	94.63%	0.0	90%	1921
Nicotiana sylvestris	XM_009776478	2374	93.92%	0.0	91%	1811
Ipomoea triloba	XM_031240175.1	1541	85.04%	0.0	88%	1881

4.2.1b Sequence Analysis of StCyG6PDH by BLASTp

Comparison between both StG6PDH isoforms (ChStG6PDH Acc NO. NP_001275038.1; CyStG6PDH Acc.No. NP_001275397) at protein level showed 90% Query coverage and 51% sequence identity. However, conventional BLASTp analysis showed significant StG6PDH sequence divergence with respect to various plant species belonging to both *Solanaceae* and non-*Solanaceae* family members. We selected 8 polypeptide sequence belonging to *Solanaceae* family for further analysis. The CyG6PDH protein showed 100% query coverage with variable sequence identities in various plant species namely tomato (99.80%, XP_004231802), *Solanum verrucosum* (99.80%, XP_049346119), *Solanum pennellii* (99.41%, XP_015066794), *Capsicum annum* (96.48%, XP_016561682), *Nicotiana tabacum* (96.48%, XP_016432237), *Nicotiana sylvestris* (96.48%,XP_009774777), *Datura stramonium* (96.28%, MCD7453132.1), *Ipomoea triloba* (88.43%, XP_031096033.1) as shown in Table 5.

Table 5 Details of some homologous StCyG6PDH amino acid sequence as available in database

Species name	Accession no	Max score	Identity%	E value	Qc %	Aa
Solanum lycopersicum	XP_004231802	1056	99.80%	0.0	100%	511
Nicotiana sylvestris	XP_009774777	1022	96.48%	0.0	100%	510

Solanum verrucosum	XP_049346119	1056	99.80%	0.0	100%	511
Solanum pennellii	XP_015066794	1052	99.41%	0.0	100%	511
Capsicum annum	XP_016561682	1020	96.48%	0.0	100%	510
Nicotiana tabacum	XP_016432237	1022	96.48%	0.0	100%	510
Datura stramonium	MCD7453132.1	1018	96.28%	0.0	100%	510
Ipomea triloba	XP_031096033	956	88.43%	0.0	99%	512

4.2.2a Sequence Analysis of *StChG6PDH* by BLASTn

The *ChG6PDH* gene showed 85%query coverage and 97.61%sequence identity with *Solanum lycopersicum (*XM_004243519), with 82%query coverage and 95.18%sequence identity with *Capsicum annum (*XM_016724979), with 84%query coverage and 91.25%sequence identity with *Nicotiana tabacum (*XM_009783930), with 72%query coverage and 85.17%sequence identity with *Ipomoea triloba* (XM_031249590.1) as shown in Table 6.

Table 6 Details of some homologous *StChG6PDH* nucleotide sequence as available in database

Species name	**Accession no**	**Max score**	**Identity%**	**E value**	**Qc%**	**Base pair**
Solanum verrucosum	XM_049518732	3720	99.27%	0.0	100%	2183
Solanum lycopersicum	XM_004243519	3007	97.50%	0.0	85%	1942
Solanum pennellii	XM_015226796	3018	97.61%	0.0	85%	1945
Solanum stenotomum	XM_049557033	3736	99.42%	0.0	100%	2189
Capsicum annum	XM_016724979	2691	95.18%	0.0	82%	2369
Nicotiana tabacum	XM_016648561	2377	91.25%	0.0	84%	1968
Nicotiana tomentosiformis	XM_009595456	2401	91.46%	0.0	84%	1914

Nicotiana sylvestris	XM_009783930	2252	93.58%	0.0	73%	2006
Ipomoea triloba	XM_031249590.1	1530	85.17%	0.0	72%	2394

4.2.2b Sequence Analysis of ChStG6PDH by BLASTp

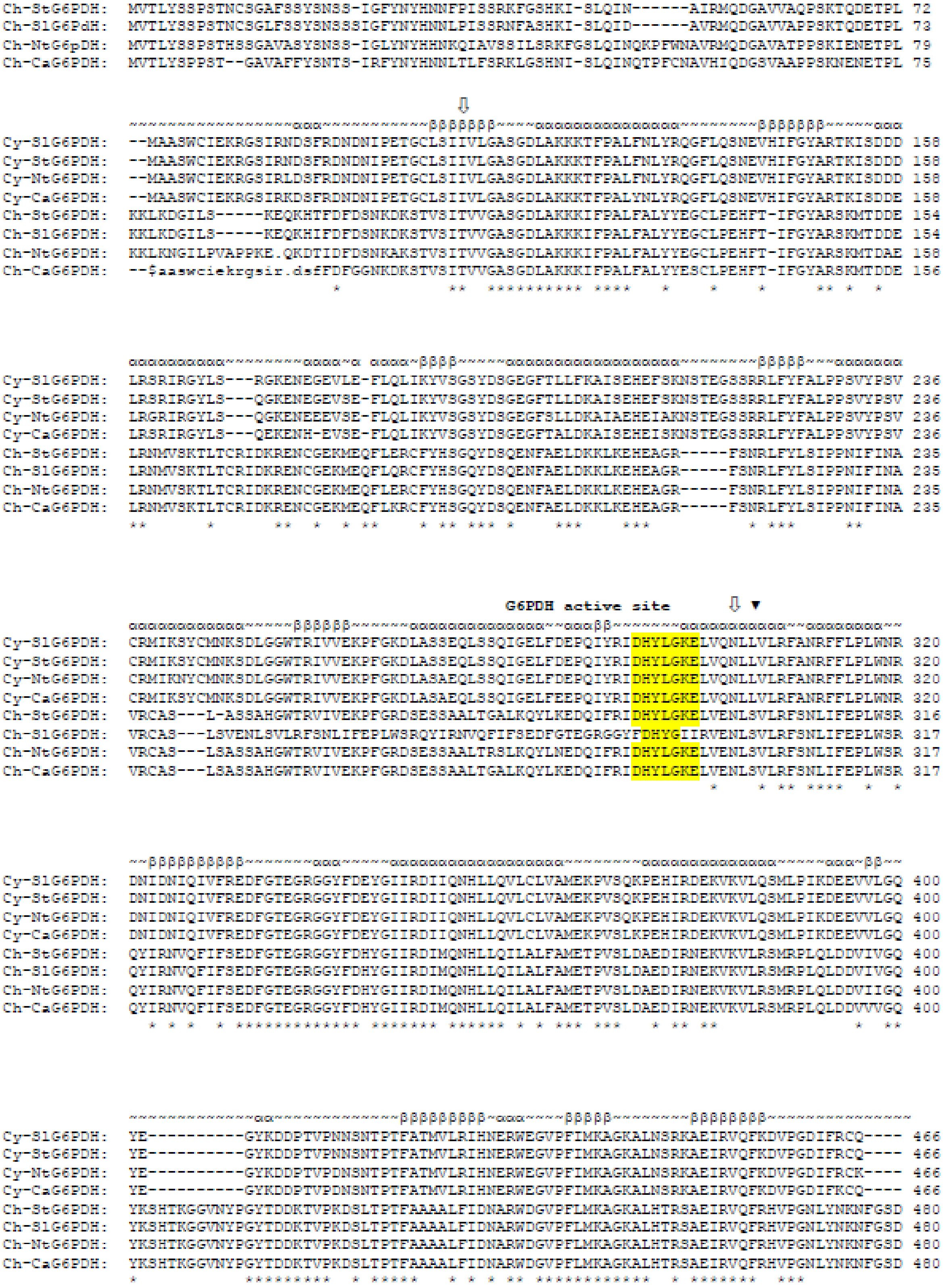
The ChG6PDH shows 100% query coverage with variable sequence identities in various plants namely tomato (98.09%, XP_004243567), *Solanum verrucosum* (99.48%, XP_049374689) *Solanum stenotomum (*99.48%, XP_015082282), *Capsicum annum (*94.45%, XP_016580465), *Solanum pennellii (*98.09%,XP_015082282), *Nicotiana tabacum (*88.83% ,NP_001312522), *Nicotiana sylvestris* (89.17%, XP_009782232) as shown in Table 7.

Table 7 Details of some homologous StChG6PDH amino acid sequence as available in database

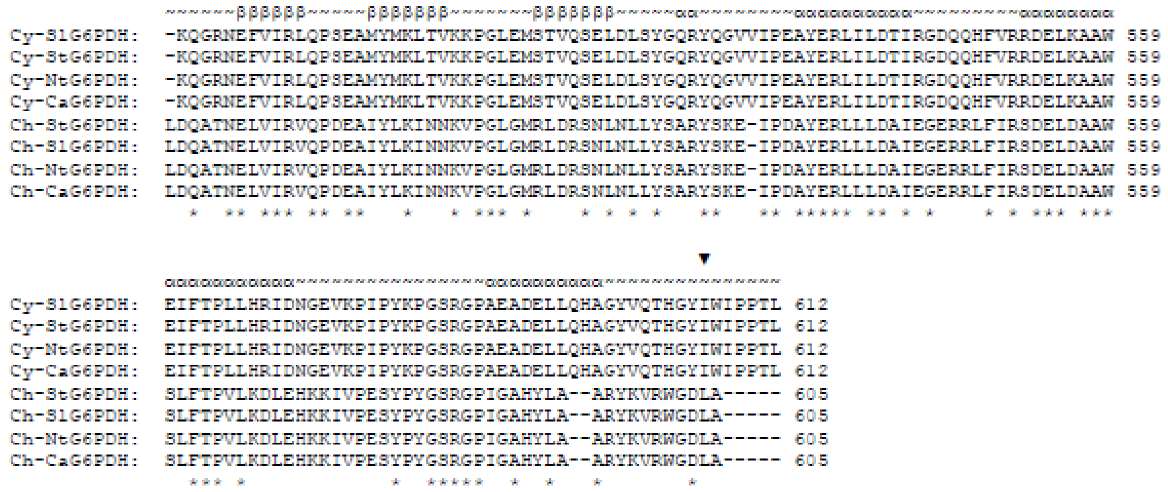
Species name	**Accession no**	**Max score**	**Per Identity%**	**E value**	**QC %**	**Amino acid**
Solanum verrucosum	XP_049374689	1191	99.48%	0.0	100%	577
Solanum lycopersicum	XP_004243567	1178	98.09%	0.0	100%	577
Solanum pennellii	XP_015082282	1176	98.09%	0.0	100%	577
Solanum stenotomum	XP_015082282	1191	99.48%	0.0	100%	577
Capsicum annum	XP_016580465	1123	94.45%	0.0	100%	573
Nicotiana tabacum	NP_001312522	1075	88.83%	0.0	100%	588
Nicotiana tomentosiformis	XP_009593751	1090	91.97%	0.0	100%	586
Nicotiana sylvestris	XP_009782232	1082	89.17%	0.0	100%	588
Ipomoea triloba	XP_031105450.1	989	85.39%	0.0	96%	591

4.2.3 Multiple Sequence Alignment

Multiple sequence analysis refers to the process or the result of aligning three or more biological sequences, such as protein, DNA or RNA, to identify evolutionary relationships and common patterns between genes. There are comparison of total 8 different protein sequence which represent the two different isoforms of G6PDH i.e. ChG6PDH and CyG6PDH from potato and the other *Solanaceae* family members such as tomato, capsicum and tobacco. Both the isoforms showed significant divergence at protein level. However, catalytic motifs were observed to be conserved among all the plant species.



```
Ch-StG6PDH:  MVTLYSSPSTNCSGAFSSYSNSS-IGFYNYHNNFPISSRKFGSHKI-SLQIN------AIRMQDGAVVAQPSKTQDETPL 72
Ch-SlG6PdH:  MVTLYSSPSTNCSGLFSSYSNSSSIGFYNYHNNLPISSRNFASHKI-SLQID------AVRMQDGAVVAPPSKTQDETPL 73
Ch-NtG6pDH:  MVTLYSSPSTHSSGAVASYSNSS-IGLYNYHHNKQIAVSSILSRKFGSLQINQKPFWNAVRMQDGAVATPPSKIENETPL 79
Ch-CaG6PDH:  MVTLYSPPST--GAVAFFYSNTS-IRFYNYHNNLTLFSRKLGSHNI-SLQINQTPFCNAVHIQDGSVAAPPSKNENETPL 75

                                               ⇩
             ~~~~~~~~~~~~~~~~ααα~~~~~~~~~~~~ββββββ~~~~ααααααααααααααα~~~~~~~~ββββββ~~~~~ααα
Cy-SlG6PDH:  --MAASWCIEKRGSIRNDSFRDNDNIPETGCLSIIVLGASGDLAKKKTFPALFNLYRQGFLQSNEVHIFGYARTKISDDD 158
Cy-StG6PDH:  --MAASWCIEKRGSIRNDSFRDNDNIPETGCLSIIVLGASGDLAKKKTFPALFNLYRQGFLQSNEVHIFGYARTKISDDD 158
Cy-NtG6PDH:  --MAASWCIEKRGSIRLDSFRDNDNIPETGCLSIIVLGASGDLAKKKTFPALFNLYRQGFLQSNEVHIFGYARTKISDDD 158
Cy-CaG6PDH:  --MAASWCIEKRGSIRKDSFRDNDNIPETGCLSIIVLGASGDLAKKKTFPALYNLYRQGFLQSNEVHIFGYARTKISDDE 158
Ch-StG6PDH:  KKLKDGILS-----KEQKHTFDFDSNKDKSTVSITVVGASGDLAKKKIFPALFALYYEGCLPEHFT-IFGYARSKMTDDE 154
Ch-SlG6PDH:  KKLKDGILS-----KEQKHIFDFDSNKDKSTVSITVVGASGDLAKKKIFPALFALYYEGCLPEHFT-IFGYARSKMTDDE 154
Ch-NtG6PDH:  KKLKNGILPVAPPKE.QKDTIDFDSNKAKSTVSITVVGASGDLAKKKIFPALFALYYEGCLPEHFT-IFGYARSKMTDAE 158
Ch-CaG6PDH:  --$aaswciekrgsir.dsfFDFGGNKDKSTVSITVVGASGDLAKKKIFPALFALYYESCLPEHFT-IFGYARSKMTDDE 156
                                     *           **  *********** ****    *    *    *     ** *  *

             ααααααααααα~~~~~~~~~αααα~α αααα~ββββ~~~~~~ααααααααααααααααααα~~~~~~~~~βββββ~~~ααααααα
Cy-SlG6PDH:  LRSRIRGYLS---RGKENEGEVLE-FLQLIKYVSGSYDSGEGFTLLFKAISEHEFSKNSTEGSSRRLFYFALPPSVYPSV 236
Cy-StG6PDH:  LRSRIRGYLS---QGKENEGEVSE-FLQLIKYVSGSYDSGEGFTLLDKAISEHEFSKNSTEGSSRRLFYFALPPSVYPSV 236
Cy-NtG6PDH:  LRGRIRGYLS---QGKENEEEVSE-FLQLIKYVSGSYDSGEGFSLLDKAIAEHEIAKNSTEGSSRRLFYFALPPSVYPSV 236
Cy-CaG6PDH:  LRSRIRGYLS---QEKENH-EVSE-FLQLIKYVSGSYDSGEGFTALDKAISEHEISKNSTEGSSRRLFYFALPPSVYPSV 236
Ch-StG6PDH:  LRNMVSKTLTCRIDKRENCGEKMEQFLERCFYHSGQYDSQENFAELDKKLKEHEAGR-----FSNRLFYLSIPPNIFINA 235
Ch-SlG6PDH:  LRNMVSKTLTCRIDKRENCGEKMEQFLQRCFYHSGQYDSQENFAELDKKLKEHEAGR-----FSNRLFYLSIPPNIFINA 235
Ch-NtG6PDH:  LRNMVSKTLTCRIDKRENCGEKMEQFLERCFYHSGQYDSQENFAELDKKLKEHEAGR-----FSNRLFYLSIPPNIFINA 235
Ch-CaG6PDH:  LRNMVSKTLTCRIDKRENCGEKMEQFLKRCFYHSGQYDSQENFAELDKKLKEHEAGR-----FSNRLFYLSIPPNIFINA 235
             **       *       **  *  * **     * ** *** *     ***     ***           * ***     **

                                                    G6PDH active site       ⇩ ▼
             αααααααααααααα~~~~~ββββββ~~~~~~~αααααααααααααα~~αααββ~~~~~~~~αααααααααααααα~~ααααααααα~~
Cy-SlG6PDH:  CRMIKSYCMNKSDLGGWTRIVVEKPFGKDLASSEQLSSQIGELFDEPQIYRIDHYLGKELVQNLLVLRFANRFFLPLWNR 320
Cy-StG6PDH:  CRMIKSYCMNKSDLGGWTRIVVEKPFGKDLASSEQLSSQIGELFDEPQIYRIDHYLGKELVQNLLVLRFANRFFLPLWNR 320
Cy-NtG6PDH:  CRMIKNYCMNKSDLGGWTRIVVEKPFGKDLASAEQLSSQIGELFDEPQIYRIDHYLGKELVQNLLVLRFANRFFLPLWNR 320
Cy-CaG6PDH:  CRMIKSYCMNKSDLGGWTRIVVEKPFGKDLASAEQLSSQIGELFEEPQIYRIDHYLGKELVQNLLVLRFANRFFLPLWNR 320
Ch-StG6PDH:  VRCAS---L-ASSAHGWTRVIVEKPFGRDSESSAALTGALKQYLKEDQIFRIDHYLGKELVENLSVLRFSNLIFEPLWSR 316
Ch-SlG6PDH:  VRCAS---LSVENLSVLRFSNLIFEPLWSRQYIRNVQFIFSEDFGTEGRGGYFDHYGIIRVENLSVLRFSNLIFEPLWSR 317
Ch-NtG6PDH:  VRCAS---LSASSAHGWTRVIVEKPFGRDSESSAALTRSLKQYLNEDQIFRIDHYLGKELVENLSVLRFSNLIFEPLWSR 317
Ch-CaG6PDH:  VRCAS---LSASSAHGWTRVIVEKPFGRDSESSAALTGALKQYLKEDQIFRIDHYLGKELVENLSVLRFSNLIFEPLWSR 317
                                                                          *    * ** ****  *  *

             ~~ββββββββββ~~~~~~~~ααα~~~~~~ααααααααααααααααααα~~~~~~~~~ααααααααααααααααα~~~~~~ααα~ββ~~
Cy-SlG6PDH:  DNIDNIQIVFREDFGTEGRGGYFDEYGIIRDIIQNHLLQVLCLVAMEKPVSQKPEHIRDEKVKVLQSMLPIKDEEVVLGQ 400
Cy-StG6PDH:  DNIDNIQIVFREDFGTEGRGGYFDEYGIIRDIIQNHLLQVLCLVAMEKPVSQKPEHIRDEKVKVLQSMLPIEDEEVVLGQ 400
Cy-NtG6PDH:  DNIDNIQIVFREDFGTEGRGGYFDEYGIIRDIIQNHLLQVLCLVAMEKPVSQKPEHIRDEKVKVLQSMLPIKDEEVVLGQ 400
Cy-CaG6PDH:  DNIDNIQIVFREDFGTEGRGGYFDEYGIIRDIIQNHLLQVLCLVAMEKPVSLKPEHIRDEKVKVLQSMLPIKDEEVVLGQ 400
Ch-StG6PDH:  QYIRNVQFIFSEDFGTEGRGGYFDHYGIIRDIMQNHLLQILALFAMETPVSLDAEDIRNEKVKVLRSMRPLQLDDVIVGQ 400
Ch-SlG6PDH:  QYIRNVQFIFSEDFGTEGRGGYFDHYGIIRDIMQNHLLQILALFAMETPVSLDAEDIRNEKVKVLRSMRPLQLDDVIVGQ 400
Ch-NtG6PDH:  QYIRNVQFIFSEDFGTEGRGGYFDHYGIIRDIMQNHLLQILALFAMETPVSLDAEDIRNEKVKVLRSMRPLQLDDVIIGQ 400
Ch-CaG6PDH:  QYIRNVQFIFSEDFGTEGRGGYFDHYGIIRDIMQNHLLQILALFAMETPVSLDAEDIRNEKVKVLRSMRPLQLDDVVVGQ 400
               *  *  *   *  *************** ******** ****** * * *** ***   * ** **             *  **

             ~~~~~~~~~~~~~αα~~~~~~~~~~~~~~ββββββββ~ααα~~~~βββββ~~~~~~~~ββββββββ~~~~~~~~~~~~~~
Cy-SlG6PDH:  YE----------GYKDDPTVPNNSNTPTFATMVLRIHNERWEGVPFIMKAGKALNSRKAEIRVQFKDVPGDIFRCQ---- 466
Cy-StG6PDH:  YE----------GYKDDPTVPNNSNTPTFATMVLRIHNERWEGVPFIMKAGKALNSRKAEIRVQFKDVPGDIFRCQ---- 466
Cy-NtG6PDH:  YE----------GYKDDPTVPDNSNTPTFATMVLRIHNERWEGVPFIMKAGKALNSRKAEIRVQFKDVPGDIFRCK---- 466
Cy-CaG6PDH:  YE----------GYKDDPTVPDNSNTPTFATMVLRIHNERWEGVPFIMKAGKALNSRKAEIRVQFKDVPGDIFKCQ---- 466
Ch-StG6PDH:  YKSHTKGGVNYPGYTDDKTVPKDSLTPTFAAAALFIDNARWDGVPFLMKAGKALHTRSAEIRVQFRHVPGNLYNKNFGSD 480
Ch-SlG6PDH:  YKSHTKGGVNYPGYTDDKTVPKDSLTPTFAAAALFIDNARWDGVPFLMKAGKALHTRSAEIRVQFRHVPGNLYNKNFGSD 480
Ch-NtG6PDH:  YKSHTKGGVNYPGYTDDKTVPKDSLTPTFAAAALFIDNARWDGVPFLMKAGKALHTRSAEIRVQFRHVPGNLYNKNFGSD 480
Ch-CaG6PDH:  YKSHTKGGVNYPGYTDDKTVPKDSLTPTFAAAALFIDNARWDGVPFLMKAGKALHTRSAEIRVQFRHVPGNLYNKNFGSD 480
             *           ********* * *****   *  *  *  ** ************  * *******  ***
```



```
             ~~~~~~ββββββ~~~~~βββββββ~~~~~~~βββββββ~~~~~αα~~~~~~~~αααααααααα~~~~~~~~~αααααααα
Cy-SlG6PDH:  -KQGRNEFVIRLQPSEAMYMKLTVKKPGLEMSTVQSELDLSYGQRYQGVVIPEAYERLILDTIRGDQQHFVRRDELKAAW 559
Cy-StG6PDH:  -KQGRNEFVIRLQPSEAMYMKLTVKKPGLEMSTVQSELDLSYGQRYQGVVIPEAYERLILDTIRGDQQHFVRRDELKAAW 559
Cy-NtG6PDH:  -KQGRNEFVIRLQPSEAMYMKLTVKKPGLEMSTVQSELDLSYGQRYQGVVIPEAYERLILDTIRGDQQHFVRRDELKAAW 559
Cy-CaG6PDH:  -KQGRNEFVIRLQPSEAMYMKLTVKKPGLEMSTVQSELDLSYGQRYQGVVIPEAYERLILDTIRGDQQHFVRRDELKAAW 559
Ch-StG6PDH:  LDQATNELVIRVQPDEAIYLKINNKVPGLGMRLDRSNLNLLYSARYSKE-IPDAYERLLLDAIEGERRLFIRSDELDAAW 559
Ch-SlG6PDH:  LDQATNELVIRVQPDEAIYLKINNKVPGLGMRLDRSNLNLLYSARYSKE-IPDAYERLLLDAIEGERRLFIRSDELDAAW 559
Ch-NtG6PDH:  LDQATNELVIRVQPDEAIYLKINNKVPGLGMRLDRSNLNLLYSARYSKE-IPDAYERLLLDAIEGERRLFIRSDELDAAW 559
Ch-CaG6PDH:  LDQATNELVIRVQPDEAIYLKINNKVPGLGMRLDRSNLNLLYSARYSKE-IPDAYERLLLDAIEGERRLFIRSDELDAAW 559
               *  ** *** ** **   *   * *** *    * * * *   **   ** ***** ** * *    * * *** ***

                                                          ▼
             αααααααααα~~~~~~~~~~~~~~~~αααααααααα~~~~~~~~~~~~~~~
Cy-SlG6PDH:  EIFTPLLHRIDNGEVKPIPYKPGSRGPAEADELLQHAGYVQTHGYIWIPPTL 612
Cy-StG6PDH:  EIFTPLLHRIDNGEVKPIPYKPGSRGPAEADELLQHAGYVQTHGYIWIPPTL 612
Cy-NtG6PDH:  EIFTPLLHRIDNGEVKPIPYKPGSRGPAEADELLQHAGYVQTHGYIWIPPTL 612
Cy-CaG6PDH:  EIFTPLLHRIDNGEVKPIPYKPGSRGPAEADELLQHAGYVQTHGYIWIPPTL 612
Ch-StG6PDH:  SLFTPVLKDLEHKKIVPESYPYGSRGPIGAHYLA--ARYKVRWGDLA----- 605
Ch-SlG6PDH:  SLFTPVLKDLEHKKIVPESYPYGSRGPIGAHYLA--ARYKVRWGDLA----- 605
Ch-NtG6PDH:  SLFTPVLKDLEHKKIVPESYPYGSRGPIGAHYLA--ARYKVRWGDLA----- 605
Ch-CaG6PDH:  SLFTPVLKDLEHKKIVPESYPYGSRGPIGAHYLA--ARYKVRWGDLA----- 605
               *** *            *  *****  *  *   *       *
```

Fig. 4.1 Multiple sequence alignment was done using MultAlin software. Gaps arisen during alignment was indicated by dashes. Asterisks '*' represent the conserved amino acids. The G6PDH active sites (251-257) was highlighted in yellow colour. The C termial domain is shown in downward traingle▼ and the NAD binding domain is highlighted in arrow⇩. 'α', 'β' and '~'denotes the tendency of the individual amino acid of Ch-StG6PDH(NP_001275038.1),Ch-SlG6PDH(XP_004243567.1),Ch-CaG6PDH(XP_016580465.1), Ch-NtG6PDH (NP_001312522.1)and Cy-StG6PDH(NP_001275397.1), Cy-SlG6PDH(XP_004231802.1),Cy-CaG6PDH (XP_016561682.1), Cy-NtG6PDH(XP_016432237.1).

4.2.4 Motif Scan

A motif is a short, conserved sequence pattern that is often associated with a specific biological function or structural feature. It can be represented as a string or regular expression and are typically identified through various methods, such as multiple sequence alignment and pattern matching. It provide information into the functional and evolutionary aspects of biological sequences. The important protein motifs that are identified by using the motif server (https://myhits.sib.swiss) of the two different isoforms of G6PDH (CyG6PDH and ChG6PDH) are discussed. All the important protein motifs play an important role in the catalytic function.

A) G6PDH Active Site: The active site of G6PDH is located in the enzyme structure which is generally found as a dimer of two identical monomers. It is responsible for catalysing the chemical reaction involves in the enzyme function. It catalyses the conversion of D-glucose 6-

phosphate to 6-phospho-D-glucono-1,5-lactone, while reducing $NADP^+$ to NADPH and H^+. It is essential for PPP and cellular redox balance.

B) C-Terminal Domain: The C-terminal domain plays a crucial role in plant acclimatization to various stresses. It involves in the removal of reactive oxygen species (ROS) and contributes to plant's antioxidant defensive system.

C) NAD Binding Domain: The NAD binding domain in G6PDH of plants plays a crucial role in the enzyme's catalytic function, coenzyme specificity, and structural stability. It facilitates the binding and transfer of electrons between the substrate (glucose-6-phosphate) and the coenzyme ($NADP^+$), leading to the production of NADPH and H^+. The NAD binding domain is essential for the role of G6PDH in the pentose phosphate pathway and the generation of NADPH for various metabolic processes in plants.

Table 8 Motifs of CyG6PDH position

Site	G6PDH active site	C terminal domain	NAD binding domain
Species			
Solanum tuberosum	207-213	219-503	33-217
Solanum lycopersicum	207-213	219-503	33-217
Capsicum annum	206-212	218-502	33-216
Nicotiana tabacum	206-212	218-502	33-216

Table 9 Motifs of ChG6PDH position

Site	G6PDH active site	C terminal domain	NAD binding domain
Species			
Solanum tuberosum	263-269	276-573	94-273
Solanum lycopersicum	263-269	272-569	94-273
Capsicum annum	259-265	287-584	90-269
Nicotiana tabacum	274-280	287-584	105-284

4.2.5 Genome Wide Study

A close inspection of potato genome database was done to identify the G6PDH family members. A total of 4 G6PDH genes encoding 13 multiple forms. Protparm tool was used to predict various physicochemical properties such as molecular weight, exons, aliphatic index, instability index and GRAVY (Table 10). The gene length of this family members ranged from 1865 bp-3631bp. Molecular weight of the proteins encoded by this family members varied from 25 to 66kDa. A protein is considered to be stable if its instability index is less than 40; if it is greater than 40, the protein may be unstable. The instability index indices of G6PDH proteins ranged from 36-47 indicating the presence of both stable and unstable proteins. Greater the aliphatic index, greater is the thermostability of the protein. All proteins with aliphatic index more than 70 are thermally stable. It is observed from the table that G6PDH are thermostable. The GRAVY score is determined by adding the hydropathy values for each amino acid in a protein and dividing the result by the total number of residues in the protein. Proteins with a negative GRAVY value are non-polar, while those with a positive value are polar. Thus all the proteins in G6PDH are non-polar with negative GRAVY. The low GRAVY index indicates the hydrophilic nature of a protein. Proteins with a negative GRAVY value are non-polar, while those with a positive value are polar. Thus all the proteins in G6PDH are non-polar with negative GRAVY. The low GRAVY index indicates the hydrophilic nature of a protein.

Table 10 Description of G6PDH members of potato

Gene id	Transcript id	Bp	aa	Exon	MW	pI	Gravy	Aliphatic index	Instability Index
PGSC0003DMG400002750	PGSC0003DMT400007105	2469	596aa	11	66.64575	5.83	-0.245	94.85	47.31
	PGSC0003DMT400007102	2987	230aa	8	25.95237	5.26	-0.457	95.74	40.49
	PGSC0003DMT400007101	2985	230aa	8	25.95237	5.26	-0.457	95.74	40.49
	PGSC0003DMT400007106	2868	230aa	8	25.95237	5.26	-0.457	95.74	40.49
	PGSC0003DMT400007103	2816	281aa	10	31.72501	5.53	-0.363	96.05	44.87
	PGSC0003DMT400007104	2374	405aa	10	45.92321	5.64	-0.308	95.53	44.87
PGSC0003DMG400010802	PGSC0003DMT400028015	2492	582aa	10	66.21330	8.57	-0.400	83.30	39.79
	PGSC0003DMT400028013	2577	239aa	9	27.07691	8.58	-0.434	91.42	36.52
	PGSC0003DMT400028014	2577	480aa	9	28.57074	9.05	-0.295	83.31	40.51
	PGSC0003DMT400028016	2112	252aa	6	28.57074	8.58	-0.355	95.24	36.91
PGSC0003DMG400017394	PGSC0003DMT400044818	2153	577aa	10	65.72091	6.88	-0.338	85.34	39.84
	PGSC0003DMT400044819	3631	255aa	9	58.48257	5.97	-0.433	85.44	47.53
PGSC0003DMG400020269	PGSC0003DMT400052250	1865	511aa	16	58482.57	5.97	-0.433	85.44	47.53

4.2.6 Expression Pattern of G6PDH

A heat map is a 2-D data visualization technique that represents the magnitude of individual values within a dataset as a colour. In *in silico* studies, heat maps are commonly used to visualize gene expression data. The approach was made to predict the organ specific expression patterns of four StG6PDH shown in the (Fig 4.2). It displays data in a grid in which each row represents a gene and each column represent a sample. The colour and intensity of the boxes used to represent changes of gene expression pattern in Transcripts per million (TPM). This observes that the G6PDH plays a very important role in the development of the plant growth. It is used to identifying genes that are commonly regulated or biological signatures associated with a particular condition. The heat map clearly showed the presence of G6PDH in present in all the organs of the of potato, but it is highly expressed in young tuber with 508TPM, followed by stem with 428TPM. The lowest expression (242 TPM) was observed in mature tubers.

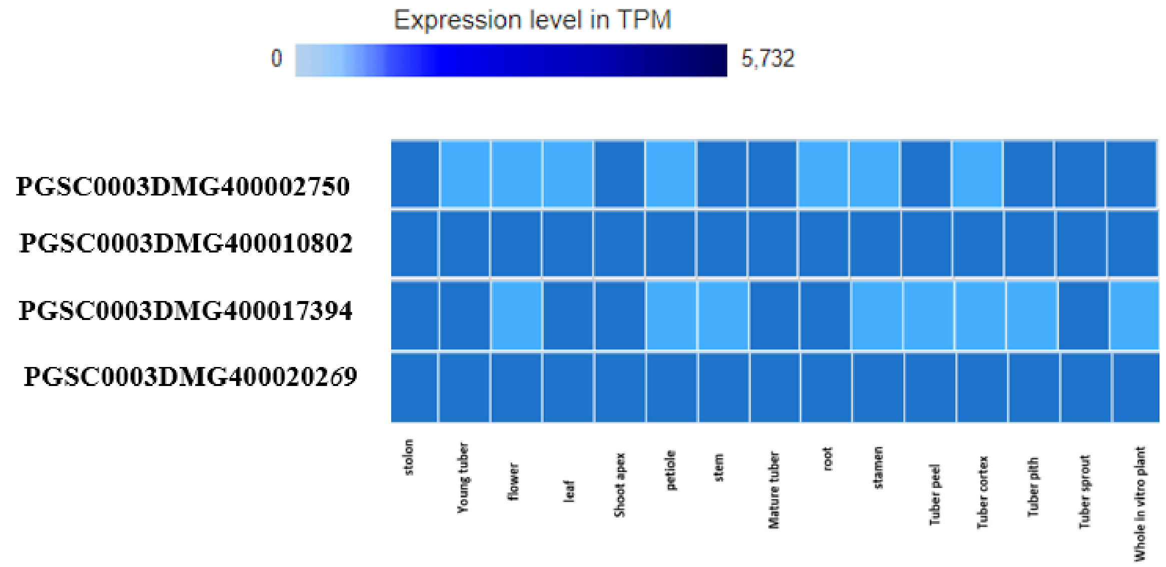


Fig. 4.2 The heatmap of G6PDH genes retrieved from motifs data.

4.2.7 Phylogenetic Tree

A phylogenetic tree is a branching diagram which represent the evolutionary relationship the evolutionary relationship among various biological species based on their similarities and difference in their physical and genetic characteristics. Phylogenetic tree was constructed using G6PDH protein sequences from potato, tomato and Arabidopsis as shown in Fig. 4.3.

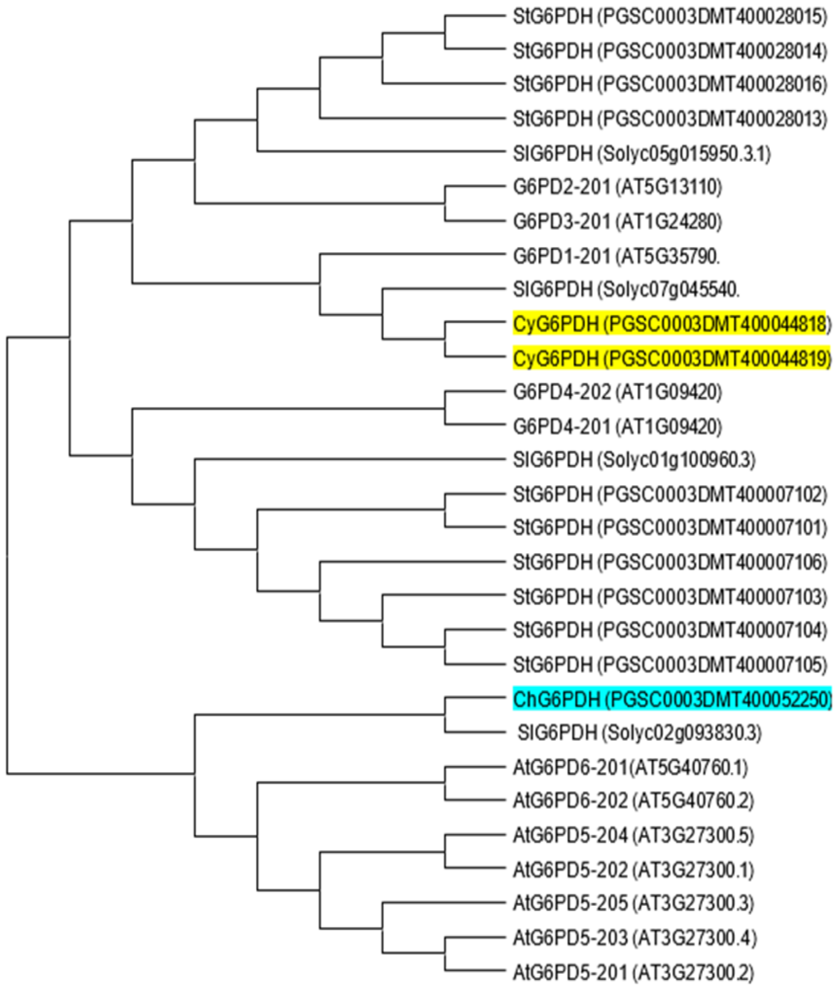


Fig. 4.3 Phylogenetic tree representing 30 G6PDH protein sequence from three different plant species: 14 from potato (StG6PDH), 5 from tomato (St G6PDH) and 11 from Arabidopsis (At G6PDH).

4.2.8 Prediction of 3-D Protein Structure

A polypeptide's overall three-dimensional structure is referred to as its tertiary structure having a protein's entire 3D structure is necessary to identify its biological activity. I-TASSER was utilised as the server to simulate the whole G6PDH sequence. It is a method for modelling

hierarchical proteins that relies on numerous threading alignments and an iterative application of the TASSER (Threading Assembly Refinement) programme (Zhang Y 2008). In order to find the best templates for modelling G6PDH, the amino acid FASTA sequences from NCBI were downloaded and uploaded into the programme. However, it was later discovered that no models could fully meet the query coverage. The best model was therefore predicted using the I-TASSER service (http://zhang.bioin-formatics.ku.edu/I-TASSER). The whole approach for the I-TASSER algorithm has been given in depth in the works Wu, Skolnick, and Zhang (2007; 2009). I-TASSER constructed a full-length model using restraints from templates that are recognised by several threading programmes. The complete protein sequences of the chloroplastic and cytoplasmic versions of G6PDH were modelled using the server I-TASSER. The models of *Solanum tuberosum*, *Capsicum annum*, *Nicotiana tabacum*, and *Solanum lycopersicum*. were made by using I-TASSER server. It is a hierarchical technique that predicts protein structure automatically and annotates activities that are dependent on protein structure.

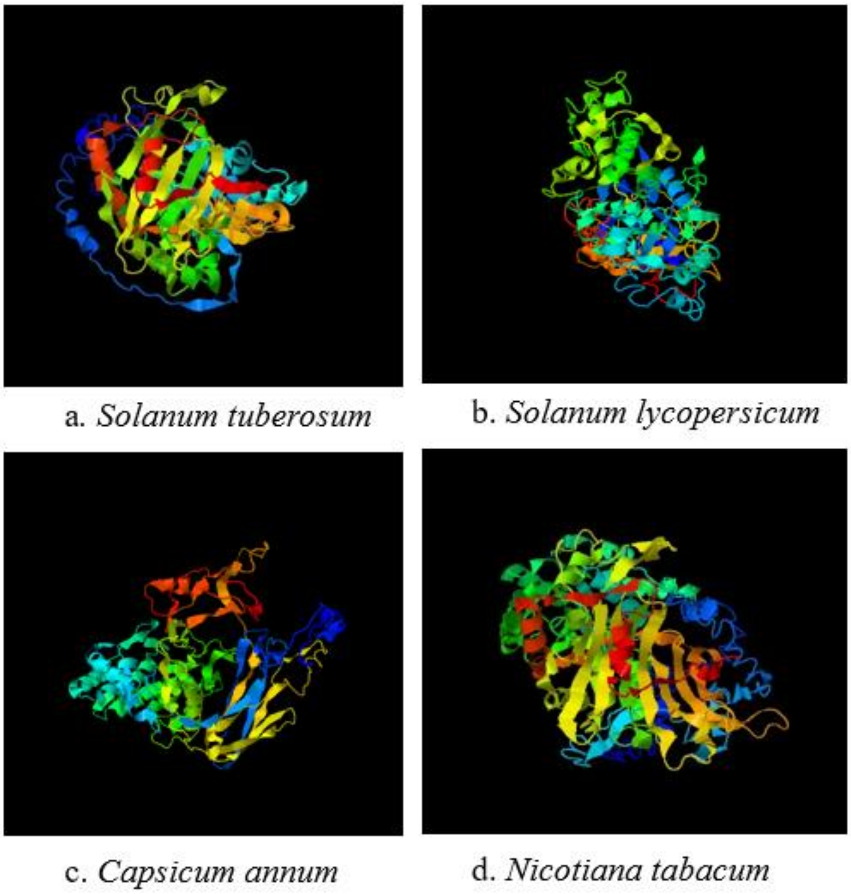


Fig. 4.4 3-D protein structure of 4 different species of *Solanaceae* family of Chloroplast G6PDH.

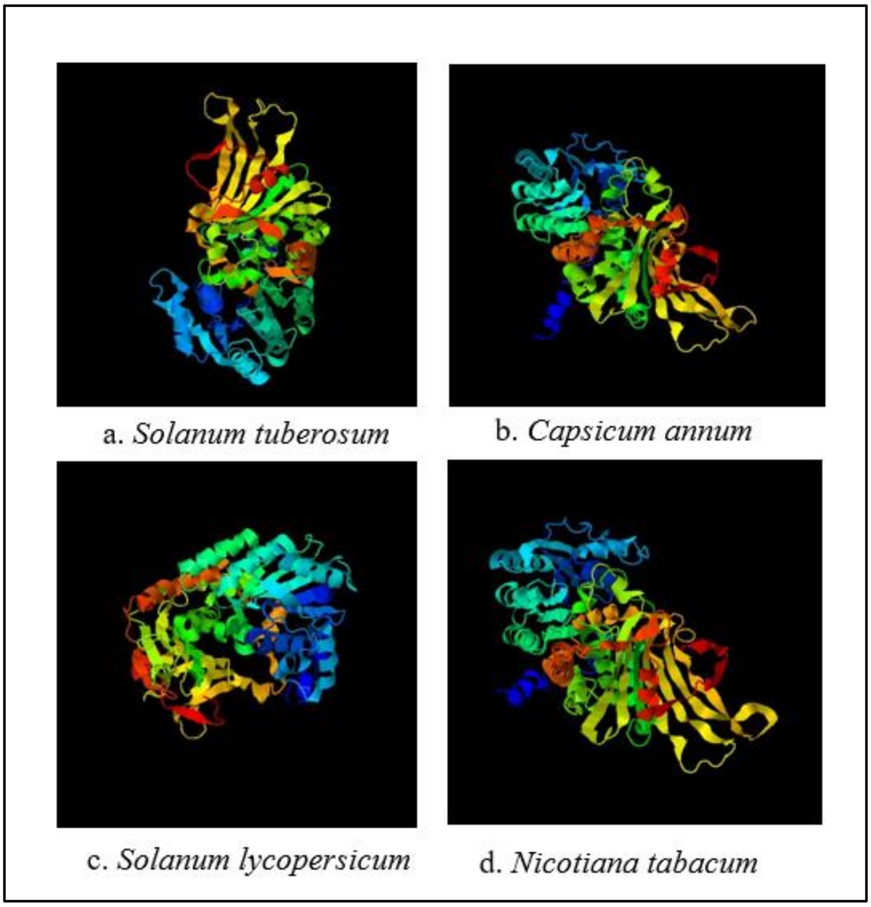


Fig. 4.5 3-D protein structure of 4 different species of *Solanaceae* family of Cytoplasmic G6PDH isoform.

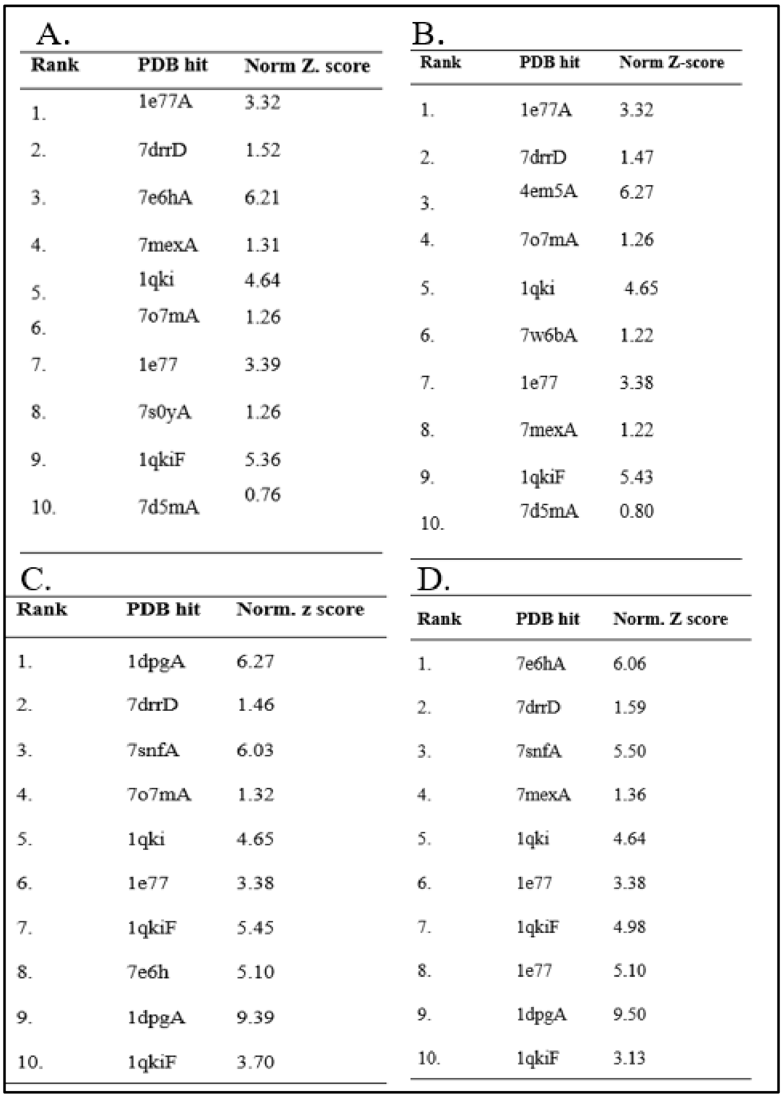
- **Confidence Score (C-score)**

The C-score is a confidence score which used by I-TASSER to estimate the quality of predicted models. The C-score is usually in the range of [-93, -18], with a higher C-score indicating a more confident model and vice versa. Greater the value of C-score better is the model.

- **Quality Estimation**

Quality estimation of the 3-D structure is necessary to determine the applications of the protein. This is done by studying the Z-score values provided by I-TASSER. The Z-score of the alignment is commonly used to judge the quality of threading alignment. It is the difference in standard deviation between the raw and average scores. The energy score's standard deviation is determined as a percentage of the statistical average of all alignments. The superior alignment is represented by a normalized Z-score <1.

Table 11 Normalized Z scores for models of ChG6PDH isoform A. *Solanum tuberosum,* B. *Solanum lycopersicum* C. *Capsicum annum* and D. *Nicotiana attenuate*



A.

Rank	PDB hit	Norm Z. score
1.	1e77A	3.32
2.	7drrD	1.52
3.	7e6hA	6.21
4.	7mexA	1.31
5.	1qki	4.64
6.	7o7mA	1.26
7.	1e77	3.39
8.	7s0yA	1.26
9.	1qkiF	5.36
10.	7d5mA	0.76

B.

Rank	PDB hit	Norm Z-score
1.	1e77A	3.32
2.	7drrD	1.47
3.	4em5A	6.27
4.	7o7mA	1.26
5.	1qki	4.65
6.	7w6bA	1.22
7.	1e77	3.38
8.	7mexA	1.22
9.	1qkiF	5.43
10.	7d5mA	0.80

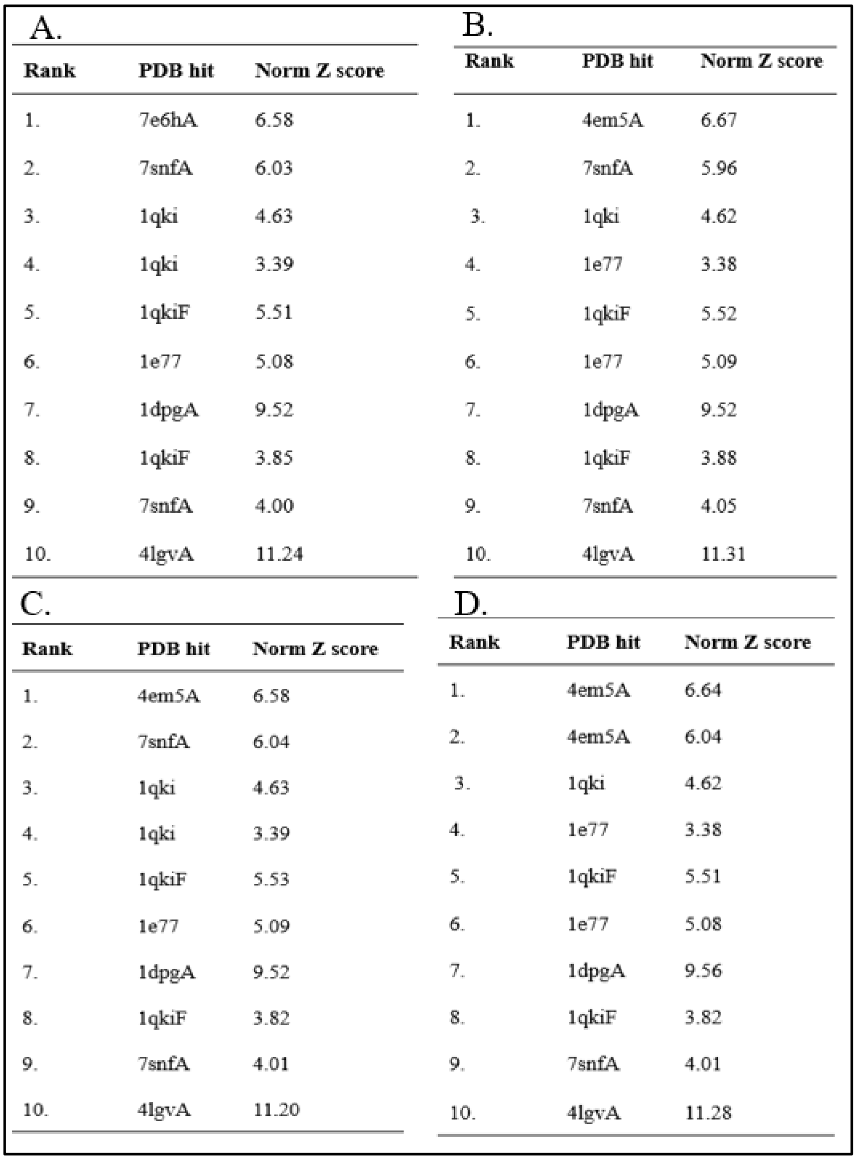
C.

Rank	PDB hit	Norm. z score
1.	1dpgA	6.27
2.	7drrD	1.46
3.	7snfA	6.03
4.	7o7mA	1.32
5.	1qki	4.65
6.	1e77	3.38
7.	1qkiF	5.45
8.	7e6h	5.10
9.	1dpgA	9.39
10.	1qkiF	3.70

D.

Rank	PDB hit	Norm. Z score
1.	7e6hA	6.06
2.	7drrD	1.59
3.	7snfA	5.50
4.	7mexA	1.36
5.	1qki	4.64
6.	1e77	3.38
7.	1qkiF	4.98
8.	1e77	5.10
9.	1dpgA	9.50
10.	1qkiF	3.13

Table 12 Normalized Z scores for models of CyG6PDH isoform A. *Solanum tuberosum,* B. *Solanum lycopersicum* C. *Capsicum annum* and D. *Nicotiana attenuate*



A.

Rank	PDB hit	Norm Z score
1.	7e6hA	6.58
2.	7snfA	6.03
3.	1qki	4.63
4.	1qki	3.39
5.	1qkiF	5.51
6.	1e77	5.08
7.	1dpgA	9.52
8.	1qkiF	3.85
9.	7snfA	4.00
10.	4lgvA	11.24

B.

Rank	PDB hit	Norm Z score
1.	4em5A	6.67
2.	7snfA	5.96
3.	1qki	4.62
4.	1e77	3.38
5.	1qkiF	5.52
6.	1e77	5.09
7.	1dpgA	9.52
8.	1qkiF	3.88
9.	7snfA	4.05
10.	4lgvA	11.31

C.

Rank	PDB hit	Norm Z score
1.	4em5A	6.58
2.	7snfA	6.04
3.	1qki	4.63
4.	1qki	3.39
5.	1qkiF	5.53
6.	1e77	5.09
7.	1dpgA	9.52
8.	1qkiF	3.82
9.	7snfA	4.01
10.	4lgvA	11.20

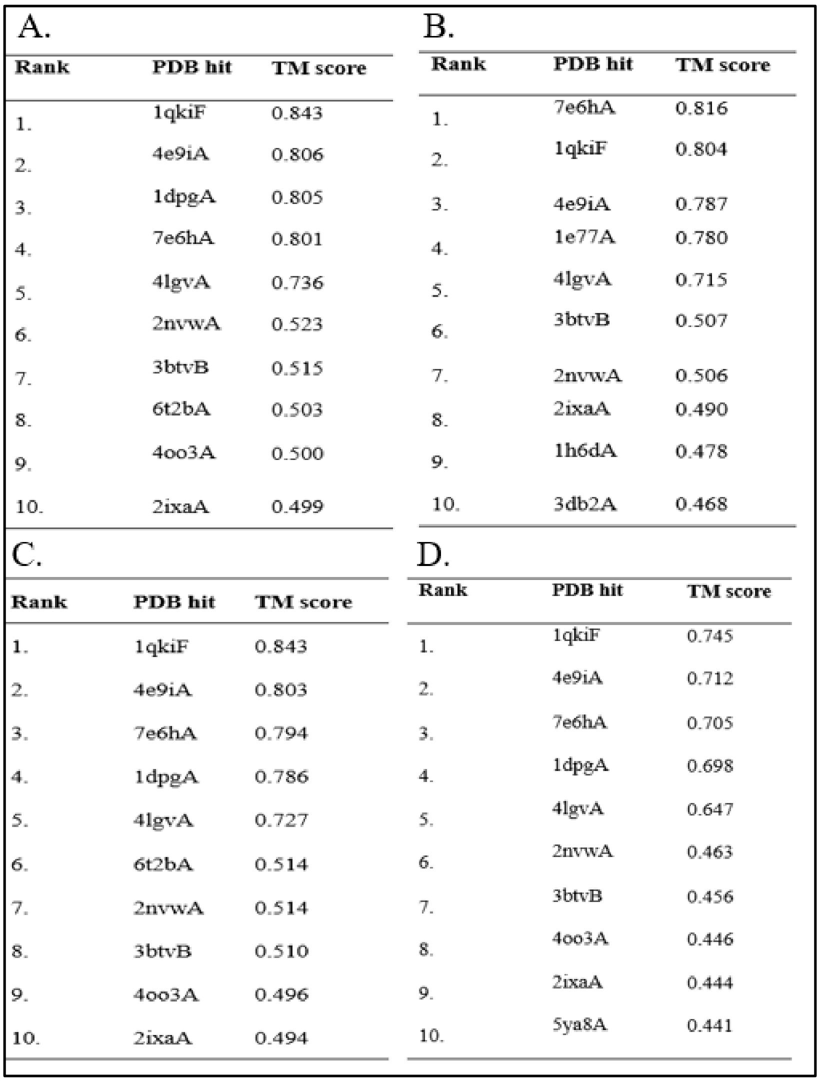
D.

Rank	PDB hit	Norm Z score
1.	4em5A	6.64
2.	4em5A	6.04
3.	1qki	4.62
4.	1e77	3.38
5.	1qkiF	5.51
6.	1e77	5.08
7.	1dpgA	9.56
8.	1qkiF	3.82
9.	7snfA	4.01
10.	4lgvA	11.28

- **Structure Similarity**

The TM-score can be used to compare the structural or topological similarity of two structures. The top PDB proteins with the greatest TM-score and the most structural similarity to the predicted I-TASSER models are listed by the Table 13,14.

Table 13 Models of TM score of ChG6PDH model of A. *Solanum tuberosum,* B. *Solanum lycopersicum* C. *Capsicum annum* and D. *Nicotiana attenuate*



A.

Rank	PDB hit	TM score
1.	1qkiF	0.843
2.	4e9iA	0.806
3.	1dpgA	0.805
4.	7e6hA	0.801
5.	4lgvA	0.736
6.	2nvwA	0.523
7.	3btvB	0.515
8.	6t2bA	0.503
9.	4oo3A	0.500
10.	2ixaA	0.499

B.

Rank	PDB hit	TM score
1.	7e6hA	0.816
2.	1qkiF	0.804
3.	4e9iA	0.787
4.	1e77A	0.780
5.	4lgvA	0.715
6.	3btvB	0.507
7.	2nvwA	0.506
8.	2ixaA	0.490
9.	1h6dA	0.478
10.	3db2A	0.468

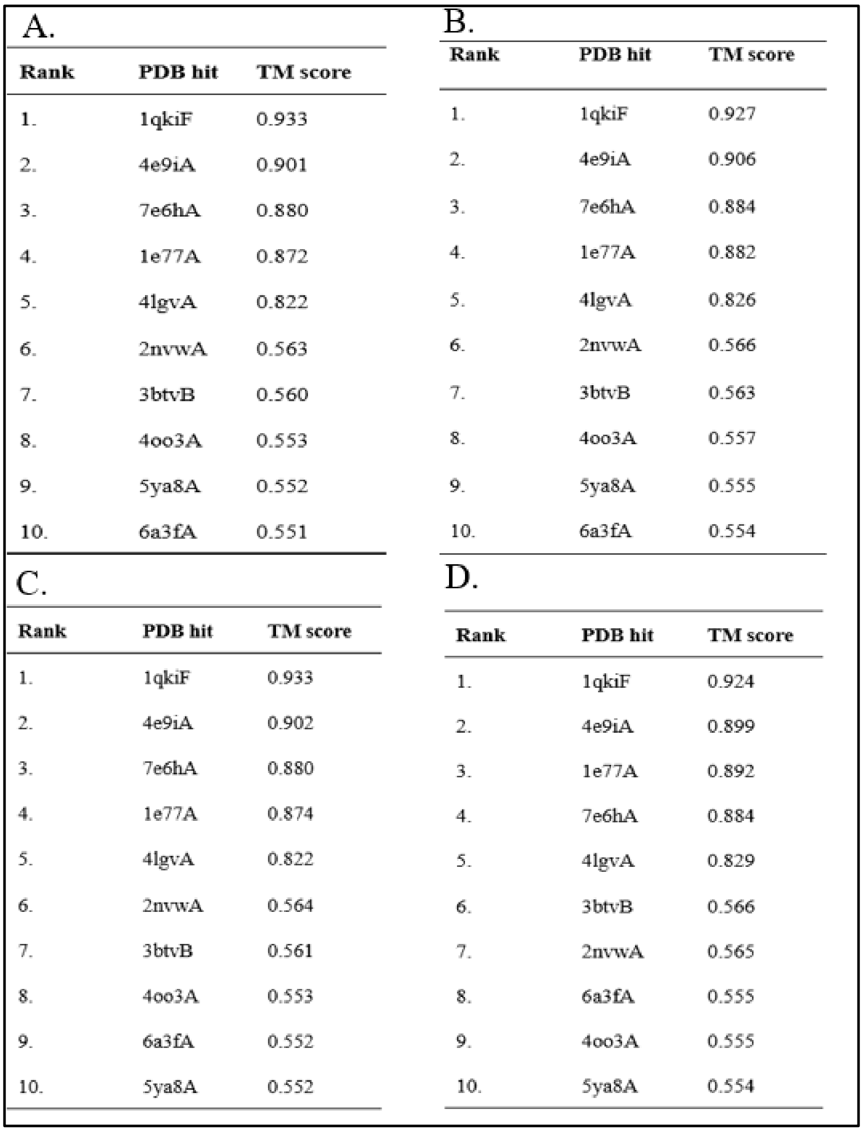
C.

Rank	PDB hit	TM score
1.	1qkiF	0.843
2.	4e9iA	0.803
3.	7e6hA	0.794
4.	1dpgA	0.786
5.	4lgvA	0.727
6.	6t2bA	0.514
7.	2nvwA	0.514
8.	3btvB	0.510
9.	4oo3A	0.496
10.	2ixaA	0.494

D.

Rank	PDB hit	TM score
1.	1qkiF	0.745
2.	4e9iA	0.712
3.	7e6hA	0.705
4.	1dpgA	0.698
5.	4lgvA	0.647
6.	2nvwA	0.463
7.	3btvB	0.456
8.	4oo3A	0.446
9.	2ixaA	0.444
10.	5ya8A	0.441

Table 14 Models of TM score of Cy-G6PDH model of A. *Solanum tuberosum,* B. *Solanum lycopersicum* C. *Capsicum annum* and D. *Nicotiana attenuate*



A.

Rank	PDB hit	TM score
1.	1qkiF	0.933
2.	4e9iA	0.901
3.	7e6hA	0.880
4.	1e77A	0.872
5.	4lgvA	0.822
6.	2nvwA	0.563
7.	3btvB	0.560
8.	4oo3A	0.553
9.	5ya8A	0.552
10.	6a3fA	0.551

B.

Rank	PDB hit	TM score
1.	1qkiF	0.927
2.	4e9iA	0.906
3.	7e6hA	0.884
4.	1e77A	0.882
5.	4lgvA	0.826
6.	2nvwA	0.566
7.	3btvB	0.563
8.	4oo3A	0.557
9.	5ya8A	0.555
10.	6a3fA	0.554

C.

Rank	PDB hit	TM score
1.	1qkiF	0.933
2.	4e9iA	0.902
3.	7e6hA	0.880
4.	1e77A	0.874
5.	4lgvA	0.822
6.	2nvwA	0.564
7.	3btvB	0.561
8.	4oo3A	0.553
9.	6a3fA	0.552
10.	5ya8A	0.552

D.

Rank	PDB hit	TM score
1.	1qkiF	0.924
2.	4e9iA	0.899
3.	1e77A	0.892
4.	7e6hA	0.884
5.	4lgvA	0.829
6.	3btvB	0.566
7.	2nvwA	0.565
8.	6a3fA	0.555
9.	4oo3A	0.555
10.	5ya8A	0.554

All potential PDB structures that are near to the target proteins are listed in Tables 4.6 A, B,C, and D. A valid topological model is suggested by a TM-score of >;0.5, whilst a TM-score of 0.17 denotes random similarity. These cutoffs are unaffected by the length of the protein. The average distance between all residue pairs in two structures is known as the RMSD which prone to local inaccuracy. This inaccuracy is solved by TM value. if the global topology is accurate, a local error will lead to a high RMSD value. The TM-score, on the other hand, makes the outcome less vulnerable to local modelling errors because the small distance is weighted more heavily than the vast distance.

4.2.9 Ramachandran Plot

The Ramachandran principle is used to measure angles in amino acids (proteins) as well as in other kind of molecules. This plot helps to determine the secondary structure of proteins. The conformation of the backbone of any peptide is governed by the values of two torsional angles which are Psi (ψ) and Phi (φ) angles. The protein sequence was uploaded on the PROCHECK server and the Ramachandran plots as shown in Fig. 4.6 and Fig.4.7 were obtained. The regions that are shown in red colour i.e., A, B and L contain most favoured residues. Residues in additional and generously allowed regions are also shown.

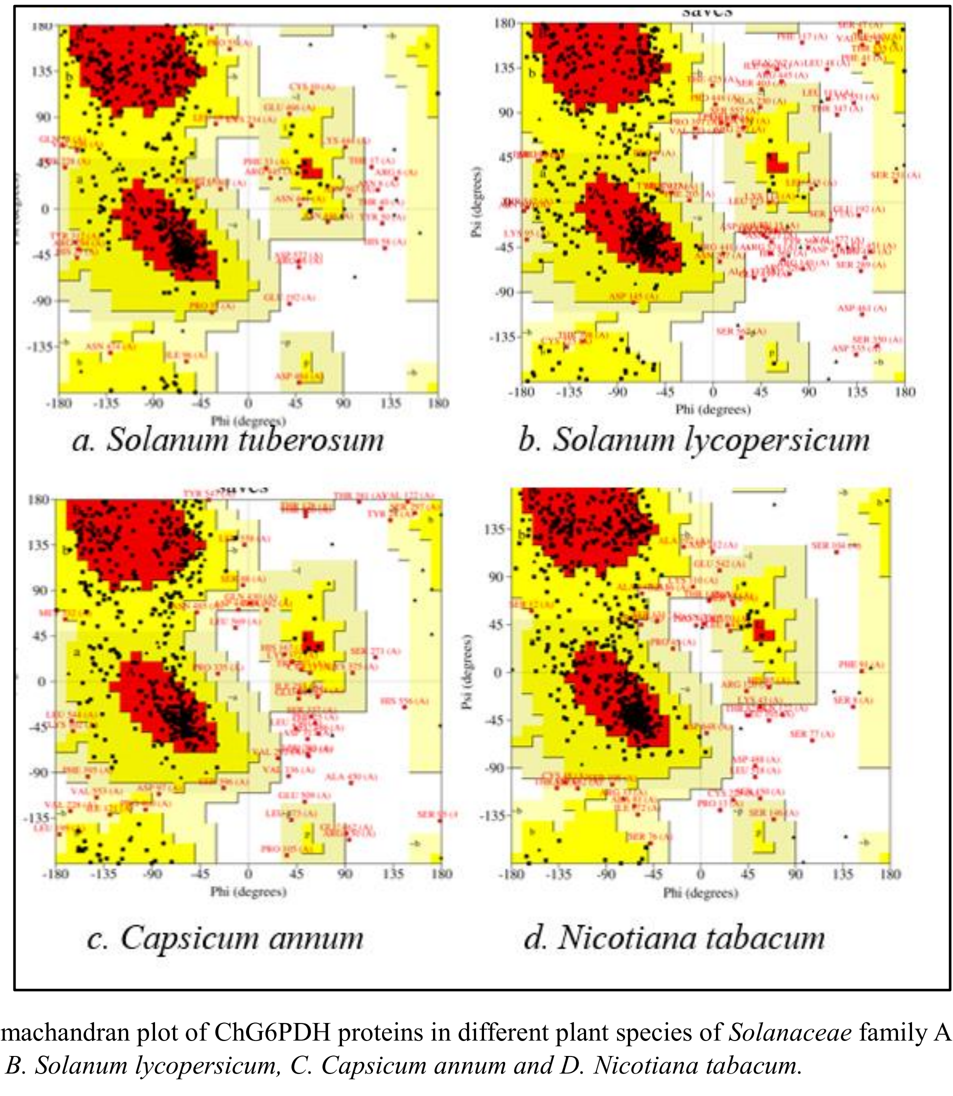
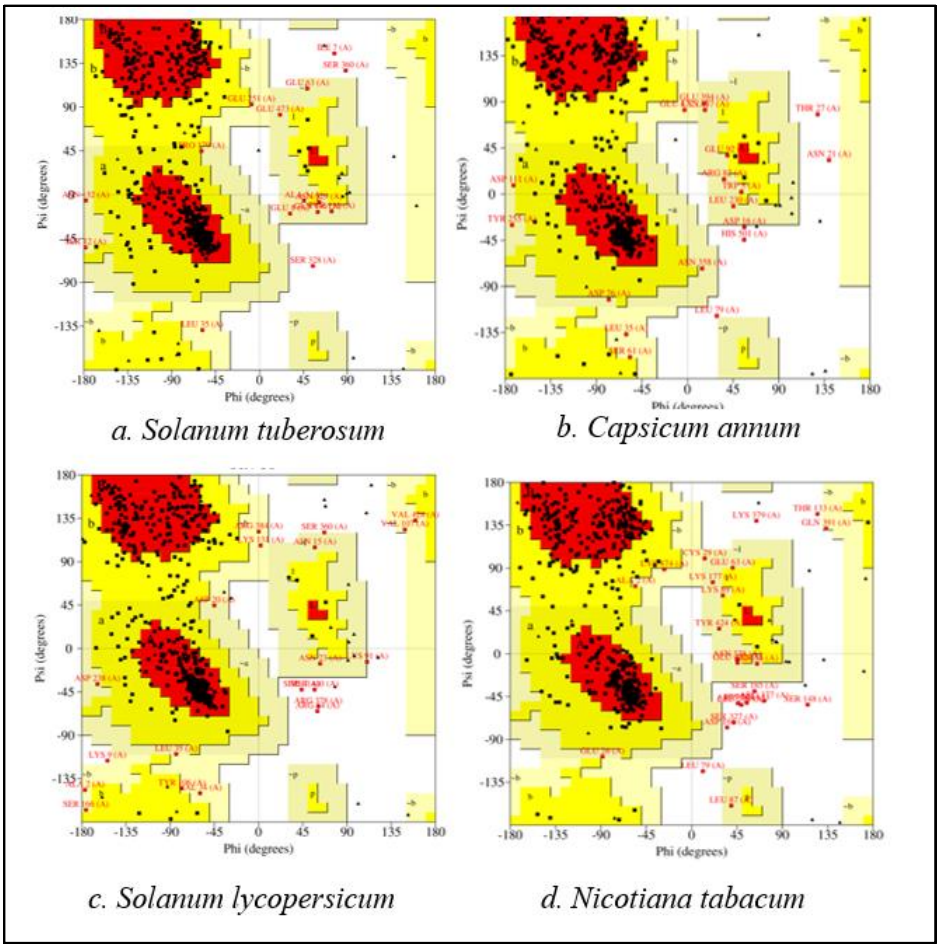


Fig. 4.6 Ramachandran plot of ChG6PDH proteins in different plant species of *Solanaceae* family A. *Solanum tuberosum, B. Solanum lycopersicum, C. Capsicum annum and D. Nicotiana tabacum.*



a. *Solanum tuberosum*

b. *Capsicum annum*

c. *Solanum lycopersicum*

d. *Nicotiana tabacum*

Fig. 4.7 Ramachandran plot of CyG6PDH proteins in different plant species of *Solanaceae* family A. *Solanum tuberosum, B. Solanum lycopersicum, C. Capsicum annum and D. Nicotiana tabacum***.**

Table 15 Statistics of plot of different plant species of ChG6PDH

		No of residues		
	Solanum tuberosum	***Solanum lycopersicum***	***Capsicum annum***	***Nicotiana tabacum***
In most favoured regions [A,B,L]	355(69.6%)	286(54.9%)	290(56.8%)	385(66.2%)
In additional allowed regions [a,b,l,p]	126(24.7%)	173(33.2%)	172(33.7%)	159(27.3%)

Iin generously allowed regions [~a,~b,~l,~p]	20(3.9%)	31(6.0%)	30(5.9%)	24(4.1%)
In disallowed regions	9(1.8%)	31(6.0%)	19(5.7%)	14(2.4%)
Non-glycine and non-proline residues	510(100%)	521(100%)	511(100%)	582(100%)
End-residues	2	2	2	1
Glycine residues	36	25	34	40
Proline residues	29	25	26	30
Total residues	577	583	573	653

Table 16 Statistics of plot of different plant species of CyG6PDH

	No. of Residues			
	Solanum tuberosum	***Solanum lycopersicum***	***Capsicum annum***	***Nicotiana tabacum***
In most favoured regions [A,B,L]	345(77.5%)	345(77.5%)	351978.5%)	336(75.7%)
In additional allowed regions [a,b,l,p]	86(19.3%)	80(18.0%)	78(17.4%)	84(18.9%)
In disallowed regions	11(2.5%)	13(2.9%)	139(2.6%)	13(2.9%)
In disallowed regions	3(0.7%)	7(1.6%)	59(1.1%)	119(2.5%)

Non-glycine and non-proline residues	445(100%)	445(100%)	44.7(100%)	444(100%)
End-residues	2	2	2	2
Glycine residues	38	38	25	38
Proline residues	26	26	26	26
Total residues	511	511	510	510

4.2.10 Protein-Protein Interaction

String map are commonly used in *in-silico* work to map the protein and the gene identifiers to their corresponding names and functional annotations. The string database performs identifier mapping on user input and displays a network covering all the mapped proteins and their interconnections. It is a useful identifier tool in *in-silico* work. Possible interactions revealed by STRING database for G6PDH are described in Fig.4.8 and Table 17.

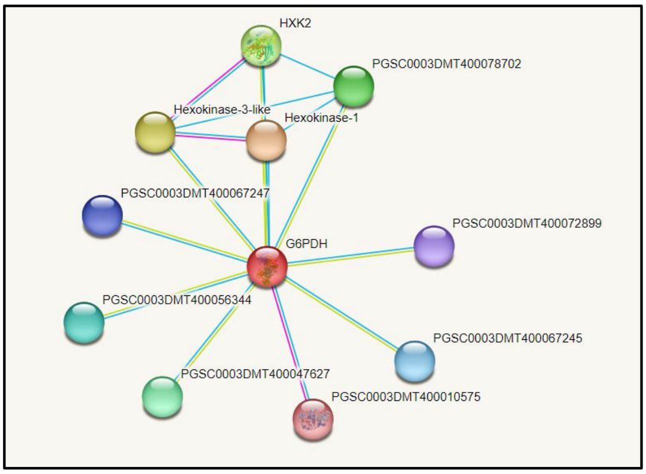


Fig 4.8 Possible protein-protein interaction between G6PDH and other various protein within potato genome sequencing consortium (PGSC) network.

Table 17 String network and their positions

S No.	Potato genome seq. consortium network (PGSC)	Position
1	PGSC0003DMT400072899	Probable 6-phosphogluconolactonase 1; Belongs to the glucosamine/galactosamine-6-phosphate isomerase family. 6-phosphogluconolactonase subfamily
2	PGSC0003DMT400067245	Probable 6-phosphogluconolactonase 4, chloroplastic; Uncharacterized protein; 6-phosphogluconolactonase
3	PGSC0003DMT400010575	Glyceraldehyde-3-phosphate dehydrogenase, cytosolic; Glyceraldehyde-3-phosphate dehydrogenase C subunit; Belongs to the glyceraldehyde-3-phosphate dehydrogenase family
4	PGSC0003DMT400047627	Belongs to the glucosamine/galactosamine-6-phosphate isomerase family
5	PGSC0003DMT400056344	Belongs to the glucosamine/galactosamine-6-phosphate isomerase family
6	PGSC0003DMT400067247	Probable 6-phosphogluconolactonase 4, chloroplastic; Belongs to the glucosamine/galactosamine-6-phosphate isomerase family
7.	PGSC0003DMT400078702	Show homologs among STRING organism
8.	Hexokinase-1	Show homologs among STRING organism
9.	Hexokinase-3-like	Shows protein sequence and homologs among STRING organism
10.	HXK2	May be involved in the phosphorylation of glucose during the export from plastids to cytosol. Seems neither to be involved in cell sugar sensing nor in carbohydrate metabolism in tuber

4.3 Biochemical Analysis of G6PDH Enzyme from Different Potato Cultivars

4.3.1 Plant Development and Harvesting of Potato Organs

Pathogen free germplasm of Indian potato cultivars namely Kufri-Chipsona-1 (KC-1), Kufri-Chipsona-2 (KC2), Kufri Pukhraj (PR), Kufri Jyoti (KJ), and Desiree (DE) procured from Central Potato Research Institute (CPRI) are routinely maintained on MS media under proper laboratory conditions (25–27 °C, ~ 70% relative humidity under 16-h photoperiod with a light intensity of 40 − 42 μmol/m2/s spectral flux photon of photosynthetically active i.e. 460 − 700 nm radiations with 2.5% sucrose at the interval of 4–5 weeks). These micropropagated plants were then hardened, acclimatized and transferred to the fields for a period of 15-16 weeks (mid of November-March). Different potato organs were harvested at various stages of development (Fig 4.9). The organs were then frozen in liquid nitrogen and stored at − 80 °C for further biochemical studies. Chemicals and enzymes were procured from Sigma-Aldrich Pvt. Ltd, Genei, SRL and HiMedia Laboratories, India.

Fig. 4.9 Process of harvesting and maintaining the Potato samples in -80ºC

4.3.2 Estimation of protein in different potato cultivars

Different organs of potato cultivars i.e. leaves, stems and tubers were crushed in liquid nitrogen and homogenized in Tris HCl buffer (pH 7.6) as mentioned in Materials and methods. Standard curve for proteins was made using BSA (Bovine serum albumin) at different concentrations and absorbance was taken at 750 nm Specific activity (per g protein) were calculated. Absorbance at various concentrations is tabulate below (Table 18)

Table 18 BSA at different concentration and absorbance 750 nm.

Concentration (µg)	**A 750 Value**
Blank	0
200	0.427
300	0.634
400	0.797
500	0.879
600	1.08
700	1.201
800	1.255
900	1.399

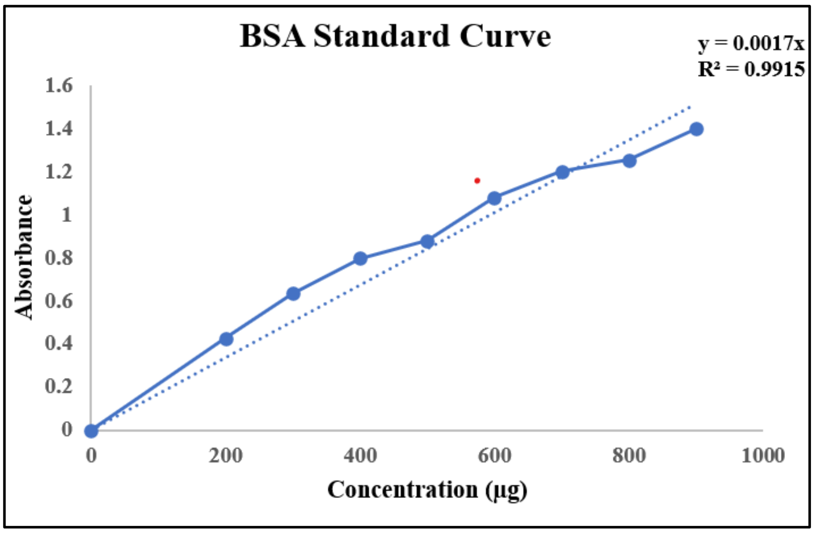


Fig.4.10 Standard curve of BSA.

4.3.3 G6PDH Activity in Various Potato Cultivars

The G6PDH enzyme assay was done from three organs namely, leaves, stems and tubers at different stages of growth. For G6PDH activity, the conversion of NADP to NADPH was recorded for 5 minutes. Specific activity of G6PDH was expressed in U/mg of protein. The specific activity of KC2 and DE showed in the Fig 4.11.

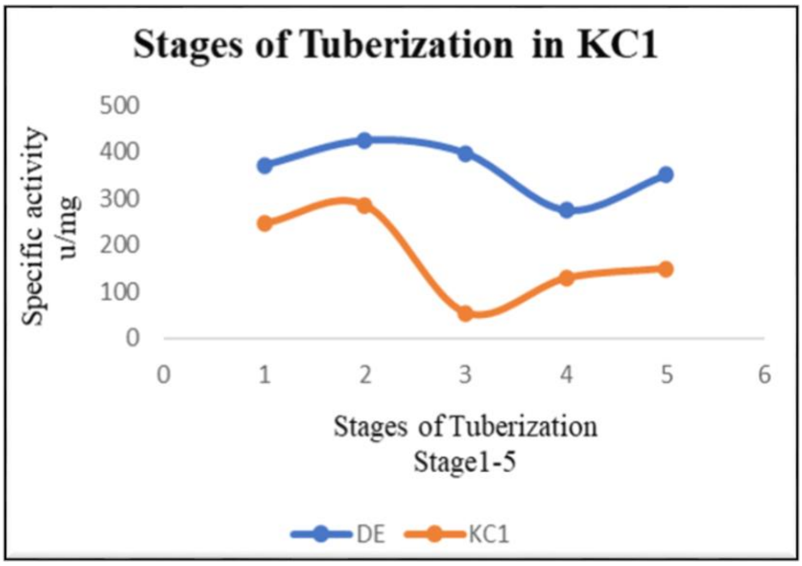


Fig. 4.11 Specific activity of G6PDH in KC1 with reference to DE

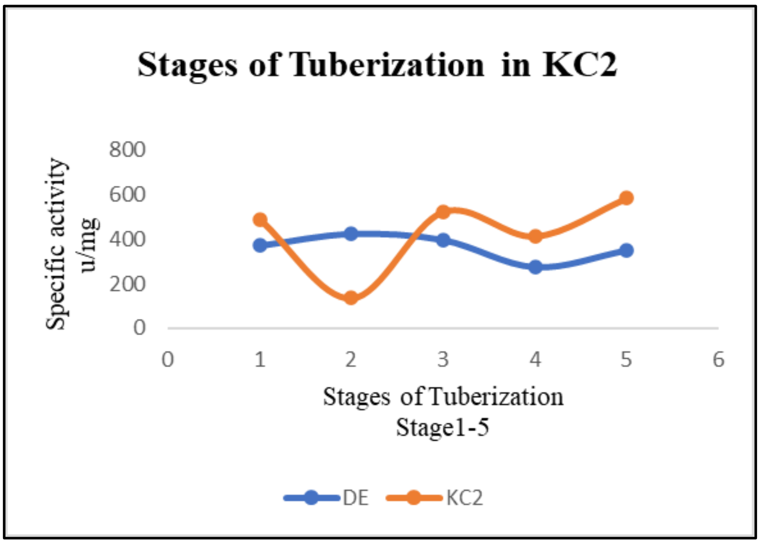


Fig. 4.12 Specific activity of G6PDH in KC2 with reference to DE

4.3.4 G6PDH Activity in Leaves

As shown in table no 15, it was observed that specific activity of G6PDH in leaves was highest in KC1 (2991 U/mg) during young stage and mature stage (3046 U/mg). It was lowest in DE (568 U/mg) during young stage and (467 U/mg) mature stage. From the data, it was observed that specific activity of G6PDH increased two folds towards maturation. The difference could be due to the different stress conditions with regards to the genetic makeup. As G6PDH is directly correlated with ROS, hence higher activity might be the resultant of the leaf efforts to combat both abiotic and biotic stress. This high activity could be used as a desirable trait in crop breeding process (Boeckx et al. 2015).

Table 19 Comparison of G6PDH activity in the leaves of the potato cultivars at different stages of growth.

Potato cultivars	**G6PDH activity in leaves (U/mg of protein)**	
	Young	**Mature**
KC1	2991	3046
KC2	1066	987
PR	1880	2093
KJ	1817	1104
DE	568	467

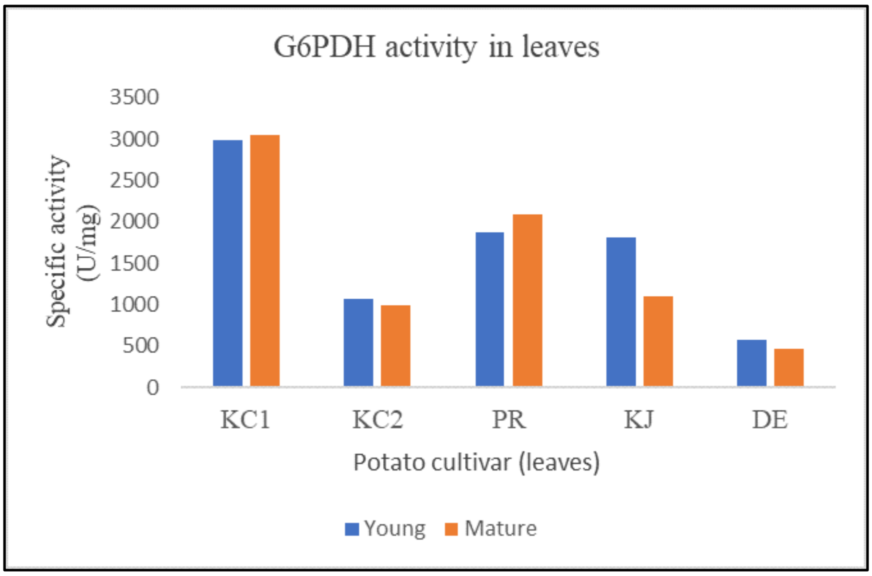


Fig. 4.13 Graph showing different stages of leaves

4.3.5 G6PDH Activity in Stem:

As presented in table 16 it was observed that specific activity of G6PDH in stem was highest in DE during young stage (1984U/mg) and mature stage (2373U/mg). It is lowest in KC1 during young stage (250U/mg) and (376U/mg) mature stage. From the data, it was observed that specific activity of G6PDH increased thrice from young stage to mature. The differences could be due to different stress conditions with regard to genetic makeup. It might be correlated to that mature stems might have high amount of ROS due to high temperature conditions.

Table 20 Comparison of G6PDH activity in the stem of the potato cultivars at different stages of growth.

Potato cultivars	**G6PDH activity in stem (U/mg of protein)**	
Potato cultivar	**Young**	**Mature**
KC1	250	376
KC2	592	777
PR	1843	642
KJ	746	1704
DE	1984	2373

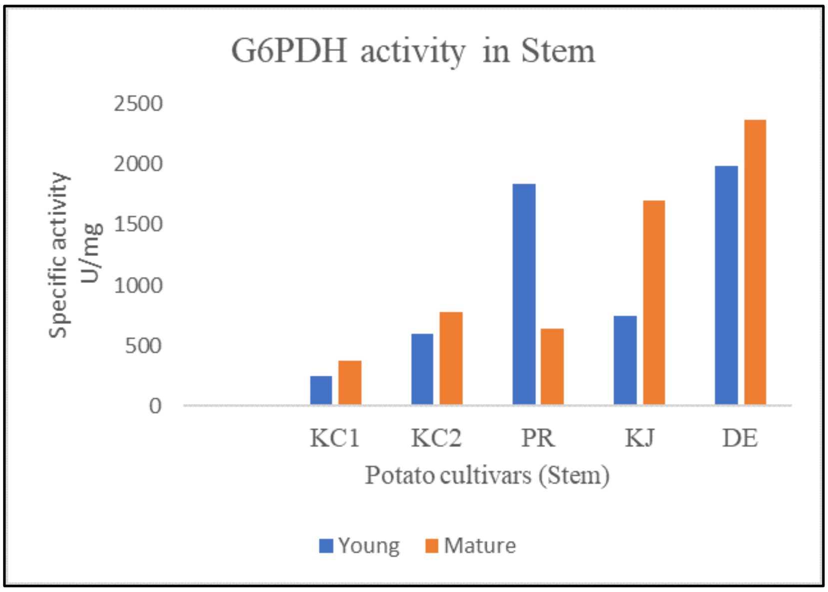


Fig. 4.14 Graph showing different stages of stem

4.3.6 G6PDH Activity in Tubers:

As shown in table, it was observed that specific activity of G6PDH in tuber was highest in PR (1204U/mg) during young stage, (796U/mg) growing stage and (535U/mg) mature stage. It was lowest in KC1 (383U/mg) during young stage, (135U/mg) during growing stage and (388U/mg) during mature stage. Approximately, four-fold increased was observed in the tubers from the growing stages towards maturity.

Table 21 Comparison of G6PDH activity in the tubers of the potato cultivars at different stages of growth.

Potato cultivars	G6PDH activity in tubers (U/mg)					
	Tuberizing stolon	Tuber stage 1	Tuber stage 2	Tuber stage 3	Tuber stage 4	Tuber stage 5
KC1	215	492	572	107	257	299
KC2	452	981	268	1045	825	1166
PR	275	1204	928	964	586	367
KJ	122	854	847	1123	979	427
DE	745	742	849	794	549	701

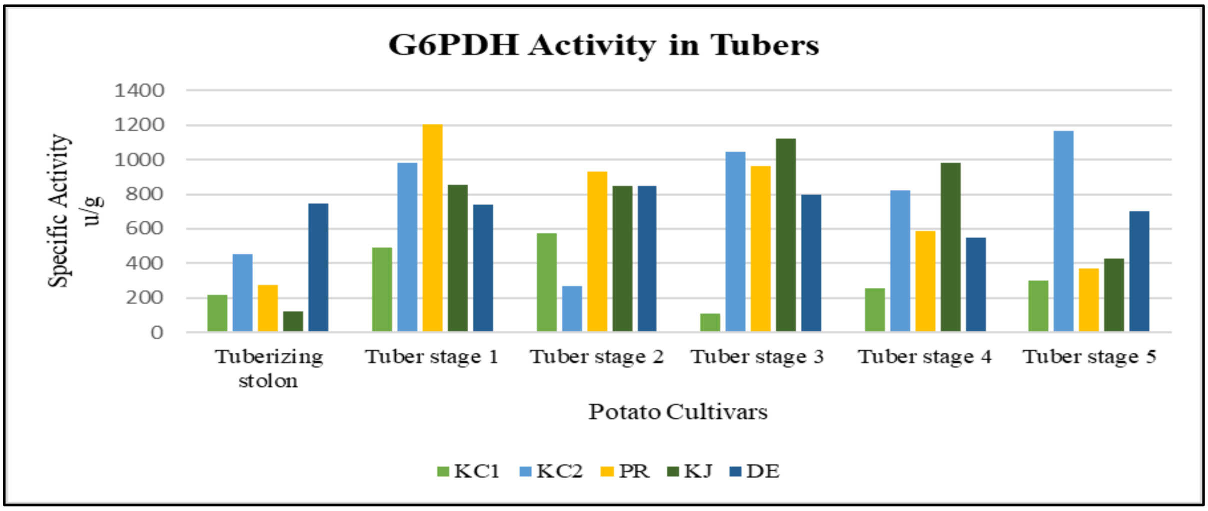


Fig. 4.15 Graph showing different stages of tubers

Concluding Remarks

Various steps of both *in silico* and biochemical approaches adopted in the thesis of are briefly stated below:

- Potato, a non-grain food crop, makes a significant contribution to the world's food security. Tubers developed through the complex process of tuberization are rich source of carbohydrates.
- Glucose 6 phosphate dehydrogenase (EC 1.1.1.49) participates in pentose phosphate pathway (PPP) responsible for supplying reducing energy (NADPH) to the cells.
- Various bioinformatics tools were employed to analyze available multiple HXK isoforms at both gene and protein level in potato.
- *In silico* analysis, revealed multiple G6PDH isoforms majorly localized in cytosol and chloroplast. Genome wide study revealed a total of 4 G6PDH encoding both stable and unstable isoforms. The proteins were found to be mostly thermostable.
- Multiple sequence alignment revealed significant divergence in crucial domains. However, amino acids present in the G6PDH activity site were found to be mostly conserved. Phylogenetic analysis was performed to reveal the relatedness of potato G6PDH members with other plant species.
- Protein-protein interactions showed the association of G6PDH with lycolysis and Pentose phosphate pathway enzymes indicating its role during starch accumulation.
- Four commercially important Indian cultivar namely Kufri chispona 1 (KC1), Kufri chispona 2(KC2), Kufri Jyoti (KJ), Kufri Pukhraj (PR) and a reference cultivar Desiree (DE) were micropropagated, hardened, acclimatized and transferred to fields (for 16 wks).
- Various organs namely leave, stems and tubers at different stages of development were collected and stored at -80º C for biochemical analysis.
- G6PDH enzyme assays were performed by adopting the method described by Bisswanger et al. (2004). Highest enzyme activity was observed in mature leaves and stems of every cultivar indicating role of this enzyme in ROS detoxification at high temperatures.
- Among various stages of tuber development, maximum activity was found to be present in young tubers. This enzyme is involved in the production of NADPH which is an essential electron donor in number of metabolic pathways taking place in developing tubers. This observation was in correlation with the Heatmap generated from Expression Atlas-an *in silico* approach for expression analysis.
- Hence, it could be concluded that this enzyme plays an important role in tuberization and stress mitigation. Further experimental expression studies could be done to deeply explore the role of this enzyme. This study will be helpful in understanding the structural and functional relationship of G6PDH in *Solanaceae* family members more importantly in potato.

Future Perspectives

- After reviewing the biochemical properties of G6PDH in Indian cultivars of potato, the significant kinetic and catalytic activities were explored.
- Comparative study of pH profiling, thermostability and storage stability of crude extract from De and KC1 were analysed.
- Denaturing SDS-PAGE revealed the two bands around 58-66 kDa in the crude extract from the cv. De and KC1 (data not shown here).
- All the potato varieties available in the country are not suitable for processing. For making superior quality processed products, the potatoes must possess certain minimum morphological and biochemical quality character.
- Further these studies could be used to raise Indian cultivars transgenics in terms of starch production and tuber enhancement using various gene editing tools such as CRISPR or enhancing environmental factors which affect the G6PDH activity.
- Also, the study done so far can be used to further analyse the cope to abiotic role of G6PDH role in potato cultivars, which is a crucial concern in the field of agriculture.

References

Abelenda JA, Navarro C, Prat S (2011) From the model to the crop: genes controlling tuber formation in potato. Curr Opin Biotechnol 22:287–292

Abelenda JA, Navarro C, Prat S (2011) From the model to the crop: genes controlling tuber formation in potato. Curr Opin Biotechnol 22:287–292

Appeldoorn N J G , De Bruijn, S M Koot-Gronsveld, E A Visser, R G Vreugdenhil, D, Plas L V D. (1999) Developmental changes in enzymes involved in the conversion of hexose phosphate and and its subsequent metabolites during early tuberization of potato. *Plant, Cell & Environment*, *22*:1085-1096. its subsequent metabolites during early tuberization of potato. *Plant, Cell & Environment*, *22*:1085-1096.

Appeldoorn N J, de Bruijn S M, Koot Gronsveld E A, Visser R G, Vreugdenhil D, van der Plas L H. (1997) Developmental changes of enzymes involved in conversion of sucrose to hexose-phosphate during early tuberisation of potato. *Planta*, *202*:220-226.

Appeldoorn NJ, Sergeeva LI., Vreugdenhil D, van der Plas LH, Visser R. (2002) In situ analysis of enzymes involved in sucrose to hexose-phosphate conversion during stolon-to-tuber transition of potato. Physiologia plantarum, 115:303-310 .

Appeldoorn NJG, De Bruijn SM ,Koot-Gronsveld EA, Visser RGF, Vreugdenhil, D, Plas LVD (1999). Developmental changes in enzymes involved in the conversion of hexose phosphate and its subsequent metabolites during early tuberization of potato. *Plant, Cell & Environment*, *22*:1085-1096.

Barba Espín G, Hernández JA, Diaz Vivancos P (2012) Role of H2O2 in pea seed germination. Plant Signal Behav. 7:193–195

Beals KA (2019) Potatoes, Nutrition and Health. Am J Potato Res 96:102–110 Biotechnology.*105*: 4017-4031.

Brown CR, Culley D, Yang CP, Durst R, Wrolstad R (2005) Variation of anthocyanin and carotenoid contents and associated antioxidant values in potato breeding lines. J Am Soc Hortic Sci 130:174–180

Chen H, Banerjee AK, Hannapel DJ (2004) The tandem complex of BEL and KNOX partners is required for transcriptional repression of *ga20ox1*. Plant J 38:276–284

Cutter EG (1978) Structure and development of the potato plant. In: Harris, P.M. (eds) The Potato Crop. Springer, Boston, pp 70-152

Dutt S, Manjul AS, Raigond P, Singh B, Siddappa S, Bhardwaj V, Kawar PG, Patil VU, Kardile B (2017) Key players associated with tuberization in potato: potential candidates for genetic engineering. Crit Rev Biotechnol 37:942–95 expression during potato tuber development in individuals and populations using the luciferase reporter

expression during potato tuber development in individuals and populations using the luciferase reporter system. Plant Mol Biol 50:653–665

Geigenberger P. (2003). Regulation of sucrose to starch conversion in growing potato tubers. *Journal of Experimental Botany*, *54*:457-465.gene from potato (Solanum tuberosum L.). Plant J 13:419–425

Hannapel, D. J Sharma, P Lin, T Banerjee, A K. (2017) The Multiple Signals That Control Tuber Formation. Plant physiology, 174: 845–856.

Jammer A, Gasperl A, Luschin Ebengreuth N, Heyneke E, Chu H, Cantero Navarro E, Roitsch T (2015) Simple and robust determination of the activity signature of key carbohydrate metabolism enzymes for physiological phenotyping in model and crop plants. J Expt Botany 66:5531–5542

Jiang Z, Wang M, Nicolas M, Ogé L, Pérez-Garcia, M D, Crespel L, Sakr S. (2022) Glucose-6- Phosphate dehydrogenases: the hidden players of plant physiology. International Journal of Molecular Sciences. 23:16128.

Junker B H, Wuttke R, Nunes-Nesi A, Steinhauser D, Schauer N, Büssis D, Fernie A R. (2006) Enhancing vacuolar sucrose cleavage within the developing potato tuber has only minor effects on metabolism. *Plant and cell physiology*, *47*(2), 277-289.

Kolomiets, M. V Hannapel, D. J Chen, H Tymeson M., Gladon, R. J. (2001) Lipoxygenase is involved in the control of potato tuber development. The Plant cell, 13:613–626.

Landi S, Nurcato R De Lillo, A Lentini ,M Grillo S , Esposito, S. (2016) Glucose-6-phosphate dehydrogenase plays a central role in the response of tomato (*Solanum lycopersicum*) plants to short and long-term drought. Plant Physiology and Biochemistry, 105:79-89

Lendzian K, Bassham JA, Regulation of glucose-6-phosphate dehydrogenase in spinach chloroplasts by ribulose 1,5-diphosphate and NADPH/NADP+ ratios. (1975) 260–275.

Liu RH (2013) Health-promoting components of fruits and vegetables in the diet. Adv Nutr 4:384–392

Masi A, Mach R L, Mach-Aigner A. R. (2021) The pentose phosphate pathway in industrially relevant fungi: Crucial insights for bioprocessing. Applied Microbiology and

Masi A, Mach RL, Mach Aigner AR (2021) The pentose phosphate pathway in industrially relevant fungi: Crucial insights for bioprocessing. Appl. Microbiol. Biotechnol. 105:4017–4031.

McGill CR, Kurilich AC, Davignon J (2013) The role of potatoes and potato components in cardiometabolic health: a review. Ann Med 45:467–473

Reddy BJ, Mandal R, Chakroborty M, Hijam L, Dutta P (2018) A review on potato (*Solanum tuberosum* L.) and its genetic diversity. Genetics 10:360–364

Rosin FM, Hart JK, Horner HT, Davies PJ, Hannapel DJ (2003a) Overexpression of a Knotted-Like Homeobox gene of potato alters vegetative development by decreasing gibberellin accumulation. Plant Physiol 132:106–117

Sarkar D (2008) The signal transduction pathway controlling in planta tuberization in potato: an emerging synthesis. Plant Cell Rep 27:1–8

Scharte J, Schön H, Tjaden Z, Weis E, von Schaewen A, Chrispeels M J. (2009) Isoenzyme Replacement of Glucose-6-Phosphate Dehydrogenase in the Cytosol Improves Stress Tolerance in Plants. *Proceedings of the National Academy of Sciences of the United States of America*, *106*:8061–8066.

Singh BP, Rana RK (2014) History of Potato and its Emerging Problems in India. In : Souvenir-National Seminar on Emerging Problems of Potato pp 7-21

Singh DK, Raigond P, Kharumnuid P (2020) Potatoes: The Food for Nutritional Security. Agric & Food :e-Newsl 2:498–501

Struik PC (2007) Above-Ground and Below-Ground Plant Development. Plant Biotechnol J 219–236 system. Plant Mol Biol 50:653–665

Taylor MA, George LA, Ross HA, Davies HV (1998) cDNA cloning and characterization of an *a*glucosidase tuberization. Theor Appl Genet 93:307–316

Van den Berg JH, Ewing EE, Plaisted RL, McMurry S, Bonierable MW (1996) QTL analysis of potato tuberization. *Theoretical and applied genetics 93*:307-316.

Viola R, Roberts AG, Haupt S, Gazzani S, Hancock RD, Marmiroli N, Machray GC, Oparkab J (2001) Tuberization in Potato Involves a Switch from Apoplastic to Symplastic Phloem Unloading. Plant Cell 13:385–398

Von Schaewen A, Langenkamper G, Graeve K, Wenderoth I, Scheibe R. (1995) Molecular characterization of the plastidic glucose-6-phosphate dehydrogenase from potato in comparison to its cytosoliccounterpart. Plant Physiol 109:1327–1335.

Graeve K, von Schaewen A, Scheibe R (1994) Purification, characterization, and cDNA sequence of glucose-6-phosphate dehydrogenase from potato (*Solatium tuberosum L.*). The Plant Journal, 5:353-361

Vreugdenhil D (1999) Initial Anatomical Changes Associated with Tuber Formation on Single-node Potato (*Solanum tuberosum* L.) Cuttings: A Re-evaluation Ann Bot 84:675–680

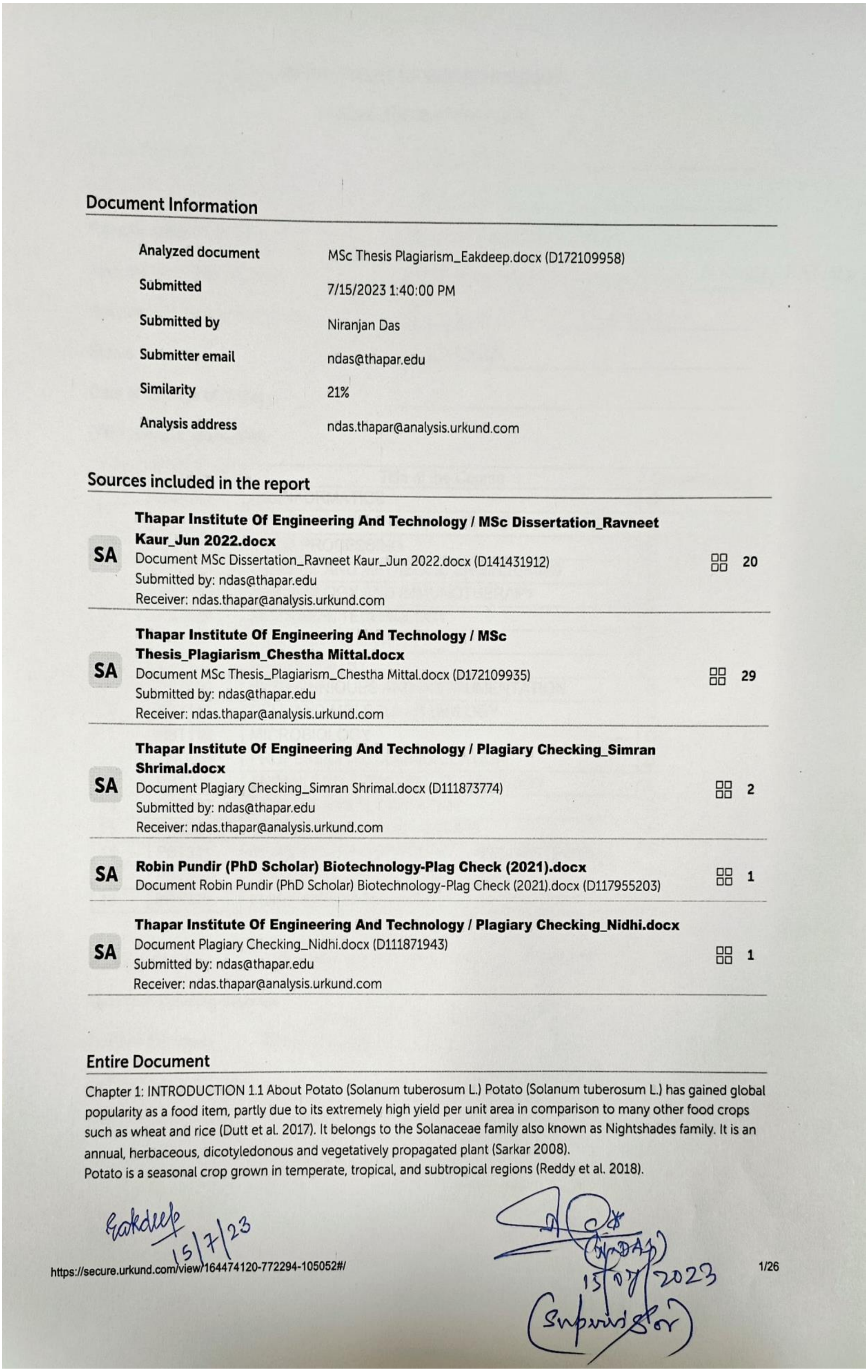
Wakao S, Benning C (2005) Genome-wide analysis of glucose-6-phosphate dehydrogenases in Arabidopsis. 41:243–256

Wang W, Vinocur B, Shoseyov O, Altman A (200 Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. Trends in plant science, 9:244-252**.**

Wenderoth I, Scheibe R, von Schaewen A (1997) Identification of the cysteine residues involved in redox modification of plant plastidic glucose-6-phosphate dehydrogenase. Journal of Biological Chemistry, 272:26985-26990.

Williams JF, Blackmore PF, Duke CC, MacLeod JK (1980) Fact, uncertainty and speculation concening the biochemistry of d-erythrose-4-phosphate and its metabolic roles. Int. J. Biochem 12:339– 344.

Zaheer K, Akhtar MH (2016) Potato Production, Usage, and Nutrition--A Review. Crit Rev Food Sci Nutr 56:711–721



Document Information

Analyzed document	MSc Thesis Plagiarism_Eakdeep.docx (D172109958)
Submitted	7/15/2023 1:40:00 PM
Submitted by	Niranjan Das
Submitter email	ndas@thapar.edu
Similarity	21%
Analysis address	ndas.thapar@analysis.urkund.com

Sources included in the report

SA	**Thapar Institute Of Engineering And Technology / MSc Dissertation_Ravneet Kaur_Jun 2022.docx** Document MSc Dissertation_Ravneet Kaur_Jun 2022.docx (D141431912) Submitted by: ndas@thapar.edu Receiver: ndas.thapar@analysis.urkund.com	**20**
SA	**Thapar Institute Of Engineering And Technology / MSc Thesis_Plagiarism_Chestha Mittal.docx** Document MSc Thesis_Plagiarism_Chestha Mittal.docx (D172109935) Submitted by: ndas@thapar.edu Receiver: ndas.thapar@analysis.urkund.com	**29**
SA	**Thapar Institute Of Engineering And Technology / Plagiary Checking_Simran Shrimal.docx** Document Plagiary Checking_Simran Shrimal.docx (D111873774) Submitted by: ndas@thapar.edu Receiver: ndas.thapar@analysis.urkund.com	**2**
SA	**Robin Pundir (PhD Scholar) Biotechnology-Plag Check (2021).docx** Document Robin Pundir (PhD Scholar) Biotechnology-Plag Check (2021).docx (D117955203)	**1**
SA	**Thapar Institute Of Engineering And Technology / Plagiary Checking_Nidhi.docx** Document Plagiary Checking_Nidhi.docx (D111871943) Submitted by: ndas@thapar.edu Receiver: ndas.thapar@analysis.urkund.com	**1**

Entire Document

Chapter 1: INTRODUCTION 1.1 About Potato (Solanum tuberosum L.) Potato (Solanum tuberosum L.) has gained global popularity as a food item, partly due to its extremely high yield per unit area in comparison to many other food crops such as wheat and rice (Dutt et al. 2017). It belongs to the Solanaceae family also known as Nightshades family. It is an annual, herbaceous, dicotyledonous and vegetatively propagated plant (Sarkar 2008).
Potato is a seasonal crop grown in temperate, tropical, and subtropical regions (Reddy et al. 2018).

Eakdeep
15/7/23

(N DAS)
15/07/2023
(Supervisor)

https://secure.urkund.com/view/164474120-772294-105052#/

1/26