GENOME WALKING FOR PUTATIVE PHYTIC ACID ($InsP_6$) UNIGENE IN BLACK PEPPER (*Piper nigrum* L.)

by

ANANDUCHANDRA GIRIDHARI
(2015-11-061)

THESIS
Submitted in partial fulfillment of the requirement for the degree of

Master of Science in Agriculture
(PLANT BIOTECHNOLOGY)
Faculty of Agriculture
Kerala Agricultural University

CENTRE FOR PLANT BIOTECHNOLOGY AND MOLECULAR BIOLOGY
COLLEGE OF HORTICULTURE
VELLANIKKARA, THRISSUR – 680 656
KERALA, INDIA
2017

DECLARATION

I, hereby declare that the thesis entitled **"GENOME WALKING FOR PUTATIVE PHYTIC ACID ($InsP_6$) UNIGENE IN BLACK PEPPER (*Piper nigrum* L.)"** is a bonafide record of research work done by me during the course of research and that this thesis has not been previously formed the basis for the award to me of any degree, diploma, fellowship or other similar title, of any other University or Society.

Vellanikkara,
Date: 24/11/2017

ANANDUCHANDRA GIRIDHARI
(2015-11-061)

CERTIFICATE

Certified that this thesis entitled **"GENOME WALKING FOR PUTATIVE PHYTIC ACID ($InsP_6$) UNIGENE IN BLACK PEPPER (*Piper nigrum* L.)"** is a bonafide record of research work done independently by Ms. Ananduchandra Giridhari under my guidance and supervision and that it has not previously formed the basis for the award of any degree, diploma, associateship or fellowship to her.

Vellanikkara,
Date: 24/11/17

Dr.Sujatha R.
(Major Advisor, Advisory Committee)
Professor
Dept.of Plant Biotechnology
College of Agriculture,
Padanakkad

CERTIFICATE

We, the undersigned members of the Advisory Committee of Ms. Ananduchandra Giridhari, a candidate for the degree of **Master of Science in Agriculture** with major field in Plant Biotechnology, agree that the thesis entitled **"GENOME WALKING FOR PUTATIVE PHYTIC ACID ($InsP_6$) UNIGENE IN BLACK PEPPER (*Piper nigrum* L.)"** may be submitted by Ms Ananduchandra Giridhari. in partial fulfillment of the requirement for the degree.

24/11/17

Dr. Sujatha R.
(Major Advisor, Advisory Committee)
Professor
Dept.of Plant Biotechnology
College Of Agriculture, Padanakkad,
Kasargode

24/11/2017

Dr. Shylaja M. R
(Member, Advisory Committee)
Professor and Head
Centre for Plant biotechnology and
Molecular Biology (CPBMB)
College of Horticulture, Vellanikkara

Dr. Deepu Mathew
(Member, Advisory Committee)
Assistant Professor
Centre for Plant Biotechnology and
Molecular Biology (CPBMB)
College of Horticulture, Vellanikkara

Dr. Joy. M
(Member, Advisory Committee)
Associate Professor
Dept. of Plant Pathology
College of Agriculture, Vellayani

EXTERNAL EXAMINER
R.L. RAVIKUMAR
Professor,
Dept. of Plant Biotechnology,
GKVK, UAS, Bangalore.

4

ACKNOWLEDGEMENT

First and foremost I bow my head before the ***Almighty God*** *for enlightening and making me confident and optimistic throughout my life and enabled me to successfully complete the thesis work in time.*

It is with immense pleasure I avail this opportunity to express my deep sense of whole hearted gratitude and indebtedness to my major advisor ***Dr. Sujatha. R,*** *for her expert advice, inspiring guidance, valuable suggestions, constructive criticisms, constant encouragement, affectionate advice and above all, the extreme patience, understanding and wholehearted co-operation rendered throughout the course of my study. I really consider it my greatest fortune in having her guidance for my research work.*

I consider it as my privilege to express my deep-felt gratitude to ***Dr. Shylaja M. R,*** *Professor and Head, CPBMB, CoH, Vellanikkara and member of my advisory committee for her constant support, valuable suggestions and critical scrutiny of the manuscript.*

I express my gratitude to ***Dr. Deepu Mathew,*** *Assistant Professor, CPBMB, CoH, Vellanikkara and member of my advisory committee for his immense help and assistance provided for constituting the manuscript.*

I sincerely thank ***Dr. Joy M,*** *Professor, Department of Plant Pathology, CoA, Vellayani and member of my advisory committee for his expert advice, constant inspiration, precious suggestions, generous support and constructive criticisms during my entire study which helped in successful completion of this work.*

I express my sincere thanks to ***Dr.Valsala P,*** *retired Professor of CPBMB for her for her support, enthusiasm and relevant suggestions.*

I am deeply obliged to the teaching assistants, research assistants and non-teaching staffs at CoA, Padannakkad as well as CPBMB, Vellanikkara especially ***Ms. Deeshma K.P*** *for their invaluable help, guidance and critical assessment throughout the period of work. I thank them for all the help and cooperation they has extended to me.*

I duly acknowledge the encouragement, moral support, precious suggestions and timely persuasions by my dear seniors, not only in my research work but also throughout my PG programme. I express my sincere thanks to my classmates ***Marjan, Arunima, Geethu, Priya, Deepali, Debashish, Basil, Ashwin, Pramod, Kishore and Swapnil*** *for their affection and kind help offered during my thesis work. I have infinite pleasure to express whole hearted thanks to my loving juniors for their love, innumerable help and support.*

I thank my dear friends ***Geethumol, Deepa, Aparna, Nusrath, Nadhika, Salpriya, Rosemary, Preethy, Zita, Nagendra, Sunil, Sunayana, Ashiba, Sherin*** *and junior friends at CoA, Padannakad for the unconditional support, help, timely valuable suggestions and encouragement which gave me enough mental strength and perseverance to get through all odds and tedious circumstances and immense thanks to all M.Sc. classmates for their moral support and encouragement.*

I am in dearth of words to express my love towards ***my beloved*** *Parents and my lovely brother* ***Ananduchandra Khanashyam*** *for their boundless affection, moral support, eternal love, deep concern, prayers and personal sacrifices which sustains peace in my life.*

I owe special thanks to Librarian, College of Horticulture, ***Dr. A.T. Francis*** *and all other staff members of Library, who guided me in several ways, which immensely helped for collection of literature for writing my thesis. Iam thankful for the service* ***Mr. Aravind*** *had done all along the academic period.*

I express my deep sense of gratitude to Kerala Agricultural University for financial and technical support for persuasion of my study and research work. It would be impossible to list out all those who have helped me in one way or another in the successful completion of this work. I once again express my heartful thanks to all those who helped me in completing this venture.

My word of apology to those I have not mentioned in person for the successful completion of this endeavor.

Ananduchandra Giridhari

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF PLATES

ABBREVATIONS

%	Percentage
^{o}C	Degree Celsius
ABA	Abscisic acid
ADAR 2	Adenosine deaminase acting on RNA 2
ADP	Adenosine diphsophate
APS	Ammonium persulphate
ATP	Adenosine triphosphate
BACE1	Beta secretase 1
BLAST	Basic local alignment search tool
bp	Base pair
BSA	Bovine serum albumin
CDD	Conserved domains database
cDNA	Complementary deoxyribonucleic acid
CTAB	Cetyl Trimethyl Ammonium Bromide
DAF	Days after flowering
DAPI	4, 6-diamidino-2-phenylindole
DBT	Department of biotechnology
DNA	Deoxy ribonucleic acid
dNTP	Deoxyibo Nucleoside Triphosphate
EDTA	Ethylene Diamine Tetra Acetic acid
E-GW	Extension based genome walking
EST	Expressed sequence tags
FAB1	Forms Aploid and Binucleate cells 1
FW	Fresh weight
g	Gram

GA3	Gibberellic acid
GO	Gene ontology
GW	Genome walking
h	Hour
ha	Hectare
HPLC	High performance liquid chromatography
IMP	Inositol monophosphate phosphatase
Ins	Inositol/ *myo* inositol
$InsP_1$	Inositol monophosphate
$InsP_3$	Inositol triphosphate
$InsP_5$	Inositol pentakisphosphate
$InsP_5$-K	Inositol pentakisphosphate kinase
$InsP_6$	Inositol hexakisphosphate
$InsP_6$-K	Inositol hexakisphosphate kinase
$InsP_7$	Diphosphoinositol pentakisphosphate
InsPs	Inositol phosphates
IPK1	Inositol pentakisphosphate 2- kinase
IPK2	Inositol trisphosphate 3/6- kinase
IPKs	Inositol phosphate kinases
IPS	*Myo* Inositol 3-phosphate synthase
ITPK	Inositol tris/tetra phosphate 5/6- kinase
kb	Kilo base
kDa	Kilo Dalton
lpa	Low phytic acid
M	Molar
mA	Milli ampere
MDA	Multiple strand displacement amplification

MDD-HPLC	Metal dye detection HPLC
mg	Milligram
MIK	*Myo* inositol kinase
min	Minute
MIPS	D- *myo* inositol 3 phosphate synthase
ml	Millilitre
mM	Millimolar
MPTP	1-methyl-4-phenyl-1, 2, 3, 6 - tetrahydropyridine
mRNA	Messenger RNA
MRP	Multidrug resistance associated protein
NCBI	National Centre for Biotechnology Information
NF-κB	Nuclear factor κB
ng	Nanogram
NGS	Next generation sequencing
nmole	Nanomole
NMR	Nuclear magnetic resonance
ORF	Open reading frame
PA	Phytic acid
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
p-ERK	phosphorylated Extracellular signal-Regulated Kinase
PGK	Phosphoglycerate kinase
P-GW	Primer based genome walking
Pi	Inorganic phosphorus
PI-3 kinase	Phosphoinositide-3 kinase
PLC	Phospholipase C
PMD	Powdery mildew disease

PPx-InsPs	Inositol pyrophosphates
PtdIns (4,5) P2	Phosphatidyl inositol 4, 5-bisphosphate
PtdInsP	Phosphatidyl inositol phosphate
PtdIS	Phosphatidyl inositol synthase
PVP	Polyvinyl pyrrolidone
R.H	Relative humidity
RCA	Rolling circle amplification
R-GW	Restriction based genome walking
RNAi	RNA interference
ROS	Reactive oxygen species
rpm	Rotation per minute
rRNA	Ribosomal RNA
rRNA	Ribosomal ribonucleic acid
SA	Salicylic acid
SBS	Sequencing by synthesis
SDA	Strand displacement amplification
SDS-PAGE	Sodium dodecyl sulpahate- polyacrylamide gel electrophoresis
sec	Second
SNP	Single nucleotide polymorphism
t	Tonnes
TAE	Tris Acetate EDTA
TBE	Tris Borate EDTA
TCOF1	Treacle ribosome biogenesis factor1
T-DNA	Transfer DNA
TEMED	Tetramethyl ethylenediamine
TIR1	Transport inhibitor response 1
TMV	Tobacco mosaic virus

UBF	Upstream binding factor
UTR	Untranslated region
UV	Ultra violet
V	Voltage
WGA	Whole genome amplification
YMV	Yellow mosaic virus
α	Alpha
β	Beta
γ	Gamma
µg	Micro gram
µL	Micro litre
µM	Micro molar

Introduction

1. INTRODUCTION

Black pepper (*Piper nigrum*), belonging to the *Piperaceae* family is one of the major spice crops of Kerala. It is renowned as the "king of spices" owing to its enormous medicinal and other uses and "black gold" as it fetches the highest return from international trade. Black pepper has many uses like dietary, medicinal, perfumery and preservative. The pungency in black pepper is due to the presence of an alkaloid piperine. Piperine and other essential oils present in pepper are responsible for its antioxidant potential, anti-inflammatory potential, anti-microbial property, anti-cancerous property, and neuroprotectant activity (Butt *et. al.*, 2013). As the crop is an economically important high value crop, it provides an ideal platform for conducting molecular studies that can effectively enhance the productivity and nutritional value of the crop.

Myo-inositol hexakisphosphate, also called phytic acid ($InsP_6$) is an important compound found in eukaryotic cells (Raboy, 2003). In plants, $InsP_6$ is a major reserve of phosphates, found mostly in seeds and other plant tissues like tubers, roots and pollen (Raboy, 1997). In mature seeds, $InsP_6$ is found as phytate salts of mineral cations like K^+, Mg^{2+}, Ca^{2+}, Mn^{2+}, and Zn^{2+} (Lott, 1984). On germination phytate gets hydrolysed to release phosphates that support the growth and development of seedlings (Murphy *et al.*, 2008). In guard cells, $InsP_6$ acts as a key signaling molecule produced in response to drought stress hormone abscisic acid and triggers the release of endomembrane stored calcium ions (Lemtiri-Chlien *et al.*, 2003). Studies conducted have found that plants defective in $InsP_6$ are more susceptible to diseases including viral, bacterial and fungal pathogens (Murphy *et al.*, 2008), thus performing a key role in plant defence mechanism also.

But the accumulation of phytic acid negatively impacts the health and nutrition of animals as well as causes environmental hazards. Phytate in seeds is a major anti-nutrient, leading to the reduced bioavailability of micronutrients like

Zn and Fe (Raboy, 2001). When seeds of cereals or pulses with high phytate content are used in animal feed, it may result in indigestion as the monogastric animals cannot utilize the $InsP_6$-P. This undigested phytic acid is accumulated in waterbodies causing eutrophication and thus polluting the environment (Filleppi *et al.*, 2010). This has led to the efforts to lower the phytate content of crops, resulting in development of low phytic acid mutants of various food crops. Similarly, crops that have altered levels of phytic acid can be developed by manipulating the genes that are coding for enzymes responsible in the synthesis of phytic acid or its intermediary compounds in the biosynthesis pathway of $InsP_6$. Identification of genes responsible for $InsP_6$ biosynthesis in plant system can therefore help in beneficial genetic manipulations.

The enormous quantity of sequence information obtained for various crops through the whole genome sequencing has provided a reference platform for the plant genome studies. The actual role played by the $InsP_6$ in plant systems and their biological conversion studies can be taken up only by sequencing the gene or gene family responsible for phytic acid metabolic pathway. Genome walking is a method used for the isolation and characterization of regions flanking a known sequence within a genome (Gadkar and Filion, 2011). There are different genome-walking strategies used for sequencing the flanking region of a known gene fragment. Inverse and ligation mediated genome walking make use of the presence of a definite restriction site in the target region whereas the randomly primed PCR walk make use of the ability of the oligonucleotide primer to bind non-specifically to the genomic DNA (Reddy *et al.*, 2008). Phi29 DNA polymerase based rolling circle amplification followed by selective amplification is a novel method of genome walking suggested by Reddy *et al.* (2008). It uses degenerate primers for whole genome amplification followed by amplification of the specific gene by PCR using different gene specific primers and their corresponding nested primers.

Hence, in realization of the enormous potential of the crop in view of its multitude of uses in the medicinal and dietary fields, its high commercial value in agricultural sector, and on identification of the importance of the biomolecule $InsP_6$ in plants, the present study was undertaken to carry out the genome walking using the sequence of the unigene *Pnc135* (995 bp) which is a putative phytic acid gene in black pepper obtained from Expressed Sequence Tags (ESTs) generated using Ion Torrent technology (Ion Personal Genome Machine™, Life Technologies) as part of a DBT project (unpublished data, Sujatha, R). The gene ontology (GO) analysis showed that the unigene had an inositol pentakisphosphate kinase ($InsP_5$-K) activity which is the enzyme responsible for the bioconversion of Inositol pentakisphosphate ($InsP_5$) into $InsP_6$. The $InsP_5$-K gene *(ipk1)* reported from other crops had a size range of 3 to 6 kb. Therefore the study encompasses the following objectives:

- To sequence the flanking regions of unigene *Pnc135* coding for Phytic acid ($InsP_6$) by primer walking and validation of the gene.
- Detection and quantification of $InsP_6$ and its derivatives [inositol pyrophosphates (PPx-InsPs)], if any, in black pepper.


Review
Of
Literature

2. REVIEW OF LITERATURE

2.1 BLACK PEPPER

Black pepper (*Piper nigrum*) is a perennial vine belonging to the family *Piperaceae*. It is known as the 'King of Spices' for its economic importance and extensive use as a spice and in medicine. The crop is found to have originated in the tropics of Western Ghats of India. Presently, black pepper is largely cultivated in countries like Thailand, Ecuador, Cambodia, India, Brazil, Indonesia, Malaysia, Sri Lanka, Vietnam and China. India is one of the leading producers of black pepper and the crop accounts for a significant share of foreign exchange. In the country, Kerala and Karnataka account for a major share of production (Bhattacharya and Bandyopadhyay, 2017). The area, production and productivity of black pepper in India is estimated to be 1,31,230 ha, 55,500 tons and 2.36 t/ha respectively and Kerala is estimated to have an annual production of 24,000 tons from an area of 85,430 hectares (Spice Board India, 2016-2017).

2.1.1 Climate and topography

Plant prefers a tropical humid climate with high rainfall and humidity and the crop grows successfully between 20° North and South latitudes and up to 1500 m above mean sea level. The optimum temperature ranges from 23 – 32°C even though the plant can tolerate temperatures of 10 – 40°C (Sivaraman *et al.*, 1999). A well distributed rainfall of 2000 mm is found ideal for the crop. Though red laterite soils are found to be best suited for black pepper, it grows well on heavy clay to light sandy clay soils.

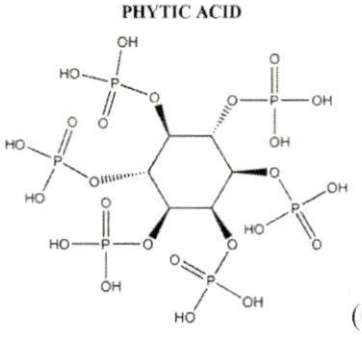
2.1.2 Uses of black pepper

Black pepper is a spice crop that is found to have numerous uses. Piperine, is the active ingredient present in black pepper which is responsible for the pungency of berries. Peppercorn is either used as a whole or its active components in the foods. The black pepper berries as well as the secondary metabolites are being used as drugs, preservatives, insecticides or larval control agents. It is found

to have an anti-oxidant, antidepressant, antibacterial and anti-fungal property. Pepper is also an effective medicine for gastric ailments, cough and cold. It is used as an anti-inflammatory drug. The alkaloids and active ingredients present in black pepper are found to have an anti-tumour action and thus a chemo preventive role. Piperine is also found to assist cognitive brain functions (Ahmad *et al.*, 2012; Butt *et al.*, 2013; Srinivasan, 2007).

2.2 PHYTIC ACID

Reddy *et al.* (1982) have reported that Phytic acid, scientifically known as inositol (1,2,3,4,5,6)-hexakis phosphate, is a member of the large family of inositol phosphates and was discovered in 1903 by Hartig. It is a cyclical ring compound with six phosphate moieties attached to the inositol ring and is present naturally in all eukaryotes (Brown *et al.*, 1961; Sasakawa *et al.*, 1995).



(Guerrero *et al.*, 2016)

There are two different isomers for inositol hexakisphosphates, the *myo*-inositol and *scyllo*- inositol forms that are present at different pH (Bauman *et al.*, 1999). Phytic acid is the major form of inositol phosphate present in developing seeds and contributes mainly to the organic phosphorous reserve in seeds. Phytate is found in storage organs like seeds, tubers, and roots and aids in germination and growth by acting as a source of phosphorus (Morton and Raison, 1963; Aberoumand, 2011). Plant seeds and grains have phytate in the range of 0.5 – 5 percent (w/w) and is mainly present as complex salts of Mg^{2+}, K^{+}, Ca^{2+}, Zn^{2+}, Fe^{3+}, and Cu^{2+} (Loewus, 2002 cited by Greiner *et al.*, 2006). Phytate when bound to Fe (as phytin) is a powerful anti-oxidant and acts as a potent inhibitor of Fe driven formation of reactive oxygen species and provides a safe binding site for

Fe during the transport through cytosol or into the cellular organelles (Graf *et al.*, 1984; Hawkins *et al.*, 1993). Phytate is known to support seedling establishment during early stages. It may play a different role in seed development and maturation during dehydration. (Valluru and Van den Ende, 2011).

As phytate is a natural compound present in plant seeds and grains, it is taken up as a part of dietary intake. Greiner *et al.* (2006) have reported that on an average, the daily intake of phytate is estimated to be 2000–2600 mg for vegetarian diets as well as diets of inhabitants of rural areas of developing countries and 150–1400 mg for mixed diets. Even though phytic acid is found to have an anti-nutrient property researches have proven many health benefits of phytic acid in diets.

2.2.1 Functions of Phytic acid (PA) and Inositol Phosphates (InsPs)

In eukaryotic cells, $InsP_6$ and its derivatives function in various signaling processes like mRNA export (York *et al.*, 1999), repair of DNA double strand breaks (Hanakahi *et al.*, 2000) and control of transcription (Shen *et al.*, 2003). It is also found that $InsP_6$ plays a role as a regulator of apoptosis, both Tumor Necrosis Factor (α mediated and Fas mediated) by inducing anti-apoptotic gene functions (Verbsky and Majerus, 2005). $InsP_6$ is also found to be a cofactor in auxin mediated gene expression, necessary for the recognition of auxin by transport inhibitor response1 (TIR1), a member of F-box protein family (Tan *et al.*, 2007) and auxin signalling has been linked to antibacterial resistance (Navarro *et al.*, 2006).

The anti-oxidant activity of phytic acid which inhibit the formation of reactive oxygen species (ROS) can prevent carcinogenesis (Shamsuddin and Vucenik *et al.*, 1999; Silva and Bracarense, 2016). Huang *et al.* (1997) have investigated the role of $InsP_6$ on tumor promoter-induced cell transformation and signal transduction pathways leading to activator protein 1 activation, which is considered to play a crucial role in tumor promotion. The results have provided

the first evidence of the molecular mechanism of $InsP_6$ antitumor promotion effect by targeting and blocking Phosphoinositide-3 kinase (PI-3 kinase) activation. They demonstrated that PI-3 kinase can serve as a molecular target for the development of cancer chemo preventive agents.

Studies in *Saccharomyces cereviseae* first revealed the *in-vivo* role of Inositol hexakisphosphate ($InsP_6$) in mRNA export. $InsP_6$ is found to bind clathrin assembly proteins AP2 and AP3, inhibiting clathrin assembly (Verbsky *et al.*, 2002). Dietary phytate is reported to protect against kidney stone formation, diabetes mellitus, caries, atherosclerosis and coronary heart disease as well as a variety of cancers (Greiner *et al.*, 2006).

Bhandari *et al.* (2008) report that inositol hexakisphosphate kinase ($InsP_6K$) mutant mice which had decreased levels of inositol pyrophosphates showed defects in spermiogenesis, insulin secretion and growth compared to the wild type. Ahmed and Ebaid (2015) observed that phytic acid had the ability to attenuate ileum injury caused by *Aeromonas hydrophila* and its endotoxins in mice sepsis models. They concluded that the protective effects may be due to increased mucous secretion, decreased apoptotic index, attenuating the inflammatory and lymphocytic cell count or increasing the renewal of the crypt cells and villous epithelial cell proliferation.

PA is found to inhibit cardiovascular diseases like dyslepidemia, and platelet aggregation. It can also influence the influx of extracellular calcium leading to release of insulin and it control progression of Alzheimer's disease (Vucenik *et al.*, 1999; Larsson *et al.*, 1997; Anekonda *et al.*, 2011). Anekonda *et al.* (2011), in their experiments conducted on mice model have found that phytic acid provided complete protection against amyloid precursor protein C- terminal fragment- induced cytotoxicity by regulating levels of increased intracellular calcium, hydrogen peroxide, superoxide, Aβ oligomers and moderately upregulating the expression of autophagy protein (beclin-1). Similar works by

Abe and Taniguchi (2014) have identified phytic acid (PA) as a β-secretase 1 (BACE1) inhibitory molecule in rice extracts and digest. BACE1 could prevent the accumulation of amyloid β peptide in brain, which can cause Alzheimer's disease. Phytic acid is also found to have protective effect against 1-methyl-4-phenyl-1, 2, 3, 6-tetrahydropyridine (MPTP) induced Parkinson's disease in mice model studies. This was correlated with its anti-inflammatory effect associated with suppression of pathways that are involved in nuclear factor κB (NF-κB) and phosphorylated extracellular signal-regulated kinase (p-ERK) (Lv *et al.*, 2015). Inositol hexakisphophate ($InsP_6$) is also found to bind with human Adenosine deaminase acting on RNA 2 (ADAR2) enzyme required for RNA editing. Inositol pentakisphosphate kinase ($InsP_5K$) has functionally significant interactions with three proteins that regulate rRNA synthesis: protein kinase CK2, Treacle ribosome biogenesis factor1 (TCOF1) and upstream-binding-factor (UBF) (Macbeth *et al.*, 2005; Brehm *et al.*, 2013).

PA is proven to provide resistance against various abiotic as well as biotic stresses in plants. It is reported to give resistance against various bacterial, fungal and viral infections (Murphy *et al.*, 2008). Reports show that transgenic potato plants with lowered $InsP_6$ had lowered resistance to avirulent pathogen Potato Virus Y and virulent pathogen tobacco mosaic virus (TMV). Similarly, in Arabidopsis, mutants for Inositol pentakisphosphate kinase ($InsP_5K$) expressed enhanced susceptibility to TMV and bacterial pathogen *Pseudomonas syringae*. *Arabidopsis* myo-inositol 3-phosphate synthase mutants *(atips2)* were also hyper susceptible to TMV, turnip mosaic virus, cucumber mosaic virus and cauliflower mosaic virus as well as to the fungus *Botrytis cinerea* and to *P. Syringae* (Murphy *et al.*, 2008).

Experiment on maize $InsP_6$ mutant line (*lpa1-241*), either un-aged or incubated for 7days at 46°C and 100 percent relative humidity showed lower germination capacity, 50 percent weakly bound iron in kernals and a higher

concentration of free radicals in embryo. Thus, phytic acid was found to provide protection against oxidative stress during seed development (Doria *et al.*, 2009).

Rice cells cultured with abscisic acid (ABA) were found to accumulate higher levels of phytic acid compared to the cells cultured without ABA. Also, myo inositol 3- phosphate synthase (*RINO1),* myo inositol kinase *(OsMIK),* inositol pentakisphosphate 2 kinase *(OsIPK1)* and low phytic acid (*OsLPA1)* mutants have shown low phytic acid phenotypes in seeds. Both the developing seeds as well as ABA cultured cells also exhibited similarity in upregulated expression of inositol tris/tetraphosphate kinase (*OsITPK4* and *OsITPK6)* (Matsuno and Fujimuro, 2014).

Aggarwal *et al.* (2015), conducted similar works on the expression studies of eight genes (*TaIMP, TaITPK1-4, TaPLC1, TaIPK2, TaIPK1*) in wheat $InsP_6$ biosynthesis in response to abcissic acid (ABA), giberilic acid (GA3) and salicylic acid (SA). The exogenous application of ABA was found to induce six out of eight phytic acid pathway genes and GA3 suppressed five genes that were induced by ABA showing an antagonistic effect. Phytic acid genes were also induced under multiple abiotic stress conditions like cold and drought. Also, biotic stress induced expression study using *Fusarium graminearum* (fungi), *Mayetiola destructor* (fly) and a fungal toxin ToxB (elicitor) could produce a heat map with significantly upregulated *TaITPK1* on exposure to *M. destructor* and other genes showing slight down regulation under biotic stresses.

Dhole and Reddy (2016) reported that high PA in seed is associated with resistance reaction to diseases like yellow mosaic disease (YMD) and powdery mildew disease (PMD) and insect pest bruchid (*Callosobruchus*) in mungean. The study also revealed that PA itself is not imparting resistance to various biotic stresses but it may enhance the expression of resistance genes already present in plant. Results showed that PA content as high as two percent in seeds had antibiosis mechanism against bruchid and to maintain resistance against YMD and

PMD, seed PA should be >8 mg g^{-1} while for bruchid resistance seed PA should be very high (~18 mg g^{-1}).

Inositol pyrophosphates (PPx-InsPs) have found to regulate ATP at cellular levels. Review by Williams *et al.* (2015), have shown that diphosphoinositol pentakisphosphate ($InsP_7$) mutant yeast and mice were found to produce altered levels of ATP which was due to effects on genes involved in glycolysis. At organism level, $InsP_7$ was found to act upon the sugar homeostasis through glucose uptake and insulin release in animals. In plants, InsPs are essential in Pi response and homeostasis but was found that $InsP_6$ was not the molecule involved, but, since PPx-InsPs are synthesized from $InsP_6$, it is considered a critical player in Pi sensing.

2.2.2 $InsP_6$ and other InsPs in cells

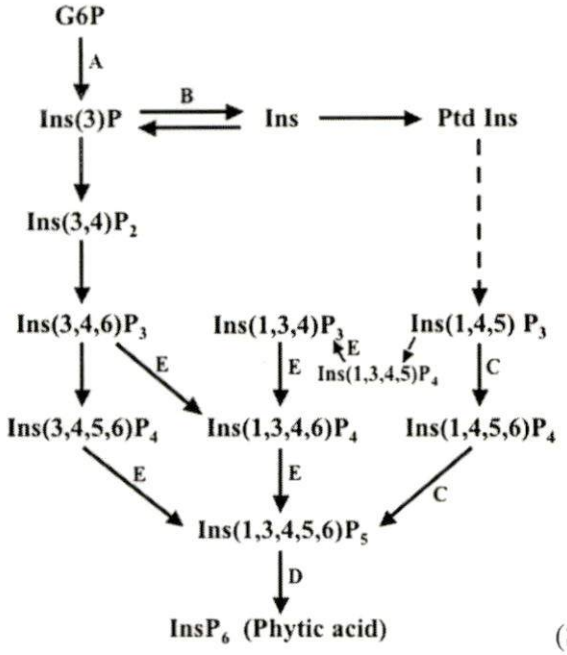
Inositol (Ins) is the major sugar present during development of kiwifruit (*Actinidia arguta*), accounting for 60-65 percent of the total sugars (Klages *et al.*, 1998, cited by Yu *et al.*, 2014). $InsP_6$ can represent from one to several percentage of typical seeds dry weight, about 75±10 percent of a seed's total phosphorus (P) and normally >90 percent of a mature seed's total acid extractable Ins phosphate (Raboy, 1997). Reddy *et al.* (1982) have reported that the amount of phytic acid varies from 0.50 to 1.89 percent in cereals (except polished rice), from 0.40 to 2.06 percent in legumes, from 2.00 to 5.20 percent in oil seeds except soybeans and peanuts (grouped under legumes), and from 0.40 to 7.50 percent in protein products. Reports have shown that seeds of cereal species such as maize (*Zea mays* L.), barley (*Hordeum vulgare* L.), and rice (*Oryza sativa* L.) contain about 3 mg per g PA phosphorus (~75%) and approximately 4 mg total P per g seed dry weight (Raboy *et al.*, 2001). Phytin is accumulated in both the embryo and living endosperm in castor bean (Organ *et al.*, 1988). The phosphorus fraction stored as phytate range from 30 percent in roots upto 80 percent in seeds and cereals (Kanani and Poonam, 2015). Studies have estimated the annual production of $InsP_6$ by seed crops to be more than 50 percent of all the Phosphorus fertilizers

applied annually, worldwide (Lott *et al.*, 2000). Approximately 65 percent of the fertilizers applied are converted to phytic acid and this account for about 65-80 percent of total seed phosphorus (Raboy, 1997). Reviews by Raboy (2003), have shown that in small grains like barley, wheat and rice, >80 percent of the whole grain's $InsP_6$ is sequestered in the aleurone layer. Thus it represents ≥20 percent of the tissue's mature dry weight. In maize, 80 percent of phytate gets accumulated in embryo and scutellum whereas, in case of legumes like common bean, >95 percent of phytic acid is present in cotyledons (Sparvolli and Cominelli, 2015). Arabidopsis seeds were found to contain less than two percent of the total inositol pool as inositol pyrophosphates (1.33% $InsP_7$, 0.24% $InsP_8$). Vegetative tissue from Arabidopsis was also analyzed and inositol pyrophosphates (PPx-InsPs) were found in both seedlings (0.64% $InsP_7$, 0.14% $InsP_8$) and mature leaves (1% $InsP_7$). $InsP_7$ was detected in other plant species as well, including another member of the Brassica family, *Camelina sativa* (1.40% in seedlings), and also in cotton (*Gossypium hirsutum*) in the leaves and in shoots and roots of seedlings (Williams *et al.*, 2015).

It is shown that $InsP_6$ is a major component of the extracellular hydatid cyst wall produced by the parasitic cestode *Echinococcus granulosus*, and represent >90 percent of total cell wall (Irigoin *et al.*, 2002 cited by Raboy, 2003) and in mammalian cells, the concentrations of $InsP_6$ are between 10 and 100 μM (Huang *et al.*, 1997)

2.2.3 InsP pathway:

In plants, $InsP_6$ can be synthesized either through a lipid-dependent or lipid-independent pathway.



(Suzuki *et al.*, 2007)

A: MIPS, B: IMP, C: IPK2, D: IPK1, E: ITP5/6K

PA biosynthesis need the production of Ins in a highly conserved method that occurs in all living organisms, where the enzyme D- myo inositol 3 phosphate synthase (MIPS) converts D-glucose 6-PO_4 to myo- inositol 3- phosphate ($InsP_3$). This InsP3 is dephosphorylated to free InsP by an Mg^{2+} dependent inositol monophosphate phosphatase (IMP). The enzyme myo inositol kinase (MIK) can phosphorylate InsP to $InsP_1$ and this can be further phosphorylated (Sparvolli and Cominelli, 2015). The main difference between the two pathways is the way in which $InsP_3$ is formed. A lipid-independent pathway is found predominant in cereals and legumes where glucose-6-PO_4 is converted to the primary precursor $InsP_3$ and the sequential phosphorylation of the inositol ring take place by the various inositol phosphate kinases (Bhati *et al.*, 2014; Sparvolli and Cominelli, 2015). Lipid-dependent pathway involves the action of phospholipase C. The common precursor of all soluble inositol phosphates in mammalian cells is $InsP_3$, produced by the cleavage of phosphatidyl inositol 4, 5-bisphosphate, (PtdIns(4,5)P2) to $InsP_3$ and diacylglycerol by the action of phospholipase C (PLC) (Verbsky *et al.*, 2002). First, InsP is converted to PtdInsP by a phosphatidyl inositol synthase (PtdIS) and is then phosphorylated to form PtdIns(4,5)P2 which

forms the substrate for PLC. In general, a lipid dependent pathway is found to operate in all plant tissues whereas lipid independent pathway occurs mainly in seeds (Sparvolli and Cominelli, 2015).

Both lipid dependent and lipid independent pathways are found active in wheat seeds (Bhati *et al.*, 2014). In developing rice seeds, lipid-independent pathway is found to be mainly active as *OsIMP-1, OsIMP-2,* and *OsIPK2*and showed low expression levels in developing embryos, whereas in anthers both lipid dependent and lipid independent pathways are present (Suzuki *et al.*, 2007).

2.2.4 Genetics of phytic acid biosynthesis

There are a number of genes involved in the production of enzymes responsible for phytic acid biosynthesis pathway. They belong to different classes of genes which encode enzymes that catalyze the addition of pyrophosphates at various positions of inositol ring resulting in the production and bioconversion of various inositol phosphates. The first class is called Inositol Phosphate Kinases (IPKs). $InsP_6$ kinases phosphorylates the 5-position of $InsP_5$, $InsP_6$ and $InsP_7$ yielding higher PPx- InsPs (Williams *et al.*, 2015).

Myo inositol 3 phosphate synthase (MIPS) convert glucose 6 phosphate to $InsP_3$, which is the first step in InsP pathway. Inositol monophosphatase (IMP), inositol tris/tetra phosphate kinase (ITPK), inositol polyphosphate kinase (IPK2), and inositol pentakisphosphate 2-kinase (IPK1) are the other enzymes that involve in the phosphorylation of inositol ring. Inositol pentakis phosphate kinase (IPK1) is involved in the final step of phytic acid biosynthesis which is common for both lipid dependent and lipid independent pathways (Aggarwal *et al.*, 2015).

Verbsky *et al.* (2002) have reported that human $InsP_5$ kinase has a role in the synthesis of $InsP_6$ through experiments on a *Saccharomyces cereviseae ipk1* null strain. $InsP_6$ levels were lowered in transgenic potato plants for *myo*-inositol 3-phosphate synthase, (IPS) - catalyzing the first step in $InsP_6$ biosynthesis

(Murphy *et al.*, 2008). *Arabidopsis thaliana* has three (*AtIPS1-3)* IPS genes (Murphy *et al.*, 2008). Similarly, several genes involved in the phytic acid biosynthesis have been identified in common bean, *Phaseolus vulgaris* (Fileppi *et al.*, 2010). This includes *PvMIPSs* and *PvMIPSv* encoding *myo*-inositol 1-phosphate synthase, inositol monophosphatase gene *(PvIMP), myo*-inositol kinase gene *(PvMIK)*, inositol 1,4,5-tris-phosphate kinase gene *(PvIPK2)*, inositol 1,3,4-triphosphate 5/6-kinase genes (*PvITPKa* and *PvITPKb*) and inositol 1,3,4,5,6 pentakisphosphate 2-kinase gene (*PvIPK1*).

Suzuki *et al.* (2007) have identified 12 genes corresponding to the phytic acid biosynthesis from rice. *RINO1 and RINO2* are the MIPS gene, *OsIMP1* and *OsIMP2* encodes inositol monophosphatase, *OsIPK1* and *OsIPK2* are the *IPK1* and *IPK2* genes respectively, *OsITP5/6K-4,* and *OsITP5/6K-6* belong to the *IPK5/6* gene. Analyses show that *RINO1*, *OsIPK1, OsITP5/6K-4*, and *OsITP5/6K-6*, are preferentially expressed in developing embryos and aleurone layers, in which synthesis of phytic acid occurs and *RINO2, OsIPK2*, *OsIMP-1*, and-2, and *OsITP5/6K-1,- 2,-3*, and-5 do not play a major role in phytic acid biosynthesis in seeds.

Six wheat genes involved in the synthesis of inositol phosphates have been identified, including four inositol tetraphosphate kinases (*TaITPK1, TaITPK2, TaITPK3*, and *TaITPK4*), and two genes coding for inositol triphosphate kinase (*TaIPK2*) and inositol pentakisphosphate kinase (*TaIPK1*) (Bhati *et al.*, 2014).

2.2.4.1 *Myo-inositol 3-phosphate synthase (MIPS) gene*

MIPS is the primary gene involved in phytic acid biosynthetic pathway. It is involved in the first step of conversion of glucose 6- phosphate to inositol. Studies have reported that MIPS gene expresses variable loci number in different species. In both monocots and dicots, MIPS gene is found to express during the early stages of seed development, shortly before the accumulation of phytic acid and then decreases (Fileppi *et al.*, 2010). Two MIPS genes have been isolated

from common bean (*Phaseolus vulgaris*) – *PvMIPSs* and *PvMIPSv*. *PvMIPS* genes were isolated from various plant tissues and were found to share similarity between each other. The two MIPS genes isolated from leaf tissues were found to have nine introns and mainly differed in the 3rd intron region (Fileppi *et al.*, 2010). Good (2001) have observed four different MIPS genes in soybean pant tissues – *GmMIPS1, GmMIPS2, GmMIPS3* and *GmMIPS4* and was isolated from various plant tissues. These genes were found to share similarity with the *PvMIPS* genes. *TaMIPS* gene with a total length of 3910 bp and the cDNA with an ORF of 1533 bp, coding for 510 amino acids were found to express in roots, stems, flag leaves and immature seeds at 15 day after flowering and was found to have four conserved domains similar to other crop species (Ma *et al.*, 2013).

2.2.4.2 *Inositol monophosphatase (IMP gene)*

IMP is involved in the myo-inositol salvage pathway. It helps the dephosphorylation of Ins3P to inositol. Comparison of *IMP* genes in different species have shown that dicots share a 12 exon- 11 intron structure whereas monocots have a 10 exon- 9 intron organisation and monocot exon 1 and exon 3 are equivalent to dicot exon 2 and exons 4 & 5 (Fileppi *et al.*, 2010).

2.2.4.3 *Inositol Phosphate Kinase2 (IPK2) gene*

IPK2 is a dual 6-/3-kinase and is specific to the lipid dependent pathway, therefore not a major contributor to the phytic acid content of seeds (Sparvolli and Cominelli, 2015). When assessed during the seed development, IPK2 was found to accumulate later than MIPS and IMP. In Arabidopsis, there are 2 homologues of IPK2- *AtIPK2α* and *AtIPK2β*. *AtIPK2α* is found to have an inhibitory effect on pollen germination and root growth and *AtIPK2β* functions in auxin signalling pathway (Xu *et al.*, 2005; Zhan *et al.*, 2015). Ajay *et al.* (2016) have isolated IPK2 gene homologues from peanut embryo and flower buds. The *AhIPK2* genes have an open reading frame (ORF) of 891bp and shows high similarity to other plant genes. *AhIPK2* have shown to share high similarity with *Aradu.24V9G* of *A. duranensis* and contained three exons with 5' and 3' UTR's on either side. Unlike

other IPK2 genes, *AhIPK2* possess conserved domains such as PxxxDxKxG and [L/M][I/V]D[F/L][A/G][H/K]. Two genes corresponding to IPK2 are found in wheat culture cells and shows similarity with the homologous genes of other members of the same family (Bhati *et al.*, 2014).

2.2.4.4 *Inositol tris/tetra phosphate kinase (ITPK) gene*

ITPK is found to cluster in three phylogenetic subgroups- α, β,γ. The ITPK group is found to have multiple substrate specificity and may also act as both isomerase and/or phosphatases of various substrates (Sparvolli and Cominelli, 2015). A similar report by Caddick *et al.* (2008) states that ITPK gene product from *Solanum tuberosum* displayed inositol 1,3,4,5,6-pentakisphosphate 1-phosphataseactivity in the absence of a nucleotide acceptor and inositol 1,3,4,5,6-pentakisphosphate–ADP phosphotransferase activity in the presence of physiological concentrations of ADP. The members of ITPK belong to the family of ATP- grasp fold proteins. Four ITPK genes are identified in wheat, and multiple sequence alignment with the maize (*ZmIP5/6K*), Arabidopsis (*AtITPK1*), and soybean (*GmITPK1-4*) genes have shown the presence of several motifs that define conserved regions among different kinases (Bhati *et al.*, 2014). Sequence analysis shows that *AhITPK1* have only a single uninterrupted exon and a cDNA fragment of 1146bp was amplified by PCR representing a putative 382-amino acid chain. Several conserved regions are found between 45 and 107 amino acid residues for *AhITPK1*. Phylogenetic analyses reveal that *ITPK1* genes are grouped into five different clusters (Ajay *et al.*, 2016).

2.2.4.5 *Inositol 1, 2, 3, 4, 5- pentakis phosphate kinase (IPK1) gene*

IPK1 is involved in the final step of phytic acid biosynthetic pathway. The enzyme catalyses the conversion of inositol pentakis phosphate into inositol hexakis phosphate by phosphorylating the inositol rings. Sun *et al.* (2007) have reported IPK1 homologues in maize. The cDNA sequence is found to have a size of 2012 bp and a predicted ORF of 440 aminoacids. Alignments of the leaf- and seed derived cDNA sequences with *ZmIPK1A* defined nine predicted exons and

eight predicted introns in the *ZmIPK1* genomic sequence. A single IPK1 gene is expressed by many crops. *Zea mays* have two IPK1 genes- *ZmIPK1A* which is expressed in a range of tissues including immature ears, seeds at 12 DAF, endosperm and developing embryos and *ZmIPK1B* is expressed particularly in roots (Sun *et al.*, 2007; Sparvolli and Cominelli, 2015).

2.2.4.6 *Low phytic acid (lpa) mutants*

Even though phytic acid has got many uses, the compound exhibits an anti nutritional effect on humans and animals. It forms complexes with minerals like K^+, Mg^{2+}, Ca^{2+}, Mn^{2+}, and Zn^{2+} etc. by chelating with these ions. Phytin thus causes Ca, Zn, P and other micronutrient deficiencies and is a significant problem in countries where legumes or grain crops contribute a major part of the diet (Bohn *et al.*, 2008). Phytate being indigestible by the monogastric animals and poultry gets accumulated in the environment leading to environmental pollution and eutrophication (Vats *et al.*, 2005; Han *et al.*, 2001). Therefore efforts have been made to lower the phytate levels in crops. This can be done either by up regulating the phytase activity that can result in degradation of phytate levels in plants or by introducing low phytic acid mutant plants. Phytase is the enzyme that can release phosphate from phytic acid. *lpa* mutants can be generated either by the over expression of phytase gene or by down regulating the gene family involved in phytic acid biosynthetic pathway. Kanani and Poonam (2015) have mentioned various methods to decrease the PA level in plants through the introduction of phytase gene. Biotechnological approach involves the identification, isolation and characterisation of genes, cloning, sequencing and manipulation of genes by various genetic engineering techniques to produce transgenic plants.

Several *lpa* mutants have been isolated in some of the important crops, such as barley, maize, rice, wheat, soybean, common bean and pea. Investigations prove that all PA in seeds are not necessary for seedling establishment and therefore *lpa* mutants of barley (Larson *et al.*, 1998; Dorsch *et al.*, 2003), maize (Shi *et al.*, 2003), soybean (Hitz *et al.*, 2002), wheat (Guttieri *et al.*, 2004), rice

(Larson *et al.*, 2000; Ali *et al.*, 2013; Kuwano *et al.*, 2009), common bean (Campion *et al.*, 2009) and arabidopsis (Stevenson-paulik *et al.*, 2005) were developed and was found to exhibit no reduction in seedling growth. Depending on the affected step in the biosynthetic pathway, these *lpa* mutants can be classified into three classes. Type1 refers to affected gene functions in MIPS, MIK and IMP. Type 2 affects 2PGK, IPK2, ITPK and IPK1 genes. Type3 mutants are affected in MRP and putative sulphate transporter genes (Sparvolli and Cominelli, 2015).

Ali *et al.* (2013) have reported RNAi-mediated seed-specific silencing of the IPK1 gene using the Oleosin18 promoter. The results revealed a 3.85-fold down-regulation in IPK1 transcripts, and a corresponding significant reduction in phytate levels as well as an increase in the amount of inorganic phosphate (Pi). The low-phytate rice seeds also accumulated 1.8-fold more iron in the endosperm due to the decreased phytic acid levels. They also found that no negative effects were observed on seed germination or in any of the agronomic traits examined.

Kuo *et al.* (2014) have reported that an incomplete loss-of-function *Arabidopsis* mutant, *atipk1-1*, exhibited disturbed Pi homeostasis and over accumulated Pi as a consequence of increased Pi uptake activity and root-to-shoot Pi translocation. Also, myo-inositol-1-phosphate synthase mutants, *atmips1* and *atmips2* with comparable reduction in vegetative $InsP_6$ to that in the *atipk1-1* mutants did not over accumulate Pi, suggesting that Pi homeostasis modulated by *AtIPK1* is not solely attributable to $InsP_6$ level. This study revealed that *AtIPK1* had important roles in growth and Pi homeostasis.

2.3 GENOME WALKING FOR UNKNOWN SEQUENCES

The conventional method of genome walking uses extensive screening of genomic libraries using labelled probes, which consumes considerable amount of time, effort and expenses. PCR mediated method of genome walking is now widely used for the identification of unknown sequence flanking the known

regions of DNA. There are different methods of genome walking that can be classified into three main categories depending on the strategy undertaken. Restriction based genome walking (R-GW) is carried out by a primary restriction digestion of genomic DNA using restriction enzymes that are situated at suitable distances from the target genes. These restriction fragments that are obtained can then be self-circularized or ligated to specific adaptors. Primer based genome walking (P-GW) is another method that is characterised by the use of a combination of primers including degenerate/random primers as well as sequence specific primers (Leoni *et al.*, 2011). Extension based genome walking (E-GW) make use of the single stranded DNA obtained by the extension of sequence specific primers and subsequent 3'- tailing as the substrate for further PCR amplification (Volpicella *et al.*, 2012).

The advances in high throughput sequencing technologies have now resulted in the application of Next Generation Sequencing (NGS) techniques to the existing genome walking methods. This has helped in the analysis of large genomic libraries. Review on genome walking by next generation sequencing by Volpicella *et al.* (2012), suggests that pyrosequencing by Roche 454 platform and Sequencing by Synthesis using illumina techniques are only used for the genome walking and, other NGS techniques like Sequencing by Oligo Ligation Detection (SOLiD) and SBS approach known as ion torrent can also be potentially used for genome walking (GW). Also, analyses of various studies have shown that pyrosequencing is exclusively used for eukaryotes while illumina-SBS method is used for both eukaryotic and prokaryotic genomes.

2.3.1 Phi29 mediated genome walking using Rolling Circle Amplification

Phi29 mediated genome walking is a non-PCR based WGA method that uses Strand Displacement Amplification (SDA)/ Multiple Displacement strand Amplification (MDA) by rolling circle replication principle of phi29 DNA polymerase. SDA was first described in 1992. It is an isothermic amplification method in which, in a single reaction, 10^9 copies of target DNA can be produced

in less than an hour. Multiple strand Displacement Amplification is an isothermal, strand displacing method based on the use of highly processive and strand displacing DNA polymerase in conjunction with modified random primers to amplify the entire genome with high fidelity. It has been developed to amplify all DNA in a sample from a very small amount of starting material. Multiple strand displacement amplification by phi29 DNA polymerase involves incubating phi29 DNA polymerase, dNTPs, random hexamers and denatured template DNA at 30°C for 16 to 18 h. The enzyme is inactivated at 65°C for 10 min and the product DNA can be used directly in downstream applications. The DNA yield and sensitivity of the method is also found very high (Fakruddin *et al.*, 2013).

Phi29 DNA polymerase is encoded by the genome of bacteriophage Phi29 which infects *Bacillus subtilis*, for its replication. Phi29 polymerase is a member of the B- family of replicative polymerases and has a molecular weight of 68 kDa (Berman *et al.*, 2007). Salas, 1991 cited by Berman *et al.*, 2007 reported that in *in-vivo* replication, phi29 polymerase initiates replication using a protein primer - terminal protein. Phi29 polymerase has got a 5'-3' polymerase activity as well as 3'-5' exonuclease activity. This provides the enzyme a high fidelity in proof reading and thus makes the replication error proof and suitable for genetic studies and the error rates are calculated to be between 10^{-5} and 10^{-6} in biochemical assays, or 9.5×10^{-6} by direct sequencing of 500,000 bp after phi29 polymerase-mediated amplification (Esteban *et al.*, 1993 cited by Johne *et al.*, 2009; Paez *et al.*, 2004). Also the isothermal reaction conditions of polymerase enzyme help to avoid the use of thermo cyclers for controlled temperatures and ensure replication of genome at normal temperatures of 30°C. Experiments have also found that phi29 DNA polymerase can be used for amplification of plant genomes that contain hyper methylated regions as phi29 polymerase can provide an efficient and representative amplification of the genomic DNA that is methylation free. This made the heavily methylated genomic DNA region amenable to physical mapping with any restriction enzyme (Adachi *et al.*, 2004).

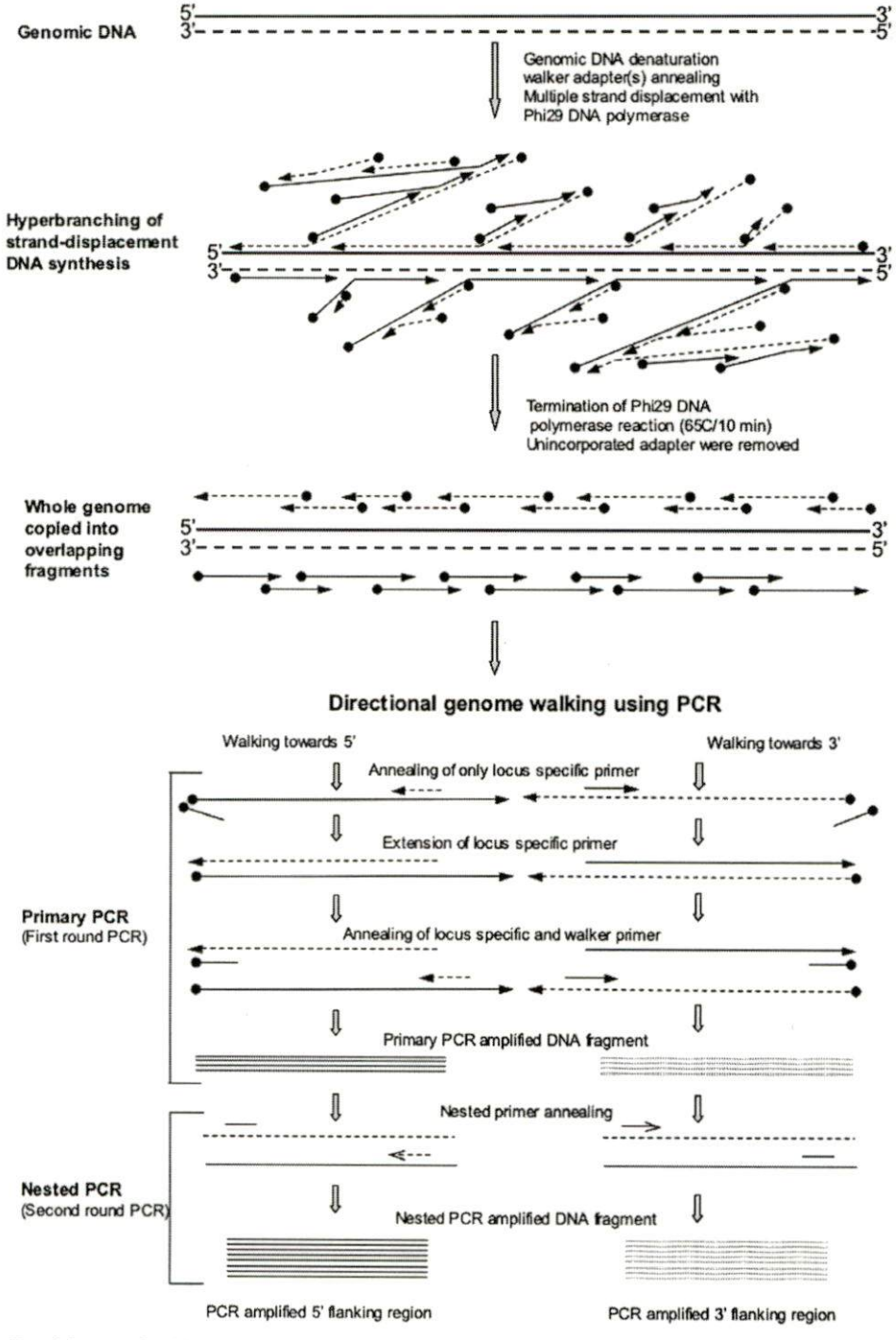
A Whole genome amplification (WGA) method should essentially be able to cover the whole genome, generate long fragments, successfully amplify small amounts of template DNA and should have high amplification fidelity and yield (Gadkar and Rilling, 2005). Experiments have shown that phi29 mediated amplification have successfully copied 99.82 percent of the total genome in normal and cancerous human cells (Paez *et al.*, 2004). Thus, phi29 mediated WGA can be utilized in amplification of clinical samples with less genome content and a larger genome size and therefore be used in detection of SNP variation if present. This will help in clinical studies, especially in cancer diagnosis in humans.

Similarly, phi29 based WGA is found useful in evolutionary and taxonomic studies. Raghunathan *et al.* (2005) have successfully amplified the genome from the cells of *E.coli* and *Myxococcus xanthus* using MSD property of phi29 polymerase. They have got five billion- fold amplification from single cells of the bacterium and have successfully sequenced a 66 bp segment of the 16S rRNA gene and is found very much helpful in the evolutionary and ecological studies of uncultured micro-organisms. They also successfully amplified the C-methyl transferase gene in *M. xanthus*.

Gadkar and Rilling (2005) have successfully amplified the whole genome of Arbuscular mycorrhizal (AM) fungi *Glomus* and *Gigaspora* species from nanogram quantities of DNA (ranging from 0.05 to 0.61 ng) present in a single fungal spore using phi29 DNA polymerase mediated SDA. The genome coverage of the method was analysed by PCR amplification of β tubulin1 gene and part of the rDNA cluster present in AM fungi. Sequencing of *Glomus intraradices* ITS1 variants to detect the four variants that were already identified in the fungus was carried out to evaluate the fidelity of SDA method.

The experiments by Reddy *et al.* (2008), have introduced a method of genome walking making use of the rolling circle model of replication by phi29

DNA polymerase. The protocol uses introduction of unique walker primer binding sites randomly into different regions of genome using modified walker adapter primers instead of the random hexamer primers commonly used. Strand displacement property of phi29 polymerase produces large overhanging fragments upto 100 kb in size compared to 10-100 bp fragments produced in other WGA. Directional genome walking is carried out using the locus specific primer designed based on the target gene and walker primers and their corresponding nested primers. The promoter regions of different genes of *Pennisetum* and *Salicornia* plant genome were amplified in the study to test the efficiency of method.



(Reddy *et al.*, 2008)

Studies conducted by Gadkar and Filion (2011), have made use of CircLigase, with a circularization property. PCR amplified flanking regions were

circularized with CircLigase and this was further used for the rolling circle amplification using Phi29 polymerase to form linear overlapping fragments that are used for the directional genome walking. Using *hncB* and *phlD* genes of *Pseudomonas sp.* LBUM300, 1kb of flanking region was successfully isolated. The method needs only 100-150 bp of prior sequence information and can be used for any microbial genome.

2.3.2 Applications of genome walking

Genome walking, due to its high potential to amplify large genomes and its capacity to produce reliable results from low initial quantity of genetic material can be used in a wide range of genetic studies. This can be broadly classified into- studies related to insertional mutagenesis and *de novo* sequencing. Both prokaryotic and eukaryotic genomes are sites of insertion by virus, transposons or T-DNA. This can be used in the field of gene marking, gene therapy and retro-viral studies (Leoni *et al.*, 2011). Rill *et al.* (1994) first described the use of GW for the retro- viral mediated gene marking of bone marrow cells of patients with neuroblastoma. Leoni *et al.* (2011) have reviewed that GW approaches can be successfully used in the study of retro-viral integration in genome. Human HIV integration sites are studied by this method. Similarly, Inverse PCR mediated genome walking is used in the identification of *Ac/Ds* transposon mutations in various plants including maize. Phi29 mediated genome walking is used in amplification and analysis of viral genomes that cannot be cultured, both circular and linear ssDNA/dsDNA like the families of circular viruses affecting plant systems (caulimoviridae, geminiviridae and nanoviridae), causing diseases like mosaic and banana bunchy top symptoms (Johne *et al.*, 2009). The whole genome amplification (WGA) methods can be most effectively utilized for the forensic studies where the initial DNA quality and quantity is very low (Lasken and Egholm, 2003). *De novo* sequencing by GW is largely used in order to identify the regulatory sequences and the multigene families rather than for identifying the genes. Coupling with the NGS programmes, GW is also used to find out nucleotide modification or mutation (SNPs) in the genomes. It can also be used in

42

the evolutionary and ecological studies of micro-organisms that cannot be cultured (Raghunathan *et al.*, 2005) as the method can compare the consensus or conserved sequences present in the whole genome of organisms.

Chawla and DeMason (2003) have reported the cloning of promoter sequences of pea genes using restriction- ligation mediated genome walking strategy. Whole genome of pea is restricted using a set of restriction enzymes and libraries are created after adaptor ligation. Specific primers are designed based on the genes and nested PCRs are carried out. They have identified upstream regions for 5 pea genes using this technique. Two of the genes are *PsPINl* and *PsPK2*. With two sets of walks, they have obtained 1500 bp of *PsPINl* and 2800 bp of *PsPK2* upstream regions. A 2405 bp of the Unifoliata (AF035163) gene promoter, PsArgonautel promoter (589 bp) and PsArgonaute2 promoter (858 bp) regions were also obtained.

Guo and Xiong (2006) have described a nested –PCR based genome walking method where partially degenerate primers are used as walker primers and multiple nested PCRs are carried out using specific nested primers. Four walker adapters are used to carry out parallel reactions in order to increase the success of walking. Known genomic loci of two highly divergent photosynthetic organisms, *Rhodobacter capsulatus* and *Heliophilum fasciatum* have been amplified successfully.

An alternative genome walking method is reported by Leoni *et al.* (2008) to identify the sequence flanking the Lhcb1.1 gene of the multigene family gene coding for light harvesting protein in spinach. Method involved extension by a sequence specific primer, C- tailing of the extended strand and nested PCR.

To identify possible variants of ChiA and P1 genes present in the genomes of maize plants of different origins, Volpicella *et al.* (2012) have reported a modification of an already established genome walking strategy in which primer

43

extension is driven by Klenow polymerase and the products can be directly sequenced. This method is found effective for the complex eukaryotic genomes.

2.4 QUANTIFICATION OF $InsP_6$

Although the Pi is limited to soil, the plant system has got a stored reserve of $InsP_6$-P which can be used in times of phosphate deficiency. But, over accumulation of Pi in crop plants is found to have an anti-nutrient effect when consumed as human and animal diets. Therefore the detection and quantification of $InsP_6$ in plants is necessary.

Several methods have been used for the detection of various InsPs. One of the common methods used is to introduce radiolabeled precursor molecule- $InsP_3$ or $InsP_6$ followed by high performance liquid chromatography (HPLC) separation. A non-radioactive chromatographic method based on metal dye detection HPLC (MDD-HPLC) can also be used. Thin layer chromatography is another method. But these are labor intensive (Williams *et al.*, 2015).

Greiner *et al.* (2006) have reported that phytate in experimental diets are found to be determined by the addition of excess Fe to precipitate phytate and further estimated by determining the inositol or phosphate contents in phytate in the earlier times. High performance liquid chromatography (HPLC) techniques have been introduced into phytate determination more recently and ion-pair reverse-phase chromatography allows the simultaneous separation and quantification of *myo*-inositol tris- to hexakisphosphates whereas anion-exchange chromatography quantifies *myo*-inositol mono- to hexakisphosphates.

Smernik and Dougherty (2007) have reported that Phosphorus-31 (^{31}P) nuclear magnetic resonance (NMR) spectroscopy of sodium hydroxide–ethylene diamine tetraacetic acid (NaOH-EDTA) extracts has become a widely used technique for the characterization of soil P and this technique can be used for the easy identification and quantification of phytate (*myo*-inositol hexakisphosphate).

44

Phytate is usually identified by its characteristic pattern of four resonances in the ratio 1:2:2:1. But there is a possibility that phytate concentrations may be overestimated using this technique due to mis-assignment of phytate resonance in the NMR spectra.

Sodium dodecyl sulphate- polyacrylamide gel electrophoresis (SDS-PAGE) is commonly used for the analysis of proteins and native PAGE is used for the nucleic acids. Alimohammadi *et al.* (2013) have reported an easy, fast and reliable method that is cost effective for the determination of $InsP_6$ in plants using native PAGE.

Materials And Methods

3. MATERIALS AND METHODS

3.1 MATERIALS

3.1.1 Plant materials

Black pepper variety Panniyur 1 maintained in the 'Hi-Tech Pepper Nursery' under the Department of Plant Biotechnology, College of Agriculture, Padannakad was utilized for the isolation of DNA as well as getting the crude extract of phytic acid for Polyacrylamide Gel Electrophoresis (PAGE) analysis. Young leaves from healthy vines were collected in early morning for the extraction.

3.1.2 Laboratory chemicals

3.1.2.1 *DNA isolation*

Liquid nitrogen, PVP, β-mercapto ethanol, chloroform isoamyl alcohol (24:1), isopropanol, 5 M NaCl, 70% ethanol, CTAB extraction buffer (100 mM Tris HCl - pH 8.0, 2 mM EDTA - pH 8.0, 1.4 M NaCl, 2% CTAB).

3.1.2.2 *Agarose gel electrophoresis*

Agarose, 6X gel loading dye, ethidium bromide (0.5µg/ml), DNA ladder of various size, 1X TAE buffer (40 mM Tris, 20 mM glacial acetic acid, 1 mM EDTA).

The chemicals used for DNA isolation and agarose gel electrophoresis were obtained from Merck India Ltd., SRL laboratories and Himedia.

3.1.2.3 *PCR amplification*

The reaction mixture (25µL) was constituted with 100 ng DNA (0.5µL), 5U Taq DNA polymerase (0.3µL, takara); 10X PCR buffer (A) with $MgCl_2$ (2.5µL, takara) and 100mM dNTPs (0.5µL, Takara), 10µM primers (0.5µL each) Integrated DNA Technologies, U.S).

47

3.1.2.4 *Rolling Circle Amplification*

10 U Phi 29 DNA polymerase (1µL, New England Biolabs), Bovine serum albumin (10mg/ml, 0.25µL), 10X phi29 polymerase buffer (2µL, New England Biolabs), dNTPs (10mM, 1µL, Takara), 100 ng walker adapter primers (1.5µL, Integrated DNA Technologies).

3.1.2.5 *Gel extraction of PCR product*

Gel elution kit (Thermofischer Scientific, Cat. No. K0691), isopropanol, absolute ethanol.

3.1.2.6 *PAGE detection of phytic acid*

Liquid nitrogen, 0.4M HCl, 40 % acrylamide/ bis acrylamide, 16 % acrylamide/ bisacrylamide, 10X TBE pH 8.0, 5X TBE pH 6.0, double distilled sterilized water, 10 % Ammonium persulphate (APS), tetramethylethylenediamine (TEMED), 6X orange G gel loading dye (Genei, cat No. 2602680061730), toluidine blue (Sigma Aldrich, cat no.T3260), phytate dipotassium salt (Sigma Aldrich, cat. no. P5681), methanol (20%), glycerol.

3.1.3 Softwares and databases used

Oligocalc (www.basic.northwestern.edu/biotools/OligoCalc.html)

ClustalW (http://www.ebi.ac.uk/Tools/msa/clustalw2/)

BioEdit sequence alignment tool (http://www.mbio.ncsu.edu/bioedit/bioedit.html)

National Centre for Biotechnology Information (NCBI)

(https://www.ncbi.nlm.nih.gov/)

BLAST (https://www.ncbi.nlm.nih.gov/blast/)

ORF Finder (https://www.ncbi.nlm.nih.gov/orffinder/)

3.1.4 Equipment and machinery used

The equipment available in the Department of Plant Biotechnology, College of Agriculture, Padannakad were used for the study. DNA quantification

48

and quality check were carried out using biophotometer (Eppendorf, Germany), polymerase chain reaction was done with the Master Cycler (Eppendorf, Germany) and gel documentation of both agarose gels and polyacrylamide gels was performed using Geldoc XR imaging system (Bio-Rad, USA). .

3.2 METHODS

3.2.1 DNA isolation

DNA was isolated from the tender leaves of variety Panniyur 1 using the modified CTAB method (Doyle and Doyle, 1987). Young leaves were collected from the healthy vines, cleaned using 70 percent ethanol and the midribs were removed. One gram of the leaf sample was weighed and transferred to a pre-chilled mortar and pestle. The sample was ground using liquid nitrogen and a pinch of PVP was added. The finely powdered sample was transferred to an Oakridge tube containing 10 ml of pre-warmed CTAB extraction buffer, followed by the addition of 40 µL of β-mercapto ethanol. The homogenate was mixed well and incubated at 65 °C for 30 minutes in a water bath. The sample was mixed intermittently by inverting the tube. After the incubation, 4.0 ml of chloroform isoamyl alcohol was added to the tube, mixed well and the tube was centrifuged at 10,000 rpm for 15 minutes at 4.0 °C. The supernatant was taken and transferred to another tube and chloroform isoamyl alcohol extraction step was repeated. The supernatant was again transferred to a fresh tube and DNA was precipitated using one and half times volume of chilled isopropanol and half volume of 5 M NaCl. The tube was incubated at -20 °C for 30 minutes and centrifuged at 10,000 rpm for 10 minutes at 4 °C. The supernatant was removed, DNA pellet was washed using 70 per cent ethanol, air dried, dissolved in 100 µL distilled water and stored at -20°C.

3.2.2 Quality and quantity of genomic DNA

The DNA samples were analyzed for its quality and quantity in a 0.7 per cent agarose gel by electrophoresis at 80-100 V for 1 h. The gel was visualized under UV light using the BioRad Geldoc XR gel documentation system and

analyzed for any shearing of DNA. The concentration was approximately checked by comparison with the concentration standard loaded along with the samples in 0.7 per cent agarose gel.

The steps involved were:

1. The gel casting tray and comb were cleaned by washing with distilled water and were again wiped with distilled water. The open ends of the tray were sealed with cello tape and the comb was placed. The tray was kept on a level surface.
2. For 100ml of 1X TAE buffer, 0.7 g agarose has been taken and heated in a microwave oven until the agarose was completely dissolved. The gel was then cooled to 45 ^{o}C and ethidium bromide (0.5 µg/ml) was added and mixed well.
3. The solution was slowly poured into the casting tray fitted with comb, without making any air bubbles and allowed to solidify under room temperature. Once the gel was solidified, the comb was removed gently after pouring a little amount of buffer.
4. The casting tray with the gel was slowly immersed into the gel tank after removing the cello tape from both ends. The electrodes were connected to the power pack. Sample DNA (2 µL) was loaded along with 1 µL of 6X gel loading dye into each well. One microlitre standard DNA fragment with 100 ng/µL concentration was used to compare the intensity of the bands produced for estimation of the quantity of DNA in the samples.

The quantification of DNA was further carried out using the biophotometer. The concentration readings as well as the quality check parameters were noted down and analysed. The quality of DNA was checked using the $A_{260/280}$ and $A_{260/230}$ readings.

The steps involved were:

1. The 'BLANK' value was set by inserting the cuvette with 50 µL of sterile distilled water

2. The DNA stock was diluted 10 times by adding 5µL of DNA into 45 µL of distilled water
3. The cuvette with sample was inserted and the absorbance was read.
4. Concentration of the sample (µg/ml) and $A_{260/280}$ values were recorded.

3.2.3 Primer designing

Primers were designed for the amplification of inositol pentakisphosphate 2-kinase (*ipk1*) gene sequence in black pepper. Different primer designing strategies were undertaken for the various steps in the study. For the confirmation of sequence information of the unigene *Pnc135*, which was the preliminary step in the study, two sets of primers were designed based on the unigene sequence. Simultaneously, for the amplification of flanking regions of the unigene sequence, two sets of primers were designed based on the pairwise alignment of the sequences of selected crops. A partial degeneracy was applied to the primers that were designed based on the pairwise aligned data. Later, three sets of primers were designed for the amplification of the *ipk1* gene sequence based on the conserved regions obtained by the multiple sequence alignment of unigene *Pnc135* with the selected sequences of 28 different crops. Finally, based on the sequence results, two sets of forward and reverse primers and their corresponding nested primers were designed for genome walking for the flanking regions. The primers used in the rolling circle amplification and the walker primer and nested walker primer in the genome walking were the same as reported in the protocol by Reddy *et al.* (2008).

The nucleotide sequences of different primers were provided under section 'Results'.

3.2.3.1 *Primer designing for confirmation of unigene sequence information*

The primers were designed based on the sequences of unigene *Pnc135* which showed similarity to phytic acid gene reported from many other crops, obtained through the homology search in GenBank. Forward and reverse primers for the unigene were designed in order to check the amplification and confirm the

sequence information. Two sets of primers were designed- 2 forward primers (PNC F1, PNC F2) from the 5' end of the sequence and 2 reverse primers (PNC R1, PNC R2) from the 3' end of the sequence. The primer designing and analysis were carried out using the software Oligocalc (www.basic.northwestern.edu/biotools/OligoCalc.html).

3.2.3.2 *Primer designing for amplification of the flanking region of unigene using selected homologous sequences*

The original unigene sequence of *Pnc135* was used for the sequence similarity search using NCBI BLASTn (https://www.ncbi.nlm.nih.gov/blast/). The BLAST homology search results with 68 per cent or more similarity with the query sequence were selected and analyzed for the primer designing of flanking regions of the unigene. The query covers as well as the similarity percentages were analyzed using the BLAST homology search tool and the homologous sequences were aligned using NCBI BLAST sequence alignment as well as Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalw2/) Multiple Sequence Alignment programmes. Primers were designed based on the consensus sequence obtained through the pair wise alignment of sequences (PNC F3/ PNC R3 for amplification of 5' flanking region and PNC F4/ PNC R4a, PNC R4b for 3' flanking region).

Sequences of *Vitis vinifera, Nelumbo nucifera, Sesmum indicum, Cicer arietinum, Camelina sativa* were selected based on BLAST homology search results.

3.2.3.3 *Primer designing for the amplification of inositol pentakisphosphate 2-kinase gene by multiple sequence alignment*

Primers were designed for the amplification of *ipk1* gene using homologous sequences of the same gene reported from different crop species. *ipk1* gene sequences from 28 different crop species belonging to different families (Table 1) were selected for multiple sequence alignment with the unigene sequence using the CLUSTALW MUSCLE program in order to obtain the

conserved or consensus sequences present. Primers were designed based on the consensus sequences using the software OligoCalc. A total of six primers- 3 forward (PNC F5, PNC F6, PNC F7) and 3 reverse primers (PNC R5, PNC R6, PNC R7) were designed based on the multiple sequence alignment, which were used in all combinations.

Table 1. List of crops selected for multiple sequence alignment to design primers for amplification of *ipk1* gene in black pepper

Description	Accession number
PREDICTED: *Daucus carota subsp. sativus ipk1*, mRNA	XM_017398551.1
PREDICTED: *Eucalyptus grandis ipk1*, mRNA	XM_010062987.2
PREDICTED: *Capsicum annuum ipk1*, mRNA	XM_016713042.1
Glycine max ipk1, mRNA	NM_001250504.1
PREDICTED: *Musa acuminata ipk1*, mRNA	XM_009391905.2
PREDICTED: *Ipomoea nil ipk1*, mRNA	XM_019296805.1
PREDICTED: *Citrus sinensis ipk1*, mRNA	XR_001507770.1
Medicago truncatula ipk1, mRNA	XM_013606692.1
PREDICTED: *Sesamum indicum ipk1*, mRNA	XM_020695367.1
PREDICTED: *Phoenix dactylifera ipk1*, mRNA	XM_008797736.2
PREDICTED: *Elaeis guineensis ipk1*, mRNA	XM_010941805.2
PREDICTED: *Cicer arietinum ipk1*, mRNA	XM_004500934.2
PREDICTED: *Vitis vinifera ipk1*, mRNA	XM_010666413.2
PREDICTED: *Nicotiana tabacum ipk1*, mRNA	XM_016582667.1
PREDICTED: *Vigna angularis ipk1*, mRNA	XM_017572635.1
PREDICTED: *Ricinus communis ipk1*, mRNA	XM_015720256.1
PREDICTED: *Cajanus cajan ipk1*, mRNA	XM_020382746.1
PREDICTED: *Populus euphratica ipk1*, mRNA	XM_011050491.1
PREDICTED: *Jatropha curcas ipk1*, mRNA	XM_020679162.1
PREDICTED: *Ziziphus jujube ipk1*, mRNA	XM_016010604.1
PREDICTED: *Citrus sinensis ipk1*, mRNA	XR_001507770.1
PREDICTED: *Gossypium hirsutum ipk1*, mRNA	XM_016893564.1
PREDICTED: *Theobroma cacao ipk1*, mRNA	XM_018126167.1
PREDICTED: *Solanum tuberosum ipk1*, mRNA	XM_015306404.1
PREDICTED: *Tarenaya hassleriana ipk1*, mRNA	XM_010542029.1
PREDICTED: *Raphanus sativus ipk1*, mRNA	XM_018596688.1
PREDICTED: *Brassica oleracea ipk1*, mRNA	XM_013751201.1
PREDICTED: *Camelina sativa ipk1*, mRNA	XM_010495936.2
Arabidopsis thaliana ipk1, mRNA	NM_123646.4

3.2.3.4 *Primer designing for genome walking*

Primers were designed for conducting genome walking for the flanking regions of the unigene using the information obtained from the sequence analysis of the amplicon (F6R5) obtained from the primer set PNC F6/PNCR5. The four walker adapter primers (WA 1, WA 2, WA 3 and WA 4) reported by Reddy *et al.* (2008) were used to generate fragments of genomic DNA with walker adapter sequence attached to 5' end of each fragment. A walker primer (WP1) and nested walker primer (WP2) specific to the walker adapter sequence were also the same as that referred in the protocol by Reddy *et al.*, (2008). Two sets of locus specific primers (LSP) each were designed at the 5' and 3' end of F6R5 sequence. LSP R1 and LSP R2 were designed at the 5' end & LSP F1 and LSP F2 were designed from the 3' end. For each locus specific primer, a corresponding nested primer (NSLP) was also designed (NSLP R1, NSLP R2, NSLP F1 and NSLP F2).

3.2.4 PCR amplification

PCR amplification was carried out with the specified primers using the DNA extracted from Panniyur 1 variety of black pepper. The reaction volume was set to 25 µL. The final reaction mix was set up with the components: 1X PCR buffer with $MgCl_2$ (2.5 µL), 200 µM dNTPs (0.5 µL), 0.05 µM each forward and reverse primers (0.5 µL), 1U Taq polymerase (0.3 µL) and 50 ng DNA template (0.5µL).

3.2.4.1 *PCR amplification for confirmation of unigene Pnc135 sequence information*

PCR amplification was carried out for the confirmation of sequence information using primer sets F1R1 and F2R2 with the PCR components in the concentration mentioned.

The thermal profile used was:

Initial denaturation	94 °C for 2 min.	
Denaturation	92 °C for 1 min.	40 cycles
Annealing	46 to 60 °C for 1 min.	
Extension	72 °C for 1.5 min.	
Final extension	72 °C for 10 min.	
Hold	4 °C.	

3.2.4.2 *PCR amplification for the flanking regions of unigene*

Flanking regions of the unigene *Pnc135* were amplified using primers PNC F3/PNC R3, PNC F4/PNCR4a and PNC F4/PNC R4b. PNC F3/PNC R3 was used for the amplification of 5' end of the unigene and PNC F4/PNC Ra and PNC F4/PNC R4b were used to amplify the 3' end of the unigene sequence. The thermal profile used was:

Initial denaturation	94 °C, 2 min.	
Denturation	94 °C, 1 min.	40 cycles
Annealing	48-62 °C, 1 min.	
Extension	72 °C, 50 sec.	
Final extension	72 °C, 10 mins.	
Hold	4 °C.	

3.2.4.3 *PCR amplification of inositol pentakis phosphate 2-kinase gene by multiple sequence alignment*

PCR amplifications were carried out in all combinations of the three sets of primers. Primer combinations PNC F5/PNC R5, PNC F5/PNC R6, PNC F5/PNC R7, PNC F6/PNC R5, PNC F6/PNC R6, PNC F7/PNC R7, PNC F7/PNC R5, PNC F7/PNC R6 and PNC F7/PNC R7 were tried.

The thermal profile used was:

Initial denaturation	94°C for 2 mins.	
Denaturation	92°C for 1 min	40 cycles
Annealing	48-62°C for 1 min	
Extension	72°C for 1 min	
Final extension	72°C for 10 mins	
Hold	4°C.	

3.2.4.4 *Rolling circle amplification*

Phi29 mediated isothermal amplification was used for random introduction of walker primer binding sites in the genomic DNA of Panniyur 1 by taking advantage of the rolling circle amplification (RCA). The reaction resulted in the synthesis of long overlapping fragments of the genomic DNA of Panniyur 1 with the walker adapter sequence attached to 5' end of all fragments. For this, the walker adapter primers reported by Reddy *et al.* (2008) were used.

RCA reaction was set up in 20 µL volume with 1X Buffer (2 µL), 200 µM dNTP mix (1 µL), 150 ng WA primers (1.5 µL), 100 ng genomic DNA (1 µL), 10 U phi29 polymerase (1 µL), 125 ng/µL Bovine serum albumin (BSA) (0.25 µL), and 13.25 µL double distilled water.

Four different walker adapter primers (WA1, WA2, WA3, and WA4) reported by Reddy *et al.* (2008) were used for the RCA reaction.

Method1

Steps:

1. Each of the walker adapter primers were added to four separate PCR tubes containing 1X buffer, template DNA and double distilled water.
2. DNA was heat-denatured at 94°C for 5 minutes.
3. The tubes were immediately transferred to ice and dNTPs, phi29 polymerase and BSA were added.

4. Reaction mix was incubated at 30°C for 1 hour 30 minutes.
5. Reaction was terminated by incubating at 65°C for 10 minutes.
6. RCA product was stored at -20°C.

Method 2:

The reaction was set up as in method 1 and in step 4, the tubes were incubated at 30°C for overnight.

3.2.4.5 *Locus specific amplification*

Flanking regions of F6R5 was amplified using directional genome walking using a walker primer and locus specific primer as per the protocol given by Reddy *et al.* (2008). The problem of non-specificity of amplification was overcome by performing a second round of PCR amplification using corresponding nested primers of the locus specific primers and walker primer. In this, each of the RCA products was amplified by locus specific primer based on F6R5 and walker primer (WP1).

The reaction was prepared in 50 µL volume containing 1X Buffer (5 µL), 200 µM dNTPs (1 µL), 150 ng each locus specific and walker primers (1.5 µL each), 50 ng RCA template (0.5 µL), 1 U Taq DNA polymerase (0.6 µL) and 39.9 µL double distilled water.

Method:

Each RCA product was separately amplified with two forward and two reverse locus specific primers. Thus each walker adapter annealed RCA product had two forward and two reverse primer PCR. This resulted in 16 different amplification results:

Amplification with WA1 RCA product			
WP1 LSP F1	WP1 LSP F2	WP1 LSP R1	WP1 LSP R2
Amplification with WA2 RCA product			
WP1 LSP F1	WP1 LSP F2	WP1 LSP R1	WP1 LSP R2
Amplification with WA3 RCA product			
WP1 LSP F1	WP1 LSP F2	WP1 LSP R1	WP1 LSP R2
Amplification with WA4 RCA product			
WP1 LSP F1	WP1 LSP F2	WP1 LSP R1	WP1 LSP R2

3.2.4.6 ***Nested locus specific PCR***

The locus specific PCR product was subjected to 1:50 dilution and this dilution was used as a template for the nested PCR reaction using nested locus specific primers (NLSP) and nested walker primer (WP2). The reaction was set to a final volume of 50 µL with template DNA (5 µL), 1X Buffer (5 µL), 200µM dNTP mix (1 µL), 150 ng each forward and reverse nested primers and nested walker primer (1.5 µL each), 1 U taq polymerase (0.6 µL) and 35.4 µL double distilled water.

The 16 nested PCR products corresponding to the locus specific PCR are as listed below:

WA1	WA2	WA3	WA4
WP2 NLSPF1	WP2 NLSPF1	WP2 NLSPF1	WP2 NLSPF1
WP2 NLSPF2	WP2 NLSPF2	WP2 NLSPF2	WP2 NLSPF2
WP2 NLSPR1	WP2 NLSPR1	WP2 NLSPR1	WP2 NLSPR1
WP2 NLSPR2	WP2 NLSPR2	WP2 NLSPR2	WP2 NLSPR2

For both locus specific and nested PCR reactions, the thermal profile used was:

Initial denaturation	94 °C, 2 min.	
Denaturation	94 °C, 1 min.	35 cycles
Annealing	51 °C, 1 min.	
Extension	72 °C, 2 min.	

Final extension	72 °C, 10 min.
Hold	4 °C

5 µL of the nested PCR products were loaded on agarose gel for analysis.

3.2.5 Sequencing and data analysis

The amplified PCR products obtained using various primer combinations were analyzed in 1 percent agarose gel along with a 100 bp or 1 kb DNA ladder and visualized under UV light. Accordingly, the bands with expected size were selected by comparing it with the DNA ladder.

The gel slice containing the DNA fragment of interest was excised out using a clean razor blade. Gel elution kit (Thermofischer scientific) was used for DNA elution from agarose gel.

Procedure:

The gel was cut as close to the DNA as possible in order to minimize the gel volume. The gel slice was placed into a pre-weighed centrifuge tube and weighed. Equal volume (volume: weight) of binding buffer was added and incubated at 50-60°C in a water bath until the gel slice dissolved completely. The tube was mixed by inverting every few minutes. Once the gel was completely melted, the tube was briefly vortexed and then transferred to the GeneJET purification column. The column was centrifuged for 1 minute and the flow-through discarded, the process was repeated with the remaining gel solution and the column was placed back into the same collection tube. Wash buffer diluted with absolute ethanol (700 µL) was added to the GeneJET purification column and centrifuged for 1 minute. Flow through was discarded and column placed back into the collection tube. The empty GeneJET purification column was centrifuged again for 1 minute to remove the residual wash buffer. The column was transferred to a new, clean micro centrifuge. DNA was eluted using 30 µL of elution buffer through the centre of the purification column membrane and centrifuged for 1 minute. The purification column was discarded and the purified DNA in the collection tube was stored at -20 °C.

The eluted products were checked for their concentration and the products of specified concentration and volume were sent for sequencing at the Eurofins India Pvt ltd. The list of PCR products, their concentration and elution volume sent for sequencing are given in Table 2.

Table 2: List of PCR products eluted from the gel and their concentration

Product	Size (bp)	Concentration (ng/μL)	Elution volume (μL)
F1R1	1000	30	50
F3R3	300	50	30
F4R4a	320	50	30
F6R5	400	60	20
F7R6	300	25	20
F7R7	400	25	20
A3F1	1000	50	30
A2F1	500	50	25
A4R1	750	50	25
A4R2	600	50	25

Expected as well as the observed amplicon size of each PCR product obtained were recorded.

3.2.6 Sequencing and validation

The eluted PCR products at various steps were sequenced and the sequence data were analyzed. Sequence information obtained by the sequencing of selected locus specific amplification products obtained by the forward primer (NLSP F1) and reverse primers (NLSPR1 and NLSP R2) were separately analyzed to find out if the sequences showed similarity to *ipk1* or not. Sequence thus obtained by genome walking which showed similarity with the *ipk1* gene (product A3F1 amplified by NLSP F1) was selected. Finally, the three sequences

viz., sequence obtained using primer combination PNC F6/PNC R5, sequence obtained using nested locus specific primers NLSP F1/WP2 and the original unigene sequence (*Pnc135*) were compared by multiple sequence alignment (Clustal Omega) to find out the overlapping regions and were manually assembled by joining the sequences with overlapping ends. The overlapping regions from either one of the sequences were removed before joining the ends of the sequences. The assembled nucleotide sequence *(Pnipk1)* was analyzed for finding the coding region of the gene using the software ORF Finder (https://www.ncbi.nlm.nih.gov/orffinder/). The amino acid sequence obtained was searched against NCBI conserved domains database (CDD) (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) to find the functional domains present, if any.

After completing the primer walking, the length of overlapping region of the amplicon with the original unigene from the cDNA was recorded.

3.2.7 Phylogenetic analysis

The assembled *Pnipk1* gene sequence and the translated aminoacid sequence of ORF were subjected to BLAST homology search and the aligned *ipk1* sequences showing similarity to query sequence were downloaded. The BLAST aligned sequences were used for phylogenetic analysis based on Neighbour joining algorithm with a bootstrap value of 2000 using MEGA7 software (http://www.megasoftware.net).

3.2.8 Detection of phytic acid using polyacrylamide gel electrophoresis

The presence of phytic acid was confirmed using PAGE analysis as per the protocol stated by Alimohammadi *et al.*, (2013) for the detection of $InsP_6$. The leaf sample from Panniyur 1 was taken, cleaned and tissue was homogenised for extracting the supernatant. The supernatant was then run in PAGE along with phytate standard for detecting the presence.

Sample preparation:

Leaf sample (200 mg) was weighed and the tissue was homogenized using 500 µL of 0.4 M HCl in liquid nitrogen. The homogenate was collected in a microfuge tube and the sample was boiled for 5 minutes. The denatured sample was briefly vortexed and centrifuged at 12,000 rpm for 10 minutes. The supernatant was then transferred to a new microfuge tube. The sample was snap frozen in liquid nitrogen and thawed at room temperature. This was centrifuged again at 12,000 rpm for 5 minutes and the supernatant was transferred to a fresh microfuge tube and stored at -20 ^{o}C.

PAGE analysis of phytic acid:

The resolving gel (33.3 %) was prepared in TBE for a total volume of 5 ml and the components involved 3.7 ml of 40 % acrylamide/bisacrylamide (19:1), 0.5 ml of 10X TBE (pH 8), 0.3 ml ddH_2O, 32 µL 10% APS and 3.5 µL TEMED. The stacking gel topped on the resolving gel had the composition of 1 ml 16% acrylamide/bisacrylamide (19:1), 120 µL 5X TBE (pH 6), 69.5 µL ddH_2O, 8.5 µL 10% APS and 1 µL TEMED.

The glass plates were fitted with the casting unit by placing the spacers at two edges and resolving gel was poured into the space between the glass plates leaving about 2 cm at the top. The gel was allowed to solidify and the excess water was drained off. The stacking gel was then poured on top of the resolving gel; comb was placed and allowed to solidify. The gel was mounted on the electrophoresis unit and pre- run was performed for 20 mins at 100 V/10 mA. The phytate extract (10 µL and 20 µL) and the standards (1-5 nmoles of phytic acid dipotassium salt) were loaded on to the wells after mixing with 5 µL of Orange G loading dye. Final volumes of the standards were made to 25 µL. The gel was run at 100 V/10 mA until the dye front reached the bottom. The gel was then stained with toluidine blue (Sigma Aldrich) staining solution (20% methanol, 2% glycerol, 0.05% toluidine blue) for 30 minutes and then destained using the destaining solution with the same components except dye for overnight in a

shaker after several changes of the solution. The destained gel was visualized and the image was captured using gel imaging system (BioRad geldoc XR).

Quantification of phytate:

Gel images were quantified using the software GelQuant.NET (biochemlabsolutions.com). Different concentrations of standard phytate dipotassium salt were prepared and loaded on the gel for preparing the standard curve. Quantity of phytate in the sample was calculated based on the standard curve obtained by ploting a graph corresponding to band intensities of different concentrations of phytate standards loaded on the gel. Standard curve was drawn by taking quantity of phytate standards (nmoles) on x- axis and volume (intensity) on y –axis.

Results

4. RESULTS

The study on 'Genome walking for putative phytic acid ($InsP_6$) unigene in black pepper (*Piper nigrum* L.)' was conducted using variety Panniyur 1 with an objective to sequence the flanking regions of unigene *Pnc135* coding for phytic acid ($InsP_6$) by primer walking and to detect and quantify the phytic acid content in black pepper.

The work was carried out as two parts. First part involved the unique technique of genome walking in order to decipher the unknown sequences on the upstream and downstream positions of a known sequence *viz*., unigene *Pnc135* obtained in a prevoius experiment (unpublished data). For this a method suggested by Reddy *et al.* (2008) based on rolling circle amplification of the whole genome by Phi29 DNA polymerase followed by locus specific amplification was used. The works include genomic DNA isolation, whole genome amplification using rolling circle mode of amplification by Phi29 DNA polymerase, PCR amplification using specific primers designed and sequence analysis of the PCR products.

Second part involved the biochemical analysis using Polyacrylamide Gel Electrophoresis for the detection of phytic acid in black pepper leaf sample by band detection and comparison using the commercially available standard phytate marker.

4.1. GENOME WALKING

4.1.1 DNA isolation

Good quality genomic DNA from variety Panniyur 1 was successfully isolated using modified CTAB method (Plate 1). The DNA pellet was dissolved in distilled water and stored at -20°C.

4.1.2 Quality and quantity of genomic DNA

Quantification and quality check of the sample was done using 0.7 per cent agarose gel electrophoresis and bio photometer. The concentration of DNA obtained was 5.05 µg/µL with $A_{260/280}$ ratio 1.9, which indicated good quality DNA. This DNA was diluted to 100 ng/µL and 1 µL was used as template for the PCR reaction.

4.1.3 Confirmation of presence of unigene and amplification of flanking regions

The unigene sequence *Pnc135* was used for the sequence similarity search using NCBI BLASTn (https://www.ncbi.nlm.nih.gov/blast/). The BLAST homology search results showed 68 per cent or more similarity of *Pnc135* with the available gene sequences of *ipk1* from other crops. Hence primers were designed simultaneously for amplification of genomic DNA of Panniyur 1 for confirmation of the presence of unigene in the genomic DNA and for the amplification of the flanking regions.

4.1.3.1 *Primer designing for amplification of unigene from genomic DNA*

The homology of inositol pentakisphosphate 2-kinase (*ipk1*) gene sequence was confirmed with the unigene (995 bp) through BLAST homology search. Since it is an assembled sequence from contigs, in order to confirm the amplification of the same sequence from the genomic DNA isolated from Panniyur 1, primers were designed manually based on the sequence as well as using the software OligoCalc. The primers PNC F1 and PNC R1 were designed to the 5' and 3' ends of the unigene respectively and PNC F2 was designed nested to PNC F1 and PNC R2 was designed as nested to PNC R1 (Fig.1). The reverse primer PNC R2 had a partial overlap with PNC R1. The size and sequence of the primers and the melting temperatures are as provided in the Table 3.

Table 3: Details of primers designed for amplification of unigene *Pnc135* sequence from black pepper

Primer	Sequence (5'-3')	Length (bp)	Tm (°C)	Expected size of the amplicon
PNC F1	AAACCCAATTTCCCGGGAATTT	22	55	959 bp
PNC R1	TGAATCAATGTCCAAGGAAGTG	22	53	
PNC F2	CAAAGGGAAACCCGGGTTAA	20	55	903 bp
PNC R2	CAAGGAAGTGCTAGGGCGATTAT	23	57	

>pnc135

5'TTGGGTTAAACCCAATTTCCCGGGAATTTCCGGGAATTTGGGCCCTTTCCCCAAA
GGGAAACCCGGGTTAACCCGGGTTAAAGGGCCCTTTGGTTTAACCCAATTTCCCGG
GAATTCCGGGAATGGCTTCAAGGACCCGGTTTCGGGTTTACGGACCGGCCGAGGT
ACTTCTTAGAAAGGCCACCTAGCTCCTTGCAGACCATGCAAGGTTGAGAAATAAT
GTTGTAATAGGCATGAATTGCCCCTTCTATATCAAGAAGGTCAAGCATTTGCGTAT
TAAGAAGACCAACCAGGACATCTGACTGTAAAATTGTCACTGAAACAAGTTCCAG
AAATCTAGATAGTCGTAGTCCAGGTGCTGCATCAATTAACACACCTAATTCATCAT
CAATAGCTTCCATGGATTTAGATTTCGTATTATCCATGGCTCCCCCCAATCCACCA
AATAACAAAGAGCCATTGAAAAAGATACGAAAATTGTTTTGCGGGATGGCATAGA
GAGATTTCATAGCACGATTTATTCTATCTTTAGAACCTGAAAACAGATCGACTGGA
TCATACTCACTTATTTCCCCTATCTCTTTCTTGTGAAATTTCAAGGATTGGTGCATT
TTAAAACGAGACACACACTTTTTGATGAAGTTATCTTCAGCTATGAATTTTGAGAC
TGGAAGAAAACCACACTTTGGCTTTATTTCCACTGCTATGCACCGATCCTCTCTAA
CACTTCCATGGATAAATAGTGAATGATCTGAAATAAGAAGAGCTGAATCACAGAG
AACATTAACCTTTGCCTCATCAACACGCCAAGATGGGCGCAAGCTCTGAACTTGTT
CTTCCAATGATTCCAGAAAGTCTCTTGACACATGCACATGTTTCCCAGCATCGACA
TGCTTTGAGCCTAGCAGTGGACTCATAACATTTAGCACATATAATCGCCCTAGCAC
TTCCTTGGACATTGATTCAGTCTGTCCTGCCGGCGCCGCGTGCCGTAC3'

Fig. 1. Unigene *Pnc135* sequence with primers for sequence confirmation in genomic DNA of black pepper

[PNC F1 (5' end): Yellow,

PNC R1 (3' end): Blue,

PNC F2 (5' end): Underlined in double lines,

PNC R2 (3'end): Underlined with single line]

4.1.3.2 *Primer designing for amplification of the flanking region of unigene Pnc135*

Primers were also designed for the amplification of flanking region of unigene *Pnc135*. For this, the 995 bp long unigene sequence was compared with the available gene sequences of *ipk1* of other crops using the BLAST homology search. BLAST search result revealed that the query sequence was aligned to the antisense strands of the database template sequences. Therefore reverse complement of the unigene sequence was taken for comparison with other sequences (Fig. 2). Based on the similarity and length, sequences from four crops *viz.*, *Nelumbo nucifera, Vitis vinifera, Sesamum indicum* and *Cicer arietinum* which showed 68 per cent or more identity with the unigene sequence in the BLAST homology search were selected (Table 4).

Table 4. Extent of homology of unigene sequence *Pnc135* from black pepper with *ipk1* gene sequence from four selected crops

Description	Accession	Length (bp)	Max score	Query cover (%)	E value	Identity (%)
Nelumbo nucifera ipk1-like, mRNA	XM_010256460	2914	288	78	3e-73	69
Cicer arietinum ipk1 -like, mRNA	XM_004500934	2245	250	58	7e-62	72
Sesamum indicum ipk1, mRNA	XM_020695367	2175	221	66	3e-53	70
Vitis vinifera ipk1, mRNA	XM_010666413	2054	259	73	1e-64	68

Alignment results showed that *Nelumbo nucifera* and *Vitis vinifera* had a longer sequence length and similar or consensus regions at the upstream 5' end of query sequence. The pairwise alignment of both the sequences had a length of 600 bp flanking the 5' end of unigene. So, these were used for designing the forward primer for amplifying 5' flanking region (PNC F3). Reverse primer for 5' flanking region (PNC R3) was designed from the unigene sequence (Fig. 2). The

forward primer for the 3' end amplification (PNC F4) was designed from the 3' end of the unigene (Fig. 2). The sequences of *Sesamum indicum* and *Cicer arietinum* were found to contain consensus regions at the 3' flanking ends of the unigene sequence and were longer than the other sequences towards the 3' end. The pair wise alignment of the two sequences had an additional length of 715 bp towards the 3' flanking region and therefore were used to design the two reverse primers (PNC R4a and PNC R4b) for the 3' flanking region.

>*Pnc135*_reverse_complement

5'GTACGGCACGCGGCGCCGGCAGGACAGACTGAATCAATGTCCAAGGAAGTGCT
AGGGCGATTATATGTGCTAAATGTTATGAGTCCACTGCTAGGCTCAAAGCATGTCG
ATGCTGGGAAACATGTGCATGTGTCAAGAGACTTTCTGGAATCATTGGAAGAACA
AGTTCAGAGCTTGCGCCCATCTTGGCGTGTTGATGAGGCAAAGGTTAATGTTCTCT
GTGATTCAGCTCTTCTTATTTCAGATCATTCACTATTTATCCATGGAAGTGTTAGAG
AGGATCGGTGCATAGCAGTGGAAATAAAGCCAAAGTGTGGTTTTCTTCCAGTCTCA
AAATTCATAGCTGAAGATAACTTCATCAAAAAGTGTGTGTCTCGTTTTAAAATGCA
CCAATCCTTGAAATTTCACAAGAAAGAGATAGGGGAAATAAGTGAGTATGATCCA
GTCGATCTGTTTTCAGGTTCTAAAGATAGAATAAATCGTGCTATGAAATCTCTCTA
TGCCATCCCGCAAAACAATTTTCGTATCTTTTTCAATGGCTCTTTGTTATTTGGTGG
ATTGGGGGGAGCCATGGATAATACGAAATCTAAATCCATGGAAGCTATTGATGAT
GAATTAGGTGTGTTAATTGATGCAGCACCTGGACTACGACTATCTAGATTTCTGGA
ACTTGTTTCAGTGACAATTTTACAGTCAGATGTCCTGGTTGGTCTTCTTAATACGCA
AATGCTTGACCTTCTTGATATAGAAGGGGCAATTCATGCCTATTACAACATTATTT
CTCAACCTTGCATGGTCTGCAAGGAGCTAGGTGGCCTTTCTAAGAAGTACCTCGGC
CGGTCCGTAAACCCGAAACCGGGTCCTTGAAGCCATTCCCGGAATTCCCGGGAAA
TTGGGTTAAACCAAAGGGCCCTTTAACCCGGGTTAACCCGGGTTTCCCTTTGGGGA
AAGGGCCCAAATTCCCGGAAATTCCCGGGAAATTGGGTTTAACCCAA 3'

Fig. 2. Reverse complement sequence of *Pnc135* with the primers designed for amplification of flanking sequences

[PNC R3 (5'end): underlined with single line,
PNC F4 (3'end): underlined with double lines]

In order to avoid any failure of amplification, partial degeneracy was applied to the forward and reverse primers (PNCF3 and PNCR4a,b) designed

using the multiple alignment result. Degeneracy was applied to the bases that showed differences between the multiple aligned sequences. The sequence and other properties of the primers are listed in Table 5.

Table 5. Details of degenerate primers designed for amplification of the flanking region of unigene *Pnc135* from black pepper

Primer	Sequence (5'-3')	Length (bp)	GC (%)	Tm (°C)	Expected size of the amplicon	Sequence details
PNC F3	TACY*TTCK*CAS*KCG GGYAKSA	21	43-67	57-67	~700 bp	Based on the sequences of *Nelumbo* and *Vitis*.
PNC R3	GAGCCTAGCAGTGGA CTCATAA	22	50	62		specific from the 5' end of unigene sequence
PNC F4	TTTCTCAACCTTGCAT GGTCTG	22	45.5	60	~320 bp	specific to the 3' end of the unigene
PNC R4a	AGCAGCTKACTATCT KCTKR*TC	22	36-45	56-64		based on the consensus regions of *Sesamum* and *Cicer*
PNC R4b	TCSATGM*AATCATAA GARTAA	21	24-33	40-48	~900 bp	

*universal degenerate code of bases Y-T/C, S-G/C, K-G/T, R-A/G, M-A/C

4.1.3.3 *PCR amplification of unigene Pnc135*

PCR reactions were carried out using the two sets of primers designed for the sequence confirmation. All 4 possible combinations of primers (PNC F1/PNC R1, PNC F2/PNC R2, PNC F1/PNC R2 and PNC F2/PNC R1) were used at a gradient of temperature ranging from 46°C - 60°C. Among the primer combinations, only PNC F1/PNC R1 produced a band of expected size. All other PCR reactions either resulted in non-specific amplification or did not produce any bands. The gradient PCR for F1R1 (amplified using PNC F1/PNC R1) produced a band of ~1kb at annealing temperatures of 46°C and 48°C (Plate 2) which was in the range of expected amplicon size.

Plate 1: Genomic DNA isolated from Panniyur 1 variety of black pepper.

(Lanes 1, 2, 3,4 : Replicates of DNA samples from Panniyur 1)

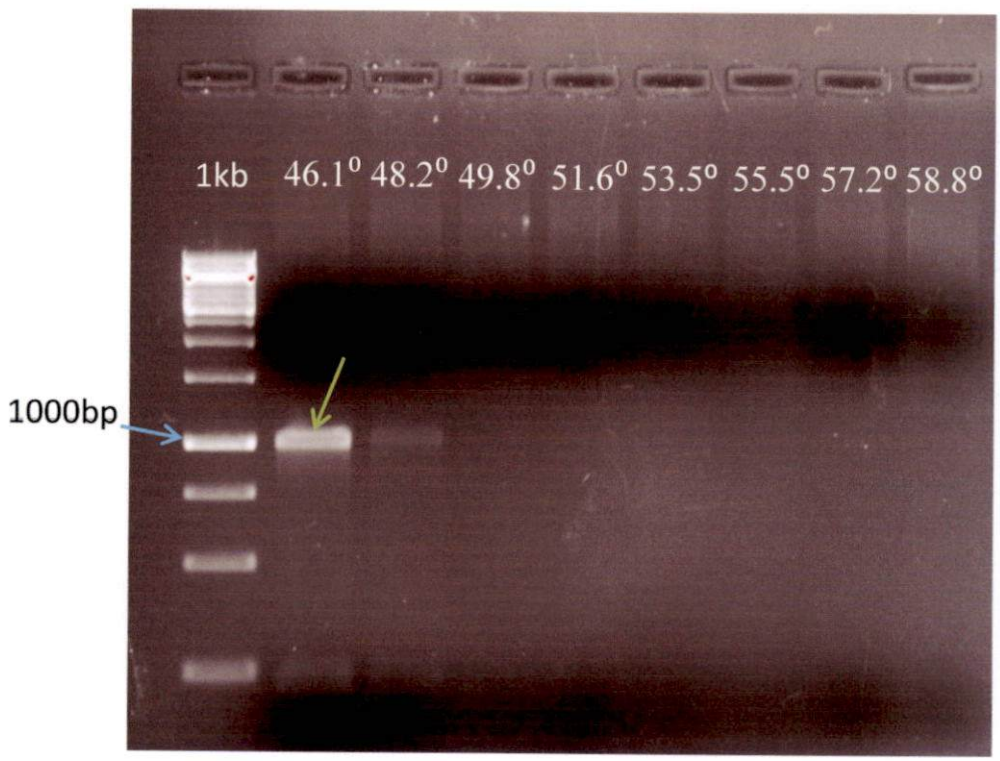


Plate 2. Gradient PCR for amplification of genomic DNA of Panniyur 1 using *Pnc135* specific primer set (PNC F1/PNC R1)

(Arrow indicates the PCR product F1R1)

19 41 235

4.1.3.4 *Sequence analysis for the confirmation of unigene sequence*

Only one PCR product (F1R1) was sequenced for the confirmation of unigene sequence information as it was the only amplicon with expected size. When sequenced with primer PNC F1, it gave a sequence of length 933bp (Fig 3). However the pair wise alignment of the F1R1 sequence with the unigene *Pnc135* did not show any similarity. BLAST homology search of the sequence also did not show similarity with the *ipk1* gene of any crops but was found to have similarity with light harvesting chlorophyll a/b binding protein of *Nicotiana sylvestris* (ABO12636), DNA sequence from chromosomes of *Solanum lycopersicum* (HG975514), complete sequences of clones of *Gossypium hirsiutum* (AC243134), whole genome shotgun sequences of *Vitis viifera* (AM 433882) etc.

>PNC135F1R1_F1_1457-1_P0200, Trimmed Sequence (933 bp)

CCGTTAGTCATCTAGATAATAGTTCTCAAAGTGAACTAGTTATCTTTGGTGAAAAT
TACCCAGTTTTCACAACAATTGGTATCAGAGCCCAAGTTGGTTTTGGTGTTGCCTT
GATTTGCCTCGCGAATTGTCGGTCGGTCCGAATTTGTGTCTCCTGATATGTCGGCC
GGTCCGAATTGGTGTTGTCTCCCAGACCCATGTTGGCTCCGCATGTGGTGAGGGGG
AGATGATTCGGGTGTGATGTGTGAGGGAGAGAATTGTTGGGTTTATCCCACATCCC
ATTTGGAGGCAATTGGTGACTTGTATATAAGCTAATGGGAAAGCTTTAATTCTTGA
GCTATATTTTGGATTAAAAGAGGCCTGTGAACCTGTTGCAAGTTATTTTGGGGCTC
GAAAGGCCCAACAAGTGGTATCAGAGCATATACCTGCTCACAACCCCAACTGCTT
GTGAAATATCGGGCCTTGAGCATAGCATTGCATACATCAAGCACTTGCATACGGTA
CCATTTTCATGTATTCCTGTTCTTCATCATTTTTTGGTGACATGGAAGCACTAAGCT
TGAAATGAGGAGCAAGTGGTGTACTAACTGGTTTTGACTTCTCATTTATGCTAAAA
CGTTTTAGTACCTTCCTCAGATATTCTTTCTGAGACAAACAAAGCCTCCCCAATTTC
CTCTCTCTGGATATCTCCATGCCAAGAATCTTTATTGCTTCACCAAGATCCTTCATC
TCAAACTCCCTATTCAATTGAGTCTTTAACTTCTCAATTTCTGTTTGACTCTTGGAA
GCTATCAGCATATCATCAACATAAAGAAGAAGATATATATAGGATCCATCCTGTA
GCTTGCGCAAATACACACAGTGATCATATTTACTCCTCGAGTACTTCTGCCCTATC
ATGAACTTGTCAAATTGCTTGTACCACTTCCTTGA

Fig. 3. Sequence of PCR product F1R1

4.1.3.5 *PCR amplification for the flanking regions of unigene*

One set of primer was designed for the amplification of 5' end and 2 sets of primers for 3' end of the unigene sequence. PNC F3/ PNC R3 were designed for the amplification of 5' flanking region and PNC F4/PNC R4a and PNC F4/PNC R4b were designed to amplify the 3' flanking end. Among the different

72

annealing temperatures tried, amplification was obtained at temperature 53 ^{o}C alone, though the band intensity was low. The PCR product F3R3 from 5' flanking region was of ~300bp against an expected product of size 700 bp whereas the amplicon F4R4a was of 320 bp which was in the expected range (Plate 3). The primer combination PNC F4/ PNC R4b produced multiple bands and therefore was not selected.

4.1.3.6 *Sequence analysis of amplicons from the flanking regions*

Two PCR products were sequenced for getting the sequence of flanking region. PCR product F3R3 was sequenced with the primer F3 and produced a trimmed sequence length of 238bp (Fig 4). The sequence was searched for its similarity with the *ipk1* gene using the NCBI BLAST homology search tool. The search result did not show homology with any gene sequences from plant kingdom.

>PNC135F3R3_F3_2268-1_P0291,Trimmed Sequence(238 bp)

GTGCAAGTTGATTTGGTCACAATCATGTTATTTATACTGATGATTAGAAGTTGGAA
CTCCAAAAATAATTTTTGCCTATGCATTAAATAAAACCCTAGCTTTTGACTAGTATT
GTGTCTATTTTTATATATATTTTTCATCAATGTCTTGTTTTTCCATCAGTTTTGTACT
TATGCTCATTCTCTTCTTATGCTGCATTACTTGTGTGTATGCCTGCAGAACGTTGAT
TATGAGTCCA

Fig. 4. Sequence of PCR product F3R3

>PNC135F4R4a_R4a_2268-2_P0291,Trimmed Sequence(272 bp)

TATAACTCTTCCCTCTTCTTATTTCTTTTCCCTTAGGTCTCTTGCATGCATGGTTTTG
CTGGTAAGGATGTAGGGTTGATGACGTTCTTCACACTTTCTTCTTTCCACATCGGTT
TTTGGGAGAGGGTTGCATGTTGGCCTGATTGCCACCATCCCCTTCCTATTTCCTCTA
GTTTTTCCTTGCTCTTGGCAGTGTTTTGGATTATTGTCTGCGGTGCCTTCATGCGTTT
CAGTTGTCGGCTATCAGATCAGACCATGCAAGGTTGAGAAAA

Fig. 5. Sequence of PCR product F4R4a

The PCR product F4R4a was sequenced with primer PNC R4a and obtained a sequence of 272bp (Fig. 5). The homology search of the sequence did

not produce any similarity. Query sequence showed similarity with only a few plant sequences including signal peptide peptidase of *Nelumbo nucifera* (XM010257083), F- box protein of *Theobroma cacao* (XM007012631), and kinesin proteins of many sequences from animal kingdom.

4.1.4. Detection of conserved regions of *ipk1* gene sequences by multiple sequence alignment with unigene *Pnc135*

As the amplification products F1R1 (for *ipk1* gene sequence) as well as the products for flanking regions (F3R3 and F4R4a) were not showing homology to *ipk1* gene, multiple sequence alignment of *ipk1* gene sequences reported from different crops along with the unigene was carried out to find out conserved or consensus regions.

4.1.4.1 *Primer designing*

From the multiple sequence alignment, the conserved regions in unigene *Pnc135* and *ipk1* gene sequences reported from different crops were identified. These regions were used for designing of forward primers PNC F6 and PNC F7 as well as reverse primer PNC R5.

Similarly, pair wise alignment of the unigene with *Nelumbo nucifera* was performed for designing the forward primer PNC F5 and reverse primer PNC R6. Reverse primer PNC R7 was designed from the pair wise alignment with *Elaeis guineensis.* The alignment results are shown in figure 6 (a-f). The sequences and other details of the primers are listed in Table 6.

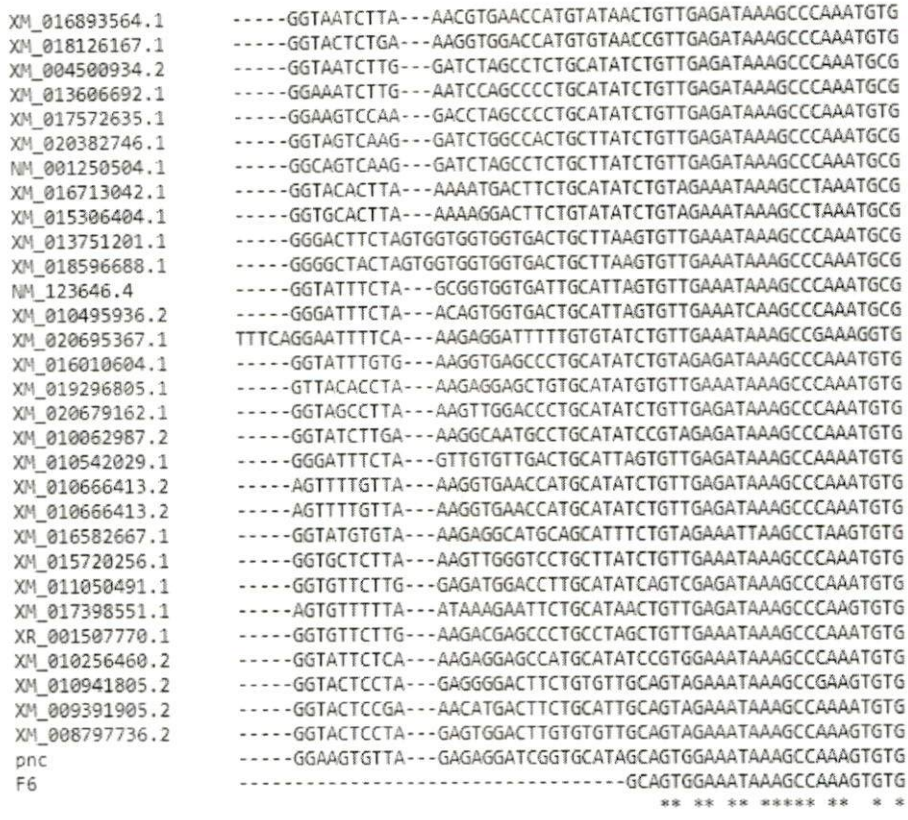


Fig. 6a: Multiple sequence alignmnet of *ipk1* sequences of different crops along with *Pnc135* for designing primer PNC F6

```
XM_015720256.1      CGGAGGTAAGCCAATATGATCCATTGGATATTTTCTCTGGTTCCAAGGAAAGGATCCAGA
XM_020679162.1      CAGAGTTGAGCCAGTATGATCCTCTGGACATTTTCTCTGGTTCGAAGGAAAGAATACACA
XM_016010604.1      TAGAATGTAGTAAATATGATCCTATTGATCTGTTCTCTCAATCCAAGGACAGAATACATA
XM_016893564.1      CAGAATATAGTGAATATAATCCGCTGGATTTATTCTCTGGATCCAGAGAAGGGATACAAA
XM_018126167.1      CAGAATATAGTGAATATAATCCACTGGATCTATTCTCTGGATCCAGAGACAGGATCTGTA
XM_017398551.1      CAGAACTAAGCAAATATGATCCATTGGATCTGTTCTCTAAATCCATAGAGAGAGTGAATA
XM_010256460.2      CCCAGATAAGTGAATATGATCCACTTGATTTATTCTCTGGATCTAGTGAGAGAATTCATA
XM_010666413.2      CACAAATTAGTGAATATGAGCCATTGGATCTATTCTCTGGATCCAAGGAGAAAATGCATA
XM_010666413.2      CACAAATTAGTGAATATGAGCCATTGGATCTATTCTCTGGATCCAAGGAGAAAATGCATA
XM_009391905.2      CACAAATAAGTGGGTATGATCCGCTTGATCTATTCTCTGGATTAAAAGATAGAATACATC
XM_013751201.1      CTGAAGTAAGTGAATATGATCCTCTTGATCTGTTCTCAGGATCCAAAGATAGAGTCTCCA
XM_018596688.1      CTGAAGTAAGCGAATATGATCCTCTTGATCTCTTCTCTGGATCCAGAGAAAGAGTATCAG
NM_123646.4         CTGAAGAAAGCGAGTACGATCCTCTTGATCTCTTCTCTGGATCAAAAGAGAGTGTTTTGG
XM_010495936.2      CTGAAGAAAGCGAATATGATCCTCTTGATCTCTTCTCTGGATCCAAAGAGAGAGTTTTGG
XM_016582667.1      CAGAGGTTAGTGCATATGATCCTCTGGATTTATTCTCTGGATCCAGAGACAGAGTACGCA
XM_010941805.2      CAAACATAAGTGAATATAACCCACTTGATCTGTTCTCTGGATCAAAAGACAGAATATATC
XM_008797736.2      CAAAGGTAAGTGAATATAACCCACTTGATCTGTTCTCTGGATCAAAAGACAGAATATATC
XM_020695367.1      GAAAAATAAGTAAATATGATCCTCTGGATATGTTTTCTGGATCCAAGGACAGAGTACAAA
XR_001507770.1      CAGAGAGAAGTGAATATGATCCACTGGACCTTTTTTCTGGGTCCAAGGAAAGAATATGTA
XM_017572635.1      CACGGCTAAGTGAGTACAATCCACTTGATCTATTCTCTGGATCCAAGGAAAGAATTCAGA
XM_020382746.1      CACAGCTGAGCGAGTACAATCCACTTGATCTGTTCTCTGGATCCAATGTAAGAATTCAGA
NM_001250504.1      CGCAACTAAGTGAGTACAATCCACTTGATCTGTTCTCTGGATCCAAGGAAAGAATTCTGA
XM_016713042.1      CAGAGATAAGTGCATATGATCCATTGGATTTGTTTTCTGGATCCAGAGATAGAGTAGACA
XM_015306404.1      CAGAGATAAGTGCATATGATCCATTGGATTTGTTTTCTGGATCCAATGATAGGGTACACA
XM_004500934.2      CGTTGCTAAGTGAGTATAATCCACTCGATTTGTTCTCTGGATCAAAGGAAAAAATTCATA
XM_013606692.1      CACTGCTAAGTGAATACAATCCACTTGATCTGTTCTCTGGTTCCAAGGAAAGAATTCATA
XM_010542029.1      CGGAACTGAGTGAATATGATCCTCTTGATCTGTTCTCAGGATCCAAGGAAAGGATACACA
XM_011050491.1      CAGAGTTAAGCGAGTATGATCCTCTGGATCTGTTTTCTGGTTCCAAGGAAAGGATACATA
XM_019296805.1      CATATATAAGTGCATATGATCCACTTGATATGTTTTCTGGGTCTAAAGACAGATTAGACA
XM_010062987.2      CAAATCTAAGTGAATATGATCCATTGGATCTGTTCTCTGGATCCAAGGAAAGAATACTTG
pnc                 GGGAAATAAGTGAGTATGATCCAGTCGATCTGTTTTCAGGTTCTAAAGATAGAATAAATC
F7                  ------------------ATCCAGTCGATCTGTTTTCAGG--------------------
                                       * **  * **  * ** **
```

Fig. 6b. Multiple sequence alignmnet of *ipk1* sequences of different crops along with Pnc135 for designing primer PNC F7

```
XM_010256460.2      ATCAGTTGAAAAGAATGTTCTTTGTCAGCGTCCTGCTTGGCGAGTTGATGCTGCAAATAT
pnc                 ATCATTGGAAGAACAAGTTCAGAGCTTGCGCCCATCTTGGCGTGTTGATGAGGCAAAGGT
F5                  ------------------------------------------------------------


XM_010256460.2      TAATATCCTCTGTGATTCTGCACTTCTTGTCTCTGATCATTCTGTCTTTCCACATGGTAT
pnc                 TAATGTTCTCTGTGATTCAGCTCTTCTTATTTCAGATCATTCACTATTTATCCATGGAAG
F5                  -AATGTTCTCTGTGATTCAGCTCTTC---------------------------------
                     *** * ********** ** ****


XM_010256460.2      TCTCAAAGAGGAGCCATGCATATCCGTGGAAATAAAGCCCAAATGTGGTTTTCTTCCATC
pnc                 TGTTAGAGAGGATCGGTGCATAGCAGTGGAAATAAAGCCAAAGTGTGGTTTTCTTCCAGT
F5                  ------------------------------------------------------------
```

Fig. 6c. Pair wise alignment of *Pnc135* with *Nelumbo nucifera* for designing primer PNC F5

```
XM_010542029.1      ATATTCCACTCCTCAGAACAATTTCCGTGTATTCTTGAATGGTGCCCTCTTATTTGGTGG
XM_013751201.1      ATACTCTATCCCACAGAACAACTTCCGCGTGTTCTTGAACGGTTCTCTCGTGTTAGGCGG
XM_018596688.1      ATACTCCATCCCACAGAACAACTTCCGAGTCTTCTTGAACGGTTCTCTGGTATTAGGCGG
NM_123646.4         ATATTCCACTCCTCAAAACAATTTCCGCGTATTCTTGAATGGTTCTCTCATATTAGGGGG
XM_010495936.2      ATATTCTACTCCCCAAAACAATTTCCGTGTATTCTTGAATGGTTCTCTCATATTAGGTGG
XM_010062987.2      GTGCAACAACCCACAGAACAATTTCCGTGTTTTCTTGAATAATTCTCTCATTTTGGGTGG
XM_020695367.1      CTTTCTTACACCTCAGAATAATTTCCGCGTTTTCTTAAATGGTTCACTTGCTTTTGGAGG
XM_020679162.1      CTACACCACCCCTCAAAACAATTTCCGTGTATTCTTAAATGGTTCTCTTATATTTGGGGG
XM_010666413.2      CTTTACCACCCCTCAGAACAATTTTCGAGTATTCTTAAATGGTTCTCTCATATATGGGGG
XM_010666413.2      CTTTACCACCCCTCAGAACAATTTTCGAGTATTCTTAAATGGTTCTCTCATATATGGGGG
XM_015720256.1      TTATGCTACCCCTCAGAACAATTTCCGTGTGTTCTTGAATGGTTCACTTATATTTGGAGG
XM_011050491.1      GTACAACACTCCTCAGAACAATTTCCGTGTATTCTTGAATGGTTCTCTTATATTTGGGGG
XM_016893564.1      GTATGCCACTCCTCAGAACAATTTCCGTGTATTTCTGAATGGTTCTATTGTATTTGGGGG
XM_010256460.2      CTTTACCACACCTCAGAACAATTTTCGCATATTCCTAAATAGCTCTCTCATTTTTGGTGG
XR_001507770.1      ATATACCAATCCTCAGAACAATTTACGTGTATTCTTGAATGGTTCTCTCATATTTGGGAG
XM_010941805.2      ATTTGCAACTCCTCAGAACAACTTCCGTATATTCTTGAATGGCTCTCTCATTTTTGGAGG
XM_008797736.2      ATTTGCAACGCCTCAGAACAACTTCCGTATATTCTTGAATGGCTCTCTCATTTTTGGGGG
XM_004500934.2      ATTTACTACACCTCAAAACAATTTCCGTGTCTTTATGAATGGGTCTCTCATATTTGGTGG
XM_013606692.1      ATTTACTACGCCTCAAAACAATTTCCGTGTATTTATGAATGGGTCTCTCATATTTGGTGG
XM_017398551.1      TTATCTTACACCTCAAAACAATTTTCGGGTATTTTTGAATGGTTCCCTCGTATATGGGGG
XM_018126167.1      ATATGCCACTCCTCAGAACAATTTCCGTGTATTTTTGAATGGTTCTCGTGTATTTGGGGG
XM_009391905.2      GTTTGCATCTCCTCAAAATAATTTCCGTATATTTTTGAATGGCTGTCTTATTTTTGGAGG
XM_016010604.1      CTTCACCACCCCTCAGAACAATTTTCGTGTATTTTTGAATGGTTCTCTAATATTTGGGGC
pnc                 CTATGCCATCCCGCAAAACAATTTTCGTATCTTTTTCAATGGCTCTTTGTTATTTGGTGG
R5                  ------------------CAATTTTCGTATCTTTTTCAATGGCTC---------------
XM_017572635.1      CCTTATAACTCCTCAAAACAATTTCCGTGTATTTTTAAATGGCTCTCTCATACTTGGAGG
XM_020382746.1      TCTTACAACTCCTCAAAACAATTTTCGTGTATTTTTGAATGGCTCTCTCATACTTGGAGG
NM_001250504.1      CCTTACAACTCCTCAAAACAATTTTCGTGTATTTTTGAATGGCTCTCTCATACTTGGAGG
XM_019296805.1      TTTTATGACCCCACAGAATAACTTTCGAGTTTTTGTGAATGGTTGTCTCATATTTGGGGG
XM_016582667.1      TTTTAAGACCCCACAGAATAATTTTCGGGTTTTCCTGGATGGTTCTCTCATAATTGGTGG
XM_016713042.1      TTTTAAGACCCCGCAGAACAATTTCCGAGTTTTCCTGAATGGCTCTCTCATATTAGGGGG
XM_015306404.1      TTTCAAGACCCCACAAAACAATTTCCGAGTTTTCCTGAATGGTTCTCTCATATTAGGGGG
                                  ** ** **  * **  *  *
```

Fig. 6d. Multiple sequence alignmnet of *ipk1* sequences of different crops along with *Pnc135* for designing primer PNC R5

```
XM_010256460.2      CAGATCAGGAGTACTGGATCGACTTCTGGAAGCCCAGAAGCTTGATTTATTGGACATAGA
pnc                 ACAGTCAGATGTCCTGGTTGGTCTTCTTAATACGCAAATGCTTGACCTTCTTGATATAGA
R6                  ------------------------------------------------------------


XM_010256460.2      AGGGGCAATTCATGCATATTACAACATTGTTTCCCAGCCTTGTGTGGCATGTAGAGATTT
pnc                 AGGGGCAATTCATGCCTATTACAACATTATTTCTCAACCTTGCATGGTCTGCAAGGAGCT
R6                  -GGGGCAATTCATGCCTATTACAAC------------------------------------
                     ************** *********


XM_010256460.2      GGCTGTTAGGGAACTCTCACACAGTCTACATTCTATTTCTTTGGAAGAAAGCCTAAAGAT
pnc                 AGGTG-------------------------------GCCTTTCTAAG-------------
R6                  ------------------------------------------------------------
```

Fig. 6e. Pair wise alignment of *Pnc135* with *Nelumbo nucifera* for designing primer PNC R6

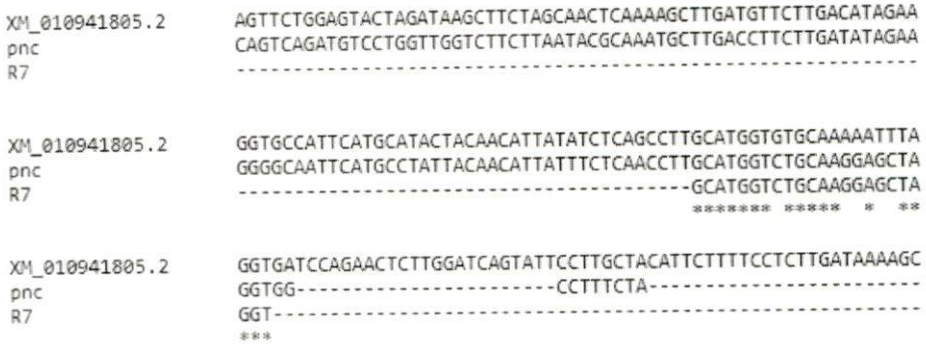


Fig. 6f. Pair wise alignment of *Pnc135* with *Elaeis guineensis* for designing primer PNC R7

Table 6. Details of primers designed for amplification of *ipk1* gene by multiple sequence alignment/ pair wise alignment

Primer	Sequence (5'-3')	Primer length (bp)	GC (%)	Tm (°C)	Sequence details
PNC F5	AATGTTCTCTGTGATT CAGCTCTTC	25	40	55	From pairwise aligned sequences of *Pnc135* and *Nelumbo nucifera*
PNC F6	GCAGTGGAAATAAAG CCAAAGTGTG	25	44	57	From multiple sequence aligned data of reported phytate sequence and *Pnc135* sequence
PNC F7	ATCCAGTCGATCTGTT TTCAGG	22	45	54	
PNC R5	GAGCCATTGAAAAAG ATACGAAAATTG	27	33	54	
PNC R6	GTTGTAATAGGCATGA ATTGCCCC	24	46	56	From pairwise aligned sequences of *Pnc135* and *Nelumbo nucifera*
PNC R7	ACCTAGCTCCTTGCAG ACCATGC	23	57	58	From pairwise aligned sequences of *Pnc135* and *Elaeis guineensis*

4.1.4.2 *PCR amplification*

Forward primers PNC F5, PNC F6, and PNC F7 and reverse primers PNC R5, PNC R6, PNC R7 were used to perform PCR amplification of the *ipk1* gene from Panniyur 1 genomic DNA. Gradient PCR was set up in the range of 48 °C to

63 °C to standardize the annealing temperatures of different primer combinations. All the primer combinations gave amplifications of different product sizes at the specified temperature gradients (Plate 3). The PCR products were selected for sequencing based on the product size obtained compared to the expected size and also based on the specificity of primers by avoiding multiple bands. The temperature gradient used, expected and observed product sizes are listed in Table.7.

Table 7. Expected and observed sizes of PCR amplification for *ipk1* gene

Product name	Primers combination	Annealing temperature (°C)	Expected size (bp)	Observed size (bp)
F5R5	PNC F5/PNC R5	51.8	330	~700
F5R6	PNC F5/PNC R6	55.6	560	~900
F5R7	PNC F5/PNC R7	51.8	600	~900
F6R5	PNC F6/ PNC R5	51.8	250	~400
F6R6	PNC F6/PNC R6	51.8	480	~650
F6R7	PNC F6/PNC R7	55.6	520	~700
F7R5	PNC F7/PNC R5	51.8 (not shown)	100	~100
F7R6	PNC F7/PNC R6	55.6	330	~300
F7R7	PNC F7/PNC R7	55.6	370	~400

The PCR products of PNC F7 primer combinations- F7R5, F7R6 and F7R7 gave bands of expected size, but F7R5 was not taken for further analysis as its product size is only 100bp (Plate 4). The products F7R6 and F7R7 were eluted from the gel and sequenced. Also the product F6R5 was selected for sequencing although it did not produce the expected size of product, as the primer combination used was based on the conserved regions in MSA and the primers showed specificity producing a single band of good intensity.

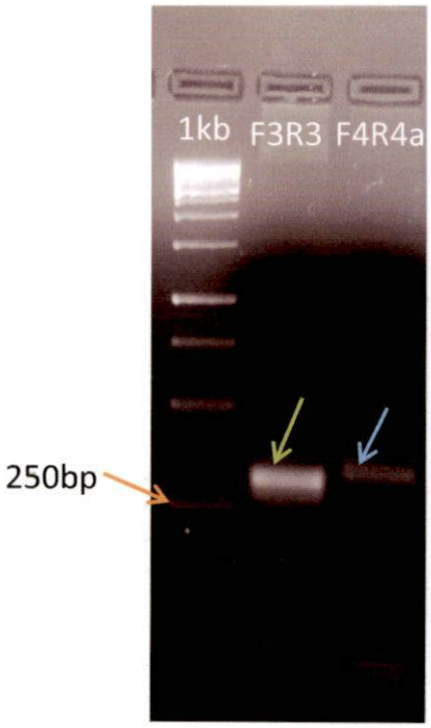


Plate 3. Amplicons from flanking regions of unigene *Pnc135* in Panniyur 1 DNA (F3R3 : amplicon at 5'end, F4R4a : amplicon from 3'end)

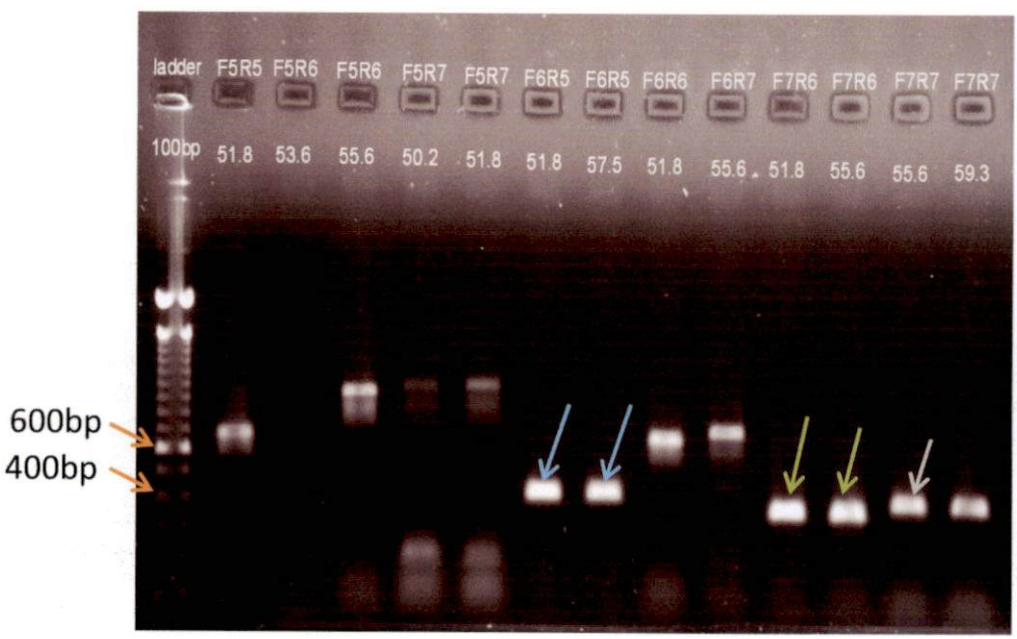


Plate 4. PCR products of different primer combinations used for amplification of *ipk1* gene at different annealing temperatures.

4.1.4.3 *Sequence analysis*

Three PCR products were selected for sequencing of *ipk1* gene. Primer PNC F6 was used for sequencing of F6R5 and primer PNC F7 was used for the sequencing of F7R6 and F7R7. Sequencing reaction of the three PCR products produced a sequence of 394 bp, 193bp and 297bp respectively for F6R5 (Fig. 7a), F7R6 (Fig. 7b) and F7R7 (Fig 7c). The analysis of the sequences of PCR products F7R6 and F7R7 using BLAST homology search revealed no similarity with the *ipk1* gene. However that of F6R5 showed similarity with the *ipk1* genes reported in crops like *Cicer arietinum* (82% identity, 49% query cover), *Beta vulgaris* (81% identity, 15% query cover), *Populus euphratica* (74% identity, 44% query cover), *Glycine max* (71% identity, 46% query cover) etc. The details of the BLAST search results are given in Table 8.

>0517_190_001_PCR_PNC_135_F6R5_PNC_135_F6_C03.ab1

GTCTTTTCGGGTCTCAAATTTCATAGCTGAGATAACTTCATCAAAAAGTGTGTGTC
TCGTTTTAAAATGCACCAACCCTTGAAATTTCACAAGAAAGAGGTAATTTTCACCT
GTGCTTACCTATTCCATTGTTTAAAAGAAAAAACTTAGATTACACTAATTGTTGAG
GATGCTTTTCTTCATCCTTTCTGTACACATGGATTTGGCTATCATTTTCTAATTATTT
TTCACTTCTCTCTCAGATAGGGGAAATAAGTGAGTATGATCCAGTCGATCTGTTTT
CAGGTTCTAAAGATAGAATAAATCGTGCTATGAAATCTCTCTATGCCATCCCGCAA
AACAATTTTCGTATCTTTTACCAATGGGCTCCAAAACTAAACGGCCGCTTTAAGCA
A

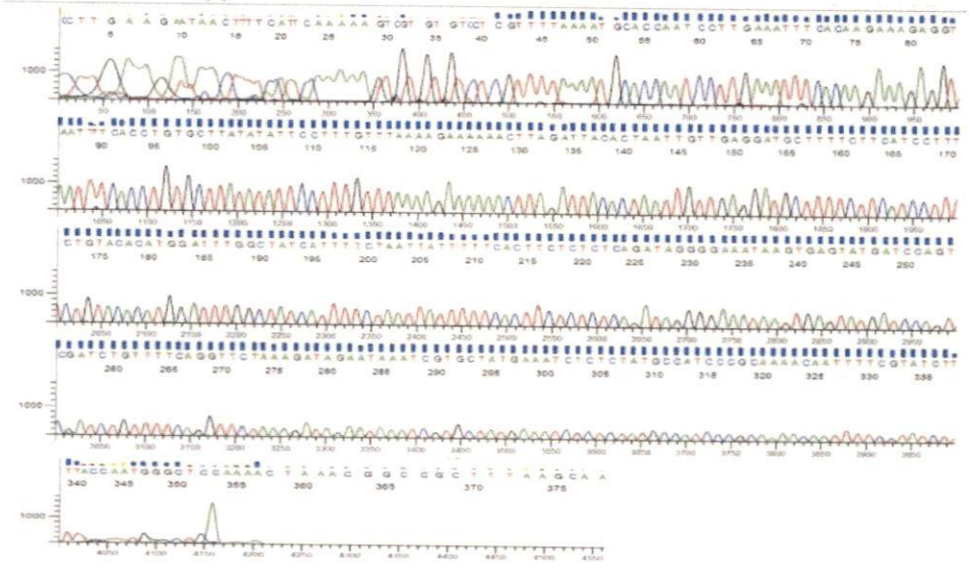


Fig. 7a. Sequence of F6R5

>0517_190_002_PCR_PNC_135_F7R6_PNC_135_F7_D03.ab1

GCCGGCCGGGGTACTGAACCGTTGGGTCCATAGTACCGTGTTTTTGTCACCCCCTGTT
GGAGGACCGACACCAGACGTTTAACAAAGGGGGTTCGGGTCTTGTTCGTATTTTGGTG
CCTTCTTTTGGGGAAGCGGCCCCGGGGGCGGGCTGTTCACCGATGATAAACATTTGTT
GGCGTGCTTCCCCCACGAG

Fig. 7b. Sequence of F7R6

>0517_190_003_PCR_PNC_135_F7R7_PNC_135_F7_E03.ab1

TTCGAATACAGTATATAAAGAATGGTTCGATGGAATAGTGTTGTCATCACCCCCCTTT
AAGAGGAAAAAATACGTAAATATGGGAACGAAAATGAGATGAGATGTGCGCGTGTT
ATGATGAATGTTAAGAAGAAAAGGCGAACAGGGGTCAAAGCATGTTCCTATAACAAT
TACAAACAAATATCAATTTATGAGCTTTACTCAATTGATGCTGCGGCCCCCGGTGTAG
CACGAGAAGGTAGAGCTATCCGATGGTGGGATGAGCGGACCTAATTGTGCATGGTGC
GCCCGACATGT

Fig. 7c. Sequence of F7R7

Fig.7. Sequences of PCR products obtained by amplification of genomic DNA of Panniyur 1 using primers designed for gene *ipk1*

Table 8. Homology of PCR product F6R5 with *ipk1* gene from crops

Description	Accession	Max score	Query cover	E value	Identity
Predicted: *Cicer arietinum* inositol pentakis phosphate kinase like, mRNA	XM 004500934	78.8	49%	3e-11	82%
Predicted: *Beta vulgaris subsp. vulgaris* inositol pentakisphosphate kinase, mRNA	XM 010674834	57.2	15%	9e-05	81%
Predicted: *Populus euphratica* inositol pentakis phosphate kinase like, mRNA	XM 011030776	68	44%	2e-07	74%
Medicago truncatula inositol pentakisphosphate kinase like protein, partial mRNA	XM 013606692	59	49%	0.004	72%
Predicted: *Arabidopsis thaliana* inositol pentakisphosphate kinase family protein, partial mRNA	NM 102060	66.2	43%	6e-04	75%
Predicted: *Camelina sativa* inositol pentakisphosphate 2 kinase, mRNA	XM 010479163	80	43%	0.014	76%
predicted: *Glycine max* inositol pentakisphosphate 2 kinase like,mRNA	XM 006581136	53.6	46%	0.004	71%
Predicted: *Lupinus angustifolius* inositol pentakisphosphate 2kinase, mRNA	XM 019581356	55.4	46%	0.001	71%

4.1.5 Rolling circle amplification (RCA) of whole genome for inserting walker adapters

Phi29 mediated isothermal amplification was used for random introduction of walker primer binding sites in the genomic DNA of Panniyur 1 by taking advantage of the RCA. The reaction had resulted in the synthesis of long overlapping fragments of the genomic DNA of Panniyur 1 with the walker adapter sequence attached to 5' end of all fragments.

4.1.5.1 *Primer designing and amplification*

To insert walker primer binding site (adapter site) randomly during rolling circle amplification of the whole genome by Phi29 DNA polymerase, walker adapter primers were required. The walker adapter primers WA1, WA2, WA3 and WA4 used in the present study were same as those reported in the protocol by Reddy *et al.* (2008), sequences of which are given in Table 9. The amplification of Panniyurl DNA was done using Phi29 DNA polymerase and the four different primers WA1, WA2, WA3 and WA4 in four different vials as described in chapter Materials and Methods. The products were used for genome walking using walker primer/locus specific primer (WP1/LSP) followed by nested walker primer/nested locus specific primer (WP2/NLSP) combinations.

4.1.6 Amplification of flanking regions of *ipk1* gene using RCA products and *ipk1* locus specific primers

4.1.6.1 *Primer designing for genome walking*

Among the various primer combinations tried for amplifying *ipk1* gene segment from genomic DNA of Panniyur 1, the best result was obtained for the product F6R5. Hence for genome walking the locus specific primers (LSP) and the nested locus specific primers (NLSP) were designed based on the sequence of F6R5. Two sets of locus specific and nested locus specific primers were designed. Primer LSP F1 was designed from the 5' end of F6R5 sequence leaving 5 bases upstream and NLSP F1 was designed 49 bp downstream of LSP F1 sequence. Primer LSP F2 was designed

36 bp downstream from the 5' end of F6R5 sequence and NLSP F2 was designed 48 bp downstream to LSP F2.

The reverse primers were designed by taking the reverse complements of the forward primers. LSP R1, NLSP R1, LSP R2 and NLSP R2 were designed from the 3' end and were the reverse complements of NLSP F2, LSP F2, NLSPF1 and LSP F1 respectively. LSP R2 differed from NLSP F1 for one base. Reverse primers were designed for the amplification of the flanking region of the 5' end and forward primers were designed for amplifying the flanking region of 3' end of the F6R5 sequence (Fig. 8). The sequence and details of the primers are listed in Table 9.

>0517_190_001_PCR_PNC_135_F6R5_PNC_135_F6_C03.ab1

GTCTTTTCGGGTCTCAAATTTCATAGCTGAGATAACTTCATCAAAAAGTGTGTGTCT

CGTTTTAAAATGCACCAACCCTTGAAATTTCACAAGAAAGAGGTAATTTTCACCTGT

GCTTACCTATTCCATTGTTTAAAAGAAAAAACTTAGATTACACTAATTGTTGAGGAT

GCTTTTCTTCATCCTTTCTGTACACATGGATTTGGCTATCATTTTCTAATTATTTTTCA

CTTCTCTCTCAGATAGGGGAAATAAGTGAGTATGATCCAGTCGATCTGTTTTCAGGT

TCTAAAGATAGAATAAATCGTGCTATGAAATCTCTCTATGCCATCCCGCAAAACAAT

TTTCGTATCTTTTACCAATGGGCTCCAAAACTAAACGGCCGCTTTAAGCAA

Fig. 8. Sequences and position of locus specific primers designed for genome walking from F6R5 sequence

Forward primers - LSP F1: yellow, LSP F2: blue, NLSP F1: green, NLSP F2: ash

Reverse primers - LSP R1, NLSP R1 and NLSP R2 are the reverse complements of NLSP F2, LSPF2 and LSP F1 respectively, LSP R2: reverse complement of the underlined sequence

4.1.6.2 PCR amplification using locus specific primer (LSP and NLSP)

Directional genome walking using the four RCA products was done using a walker primer and locus specific primer as per the protocol stated by Reddy *et al.* (2008). The problem of non-specificity of amplification was overcome by performing a second round of PCR amplification using corresponding nested primers of the locus specific primers and walker primer.

Each RCA product was separately amplified with the two forward and two reverse locus specific primers and walker primer (WP1). This resulted in 16 different amplifications as detailed below (corresponding product name given in bracket).

WA1 RCA product:	WP1/LSP F1 (A1F1)	WP1/LSP F2 (A1F2)	WP1/LSP R1 (A1R1)	WP1/LSP R2 (A1R2)
WA2 RCA product:	WP1/LSP F1 (A2F1)	WP1/LSP F2 (A2F2)	WP1/LSP R1 (A2R1)	WP1/LSP R2 (A2R2)
WA3 RCA product:	WP1/LSP F1 (A3F1)	WP1/LSP F2 (A3F2)	WP1/LSP R1 (A3R1)	WP1/LSP R2 (A3R2)
WA4 RCA product:	WP1/LSP F1 (A4F1)	WP1/LSP F2 (A4F2)	WP1/LSP R1 (A4R1)	WP1/LSP R2 (A4R2)

The 16 locus specific PCR products were subjected to 1:50 dilution and amplified with their corresponding nested primers at 51 °C. Five microlitres of PCR products were analyzed in 1percent agarose gel and those with greater sizes were selected for sequencing.

The nested PCR product A3F1 and A2F1 which were based on RCA method 1 mentioned in section 3.2.4.4 gave a single band of product size 1 kb and 500 bp, respectively (Plate 5). These two PCR products were selected for sequencing of the 3' flanking region of the sequence F6R5. The reverse primers produced bands of only smaller size and were not selected for sequencing. In the RCA method 2 based nested PCR, a greater amount of amplification was observed with the reverse primers compared to the forward primers. The PCR products showed multiple bands corresponding to different sizes (Plate 6).

From the amplicons, those with greater fragment size were selected. A4R1 produced multiple bands of product size about 750 bp, 600 bp and 300 bp. A4R2 produced bands of 600 bp and 100 bp. The products of size 750 bp and 600 bp were selected for sequencing of the 5' flanking region of F6R5.

Table 9. Details of primers designed for genome walking

Sl. No	Primer	For/ Rev	Sequence (5'---3')	Primer length (bp)	GC%	Tm °C
1.	WA 1		GTGAGCGCGCGTAATACGACTCACTATAGGGNNNNATGC	39	53.8	67
2.	WA 2		GTGAGCGCGCGTAATACGACTCACTATAGGGNNNNGATC	39	53.8	67
3.	WA 3		GTGAGCGCGCGTAATACGACTCACTATAGGGNNNNTAGC	39	53.8	67
4.	WA 4		GTGAGCGCGCGTAATACGACTCACTATAGGGNNNNCTAG	39	53.8	67
5.	WP 1		GTGAGCGCGCGTAATACGA	19	57.9	58
6.	WP 2		GTAATACGACTCACTATAGGG	21	42.9	49
7.	LSP R1	R	ACCTCTTTCTTGTGAAATTTCAAG	24	33.3	52
8.	LSP R2	R	AACAATGGAATAGGTAAGCACAGG	24	41.7	55
9.	LSP F1	F	TTCGGGTCTCAAATTTCATAGCTG	24	41.7	55
10.	LSP F2	F	TTCATCAAAAAGTGTGTGTCTCGT	24	37.5	55
11.	NLSP R1	R	CAGCTATGAAATTTGAGACCCGAA	24	41.7	55
12.	NLSP R2	R	ACGAGACACACACTTTTTGATGAA	24	37.5	55
13.	NLSP F1	F	TTGAAATTTCACAAGAAAGAGGTA	24	29.2	51
14.	NLSP F2	F	CCTGTGCTTACCTATTCCATTGTT	24	41.7	55

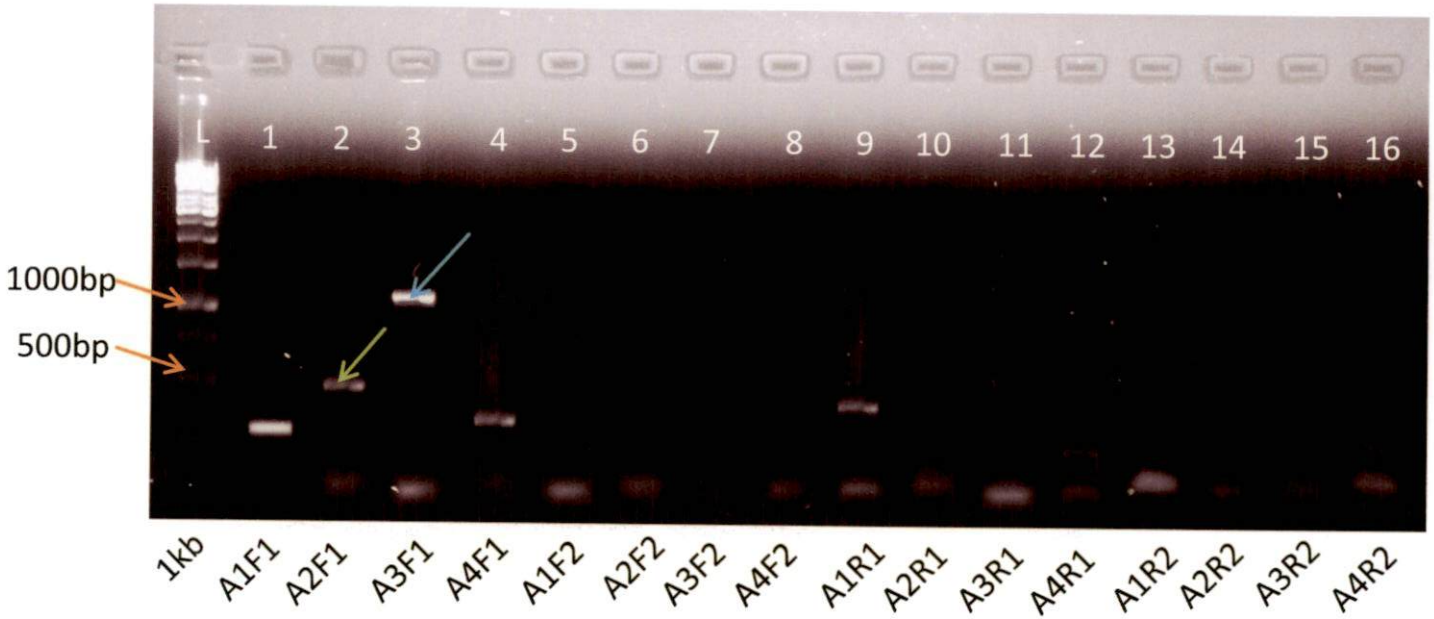


Plate 5. Nested PCR amplification of RCA product incubated for 90 min. at 30 °C

The products obtained in each lane corresponds to different primer combinations: lane1-WA1: NLSPF1, lane2-WA2: NLSPF1, lane3-WA3: NLSPF1, lane4-WA4: NLSPF1, lane5-WA1: NLSPF2, lane6-WA2: NLSPF2, lane7-WA3: NLSPF2, lane8-WA4: NLSPF2, lane9-WA1: NLSPR1, lane10-WA2: NLSPR1, lane11-WA3: NLSPR1, lane12-WA4: NLSPR1, lane13-WA1: NLSPR2, lane14-WA2: NLSPR2, lane15-WA3: NLSPR2, lane16-WA4: NLSPR2, laneL- 1 kb ladder. The products selected for sequencing are marked by arrowheads.

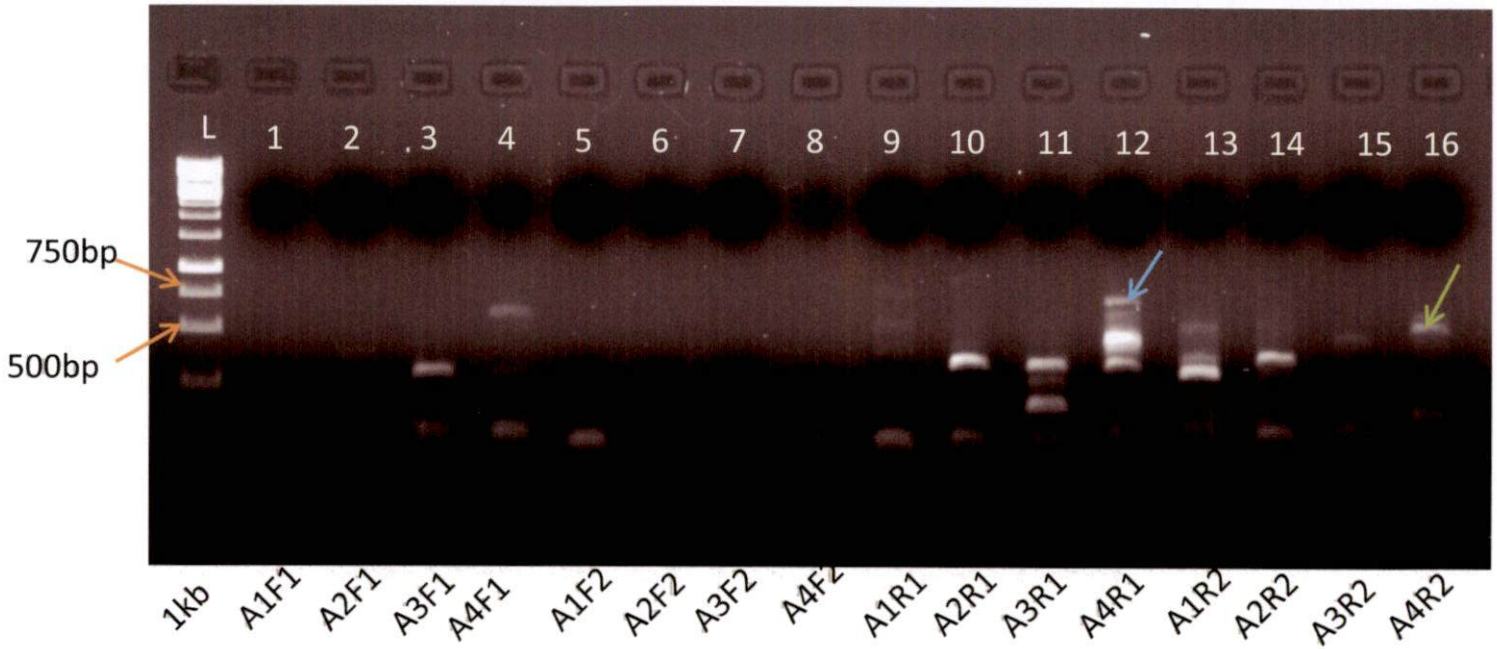


Plate 6: Nested PCR ampilfication of RCA product incubated overnight at 30°C

The selected products are indicated by the arrowheads and each lane corresponds to the products: lane1 –WA1: NLSP F1, lane2 –WA2: NLSP F1, lane3 –WA3: NLSP F1, lane4 –WA4: NLSP F1, lane5 –WA1:NLSP F2, lane6 –WA2: NLSP F2, lane7 –WA3: NLSP F2, lane8 –WA4: NLSP F2, lane9 –WA1: NLSP R1, lane10 –WA2: NLSP R1, lane11 –WA3: NLSP R1, lane12 –WA4: NLSP R1, lane13 –WA1: NLSP R2, lane14 –WA2: NLSP R2, lane15 –WA3: NLSP R2, lane16 –WA4: NLSP R2, laneL-1 kb ladder

4.1.6.3 *Sequence analysis of genome walking at 3' flanking region of F6R5*

There were two products from 3'end selected for sequencing – A3F1 and A2F1. A3F1 which is a 1 kb long product from 3' end of *ipk1* fragment F6R5 was sequenced with NLSPF1 and WP2. The 5' end of product A3F1 was sequenced with the primer NLSPF1. It produced a sequence of 475 bp (Fig 9a). BLAST homology search of the sequence result revealed sequence similarity with *ipk1* gene of several crops (Table 10) like *Nelumbo nucifera* (69% identity, 76% query cover), *Phoenix dactylifera* (66% identity, 76% query cover), *Musa acuminate* (75% identity, 44% query cover) etc. at the 5' end of sequence.

>0617_596_001_PCR_A3F1_NLSPF1_A09.ab1

AAAAAACTTAGATTACACTAATTGTTGAGGATGCTTTTCTTCATCCTTTCTGTACA
CATGGATTTGGCTATCATTTTCTAATTATTTTTCACTTCTCTCTCAGATAGGGGAA
ATAAGTGAGTATGATCCAGTCGATCTGTTTTCAGGTTCTAAAGATAGAATAAATC
GTGCTATGAAATCTCTCTATGCCATCCCGCAAAACAATTTTCGTATCTTTTTCAAT
GGCTCTTTGTTATTTGGTGGATTGGGGGGAGCCATGGATAATACGAAATCTAAAT
CCATGGAAGCTATTGATGATGAATTAGGTGTGTTAATTGATGCAGCACCTGGACT
ACGACTATCTAGATTTCTGGAACTTGTTTCATTGACAATTTTACAGTCAGATGTCC
TGGTTGGTCTTCTTAATACGCAAATGTTTGACCTTCTTGATATAGAAGGGGCAATT
CATGCCTATTACAACATTATTTCTCAACCT

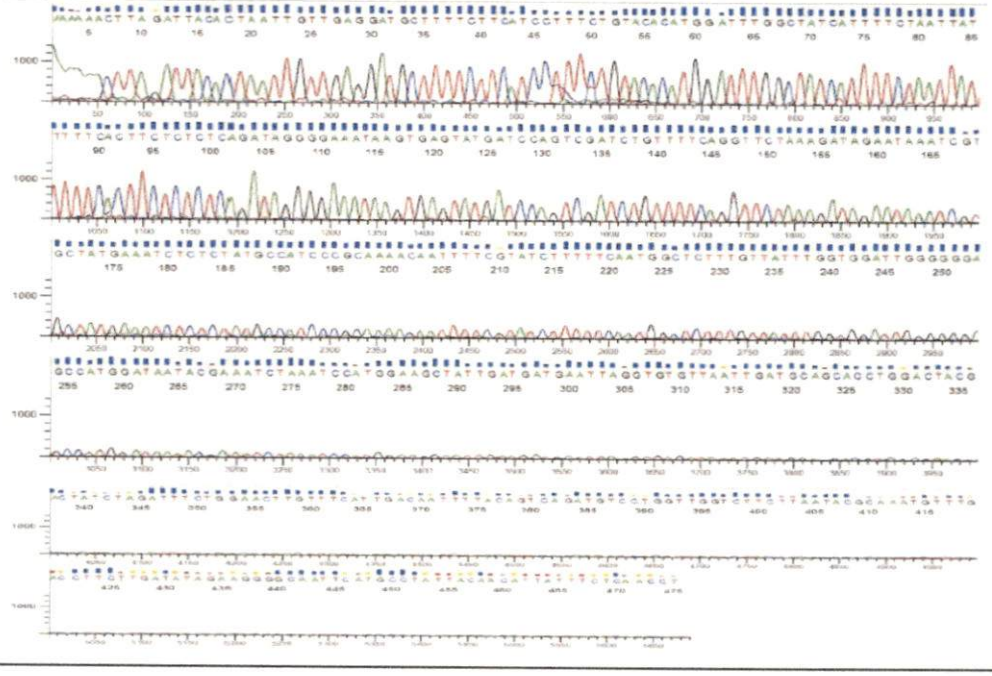


Fig. 9a: Sequence of 5'end of nested PCR product A3F1 sequenced with NLSPF

Table10. Homology of 5'end of A3F1 with 5'end of *ipk1* gene from other crops

Description	Accession	Max score	Query cover	E value	Identity
Predicted: *Nelumbo nucifera ipk1* like, mRNA	XM 002306847	111	76%	2e-20	69%
Predicted: *Phoenix dactylifera ipk1* like, mRNA	XM 008797737	107	76%	2e-19	66%
Predicted: *Camelina sativa ipk1*, mRNA	XM 010479163	102	31%	1e-17	75%
Predicted: *Musa acuminata subsp.malaccensis ipk1*, mRNA	XM 009391905	102	44%	1e-17	75%
Predicted: *Eucalyptus grandis ipk1*, mRNA	XM 010062987	102	76%	1e-17	67%
Predicted: *Vitis vinifera ipk1*, mRNA	XM 010666413	100	78%	4e-17	67%
Predicted: *Populous euphratica ipk1* like, mRNA	XM 011050491	100	76%	4e-17	68%
Predicted: *Tarenaya hassleriana ipk1*, mRNA	XM 010542030	93.3	29%	5e-15	74%
Predicted: *Cicer arietinum ipk1* like, mRNA	XM 004500934	93.3	29%	5e-15	75%
Glycine max ipk1, mRNA	NM 001250504	89.7	31%	7e-14	73%
Arabidopsis thaliana ipk1 family protein partial mRNA	NM 102060	87.8	31%	2e-13	74%
Predicted: *Lupines angutifolius ipk1* like, mRNA	XM 019609521	82.4	29%	1e-11	73%
Predicted: *Jatropha curcasipk1*, mRNA	XM 020679162	78.7	74%	1e-10	66%
Predicted: *Brassica rapa ipk1*, mRNA	XM 009111902	78.8	31%	1e-10	72%
Predicted: *Arachis duranensis ipk1* like, mRNA	XM021128459	77	28%	4e-10	72%
Medicago truncatula ipk1 like protein partial mRNA	XM 013606692	77	27%	4e-10	73%

The 3' end of A3F1 was sequenced with WP2 and a length of 234 bp was obtained (Fig 9b). BLAST homology search of the sequence (reverse complement) did not produce any sequence similarity with *ipk1* genes of any crop.

>0617_596_001_PCR_A3F1_WP2_E01.ab1

TCCCAAAGATATAATAGACATAAAAATAAATTAATTCATGCAATTTCAGAACTTCATTG
ATGCATGTTAAATTTATTAATAACACATGTTTTCTTCATCTACATAATTCTTAATAGGCA
GCATTGCAAAATTCAGATCATAGGAGTAATGAAAATGGTTCAAGGGCAATATCATCAC
AAAAACAAACCAAGGAGATGAGCACATCAAGTACCTTATAGCAAAAAACTTTGGCCA

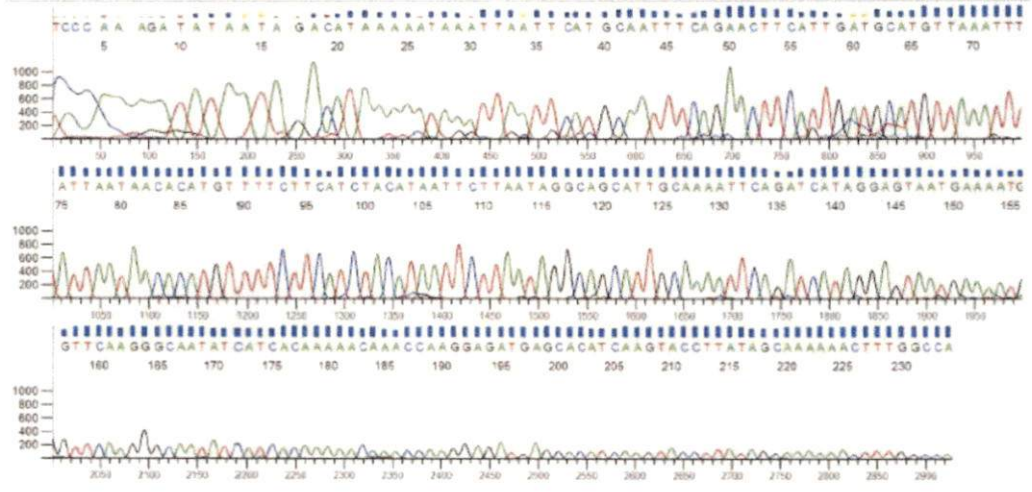


Fig. 9b: Sequence of 3'end of nested PCR product A3F1 sequenced with WP2

Sequencing of A2F1, the second product from 3' end of F6R5, was also carried out with the primer NLSPF1 but produced a sequence of only a short read length of 198 bp (Fig 9c). The sequence when pair wise aligned with A3F1 did not show any similarity with the sequence. BLAST homology search also did not produce any significant similarity with *ipk1* gene sequence of any crop.

>0617_713_001_PCR_A2F1_NLSPF1_E09.ab1

CAACAAAAAACAAAAAAAGCAAATTGTTTATTAGAAACGAGCATGTATGAAACGATG
CAAATTGTTAACAAAAAGCCAAGCAGGTGAGCAAAAGGAAAAAGAAAACCACACGT
AAAACGGTAAAAACTCAAACAAGCAAGCATCATTCATTCAAACGATGATGATCACTA
CCTCTCCCAGTTTGTGGAGAACTTTTTC

Fig. 9c. Sequence of nested PCR product A2F1

4.1.6.4 *Sequence analysis of genome walking at 5' flanking region of F6R5*

There were two products from 5'end selected for sequencing – A4R1 and A4R2. A4R1 was sequenced with the primer NLSPR1 and got a read length of 624 bp (Fig. 10a). The sequence was analyzed by BLAST homology search. No similarity was observed with the *ipk1* gene sequence of any crops.

>0617_713_002_PCR_A4R1_NLSPR1_F09.ab1

CGTCCCGGGTATAACCTCATTTTACAACAATTAAAAGGATTTTATACTACAAACAGGC
TAGGGTATTTCTTGGCCATCTCTTGTTCATTCTCCCAGGTACCTTCTTCTAATTCCTGAT
TTCTCCACAATACCTTTACCAAAGGGATCTCCTTGTTCCGAATTGTCTTATTTTCTCTTG
CCAGGATGCCAATCGGCTTTTCTTCGAAGGTCAAATCCAATTGCACCTCTAATTCTGC
AAAATTTGTGATGATATGATTCTCATCTCGTACATATTTCTTTACCTGTGACACATGGA
AAACATTGTGTACCCCATCAAGGTTTGGTGGTAATGCTAATCTATATGCACAATTTCT
GACCCAATCCAAAATAGGAAATATTCCTATGTATCTTGGCTTCAACTTCCCTGATGTTT
TAAATCTGATCACTCCCTTCTTTGGTGAGGTCTTTAGGTAAACAAACTCCCCAACTTG
AAACTCTATTTCCCTCCTTCAAGTGTCACCATAACTTTTTTGTCAATCTTGCGCTGCCT
TAACCCTTTGCTGAATGATTTGCACTTTTTCACCTGCTTCTCTCGCCATCTCTAGCCCG
CCCTATAGTGAGTCGTATTAAAAGCGACGGGTCCGG

Fig. 10a. Sequence of nested PCR product A4R1 sequenced with NLSPR1

The product A4R2 was sequenced with primer NLSPR2 and got a read length of 527 bp (Fig 10b). The homology search did not give any significant similarity with *ipk1* gene or with any sequences in the plant kingdom.

>0617_713_003_PCR_A4R2_NLSPR2_G09.ab1

TCACTCGAACATGGATAATTCGGCCACGGATAGTAACGTAATATGAAACATAATCAA
ATTTGTATGAATAGGATTAGTTAGAATTATTTTTACATTGTAATTATTTGACTATTAAT
GGATGTTGATGATTTCGTTGTATGTGATGTTGATTGAATTTGATGCATCATTTACTTAG
CGGGCATTAATATTTAGTTTTATTGTGTAGGTACGGAGTAAGAGCTCGCCACCTAGGC
ACGATGACACGCTTGGAGCTTCCACCAAGTCCGTTTATGACTCGTCTCTATTTTACTAT
CTATTCTATCAGTTGTTCGGTTTTTTACTTGTAGGACGGGTTAAGTGGTGTTTTGCCGG
ATTCTTTTTAACCCGGCCCATCTTAACTAACTGGTTTCCGAGATTAATAACCCTCGAGG
ATGAAATCCCTCCAGTTTGTTTTTCGAAATTCAGATCCAGTTCACGATGCCTTCGGGCC
TATTATATTAAATTCCCTAGCCAGCCCTATAGTGAGTCGTATTACAAGTGTAAGATTG

Fig. 10b: Sequence of nested PCR product A4R2 sequenced with NLSPR2

4.1.7 Assembly and analysis of sequence based on overlapping regions

Sequence assembly was carried out, confirming the sequence by comparing the results of at least two sequencing reactions. The assembled sequence had a length of 1072 bp (Fig.11) and was analyzed by BLAST homology search which showed similarity to *ipk1* genes of different crop species (Table 11). Analysis by ORF Finder for the coding region showed that the nucleotide produced six different open reading frames (ORFs) and the longest was ORF4 with 522bp, coding for 173 amino acids. The ORF belonged to 5'-3' frame 3 with the start and end site at 312 and 833 bp respectively (Fig 12a, b). Conserved domains search performed was found to show similarity with two super families- InsP5 kinase superfamily (cl05564) and Fab1

super family (cl27754). BLASTp homology search was conducted using the translated amino acid sequence obtained by analysis using ORF Finder (Table 12).

> F6R5- NLSPF1 sequence assembly
5'GTCTTTTCGGGTCTCAAATTTCATAGCTGAGATAACTTCATCAAAAAGTGTGTGTCTC
GTTTTAAAATGCACCAACCCTTGAAATTTCACAAGAAAGAGGTAATTTTCACCTGTGCT
TATATATTCCTTTGTTTAAAAGAAAAAACTTAGATTACACTAATTGTTGAGGATGCTTTT
CTTCATCCTTTCTGTACACATGGATTTGGCTATCATTTTCTAATTATTTTTCACTTCTCTC
TCAGATAGGGGAAATAAGTGAGTATGATCCAGTCGATCTGTTTTCAGGTTCTAAAGATA
GAATAAATCGTGCTATGAAATCTCTCTATGCCATCCCGCAAAACAATTTTCGTATCTTTT
TCAATGGCTCTTTGTTATTTGGTGGATTGGGGGGAGCCATGGATAATACGAAATCTAAA
TCCATGGAAGCTATTGATGATGAATTAGGTGTGTTAATTGATGCAGCACCTGGACTACG
ACTATCTAGATTTCTGGAACTTGTTTCATTGACAATTTTACAGTCAGATGTCCTGGTTGG
TCTTCTTAATACGCAAATGTTTGACCTTCTTGATATAGAAGGGGCAATTCATGCCTATTA
CAACATTATTTCTCAACCTAACATGGTCGGCAAGGTCCTAGGCGGCCTTTCTAAGAAGT
ACTTTTCCCTACATTATCTTTCTTTGGAAGCAAGTGTTAAGATTGTTCGCGATTTTTTGAT
CGCTGCAACTGCAAAGGACTGTAGCTTGATGATCAGTTTCAGACTGTTATCAGGGGAGA
TGATATCTGAAATTGATGTTGTGCACTTAGAGACAACTGGCCAAAGTTTTTTGCTATAA
GGTACTTGATGTGCTCATCTCCTTGGTTTGTTTTTGTGATGATATTGCCCTTGAACCATTT
TCATTACTCCTATGATCTGAATTTTGCAATGCTGCCTATTAAGAATTATGTAGATGAAGA
AAACATGTGTTATTAATAAATTTAACATGCATCAATGAAGTTCTGAAATTGCATGAATT
AATTTATTTTTATGTCTATTATATTTTGTGAAACCATTTTAAGGTTAGTGTTAAACACC3'

Fig. 11. *ipk1* sequence assembly using flanking sequences and F6R5 amplified from Panniyur 1 genomic DNA

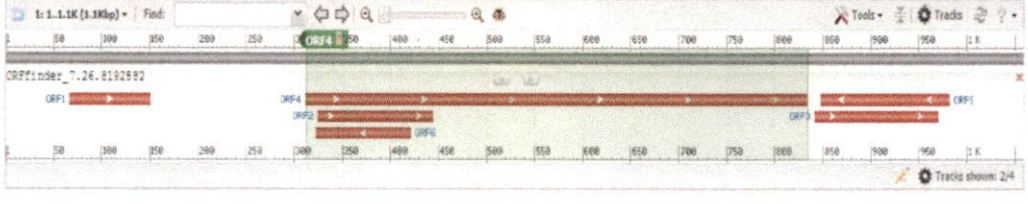


Fig. 12a: Overview of 6 ORFs assembled by ORF Finder

>lcl|ORF4:312:833 unnamed protein product

MKSLYAIPQNNFRIFFNGSLLFGGLGGAMDNTKSKSMEAIDDELGVLIDAAPGLRLSRFLE
LVSLTILQSDVLVGLLNTQMFDLLDIEGAIHAYYNIISQPNMVGKVLGGLSKKYFSLHYLS
LEASVKIVRDFLIAATAKDCSLMISFRLLSGEMISEIDVVHLETTGQSFLL

Fig. 12b. Amino acid sequence of ORF4

Table 11. Homology of assembled sequence with *ipk1* gene from different crops

Description	Max score	Query cover	E value	Identity	Accession
Predicted: *Vitis vinifera ipk1*, mRNA	154	48%	3e-35	67%	XM010666413.2
Predicted: *Carica papaya ipk1* like. mRNA	143	47%	8e-30	67%	XM022036758.1
Predicted: *Nelumbo nucifera ipk1* like, mRNA	141	47%	3e-29	67%	XM010256463.2
Predicted: *Phoenix dactylifera ipk1*, mRNA	132	47%	1e-26	66%	XM008797736.2
Morus notabilis ipk1 partial mRNA	131	47%	5e-26	67%	XM010111739.1
Predicted: *Musa acuminata subsp. malaccensis ipk1*,mRNA	122	47%	3e-23	66%	XM009391905.2
Predicted: *Manihot esculenta ipk1* like,mRNA	116	46%	1e-21	66%	XM021745380.1
Predicted: *Ricinus communis ipk1*, mRNA	114	47%	4e-21	66%	XM015720251.1
Predicted: *Ipomea nil ipk1* like, mRNA	113	47%	1e-20	66%	XM019296805.1
Predicted: *Populus euphratica ipk1* like,mRNA	104	55%	7e-18	65%	XM011030776.1
Predicted: *Camelina sativa ipk1*,mRNA	102	25%	2e-17	75%	XM010479163.2
Predicted: *Eucalyptus grandis ipk1*,mRNA	102	38%	2e-17	67%	XM010062987.2
Predicted: *Dendrobium catenatum ipk1,* mRNA	125	26%	2e-24	70%	XM020819929.1
Predicted: *Tarenaya hassleriana ipk1*,mRNA	93.3	28%	1e-14	74%	XM010542030.2
Predicted: *Cicer arietinum ipk1* like,mRNA	93.3	25%	1e-14	75%	XM004500934.2
Predicted:*Phalaenopsis equestris ipk1* like,mRNA	116	26%	1e-21	69%	XM020742471.1

Table 12. BLASTp hit of the predicted ORF with IPK1 of different crops

Description	Accession	Max score	Query cover	E value	Identity
Predicted: *Phoenix dactylifera* IPK1	XP 008795958.1	179	98%	3e-52	59%
Phalenopsis equestris IPK1 like	XP 020598130.1	172	97%	2e-51	56%
Predicted: *Gossypium arboretum* IPK1 like	XP 017611552.1	174	98%	2e-49	53%
Predicted: *Populous euphratica* IPK1 like	XP 011048791.1	173	98%	2e-49	54%
Predicted: *Malus domestica* IPK1 like	XP 017189442.1	174	98%	3e-49	53%
Predicted: *Nicotiana tabacum* IPK1 like	XP 016435252.1	172	98%	3e-49	52%
Gossypium hirsutum IPK1	XP 016687593.1	174	98%	4e-49	53%
Predicted: *Ipomea nil* IPK1 like	XP 019152350.1	172	98%	2e-48	53%
Predicted: *Elaeis guineensis* IPK1	XP 010940107.1	169	97%	2e-47	57%
Predicted: *Theobroma cacao* IPK1	XP 017981655.1	169	98%	2e-47	53%
Predicted: *Vitis vinifera* IPK1	XP 010664715.1	168	98%	5e-47	52%
Dendrobium catenatum IPK1	XP 020675588.1	167	97%	9e-47	52%
Ricinus communis putative IPK1	XP 015575737.1	166	97%	3e-46	52%
Predicted: *Musa acuminate* IPK1	XP 009390180.1	166	98%	4e-46	56%
Predicted: *Oryza sativa* IPK1	XP 015633432.1	164	98%	2e-45	53%
Predicted: *Brassica oleraceae* IPK1	XP 013606655.1	163	98%	3e-45	50%
Manihot esculenta IPK1 like	XP 021613502.1	163	98%	4e-45	50%
Predicted: *Nelumbo nucifera* IPK1 like	XP 010254763.1	162	98%	6e-45	52%
Sorghum bicolor IPK1 like	XP 021319382.1	162	98%	7e-45	53%
Zea mays IPK1	ONM 12613.1	157	98%	3e-45	52%

4.1.8 Phylogenetic analysis

In the BLAST homology search, *Pnipk1* sequence showed similarity with *ipk1* gene sequences of 22 different crops with an identity ranging from 65 to 82 percent and query coverage in the range of 7 to 49 percent. From this aligned sequence, those showing an identity of 65 percent or more and a query cover of at least 25 percent were selected for phylogenetic analysis (Table 13).

Table 13. List of nucleotides used for phylogeny tree construction

Organism	Accession No.
Manihot esculenta	XM 021745380.1
Ricinus communis	XM 015720256.1
Populus euphratica	XM 011030776.1
Vitis vinifera	XM 010666413.2
Morus notabilis	XM 010111739.1
Ipomea nil	XM 019296805.1
Nelumbo nucifera	XM 010256463.2
Phoenix dactylifera	XM 008797737.2
Musa acuminate	XM 009391905.2
Dendrobium catenatum	XM 020819929.1
Phalaenopsis equestris	XM 020742471.1

The phylogenetic tree was clustered into two main groups (Fig 13). Group I included sequences of *Manihot esculenta, Ricinus communis, Populus euphratica, Vitis vinifera, Morus notabilis, Ipomea nil* and *Nelumbo nucifera*. Group II contained 5 members including *Phoenix dactylifera, Musa acuminate, Dendrobium catenatum, Phalaenopsis equistris* and *Piper nigrum. P. nigrum* was clustered into a separate sub group II a, and the remaining members of the group belonged to the sub group II b.

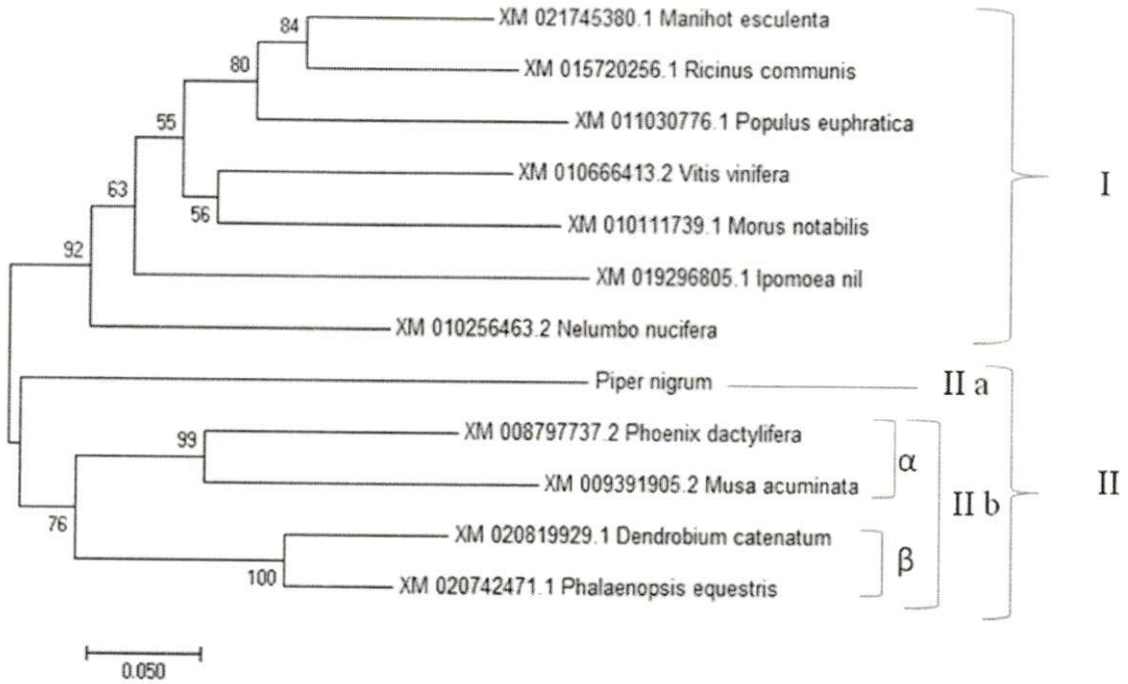


Fig. 13. Phylogeny tree constructed based on the aligned nucleotide sequences of selected selected crops with *Pnipk1* sequence

Evolutionary analysis was carried out using MEGA7 software and the phylogenetic tree was constructed using neighbour joining method giving a bootstrap replication value of 2000. Bootsrap values for each sub group are shown at the nodes. The main groups are denoted as I and II and the sub groups as a and b.

For construction of aminoacid based phylogenetic tree, the translated amino acid sequence obtained by the ORF Finder was used. BLASTp result showed similarity to IPK1 of 40 different crops with a query coverage ranging from 97 to 98 percent and identity of 49-59 percent. Those sequences showing 51 percent or more identity with the *Pnipk1* ORF sequence were selected for construction of pylogenetic tree (Table 14).

Table 14. List of amino acid sequences of different crops used for phylogeny tree construction

Organism	Accession No.
Predicted: Phoenix dactylifera	XP 008795958.1
Phalenopsis equestris	XP 020598130.1
Predicted: populous euphratica	XP 011048791.1
Malus domestica	XP 017189442.1
Nicotiana tabacum	XP 016435252.1
Gossypium hirsutum	XP 016687593.1
Ipomea nil	XP 019152350.1
Elaeis guineensis	XP 010940107.1
Vitis vinifera	XP 010664715.1
Ricinus communis	XP 015575737.1
Musa acuminate	XP 009390180.1
Oryza sativa	XP 015633432.1
Manihot esculenta	XP 021613502.1
Nelumbo nucifera	XP 010254763.1
Sorghum bicolor	XP 021319382.1
Zea mays	ONM 12613.1

The evolutionary tree constructed based on the amino acid sequences had 2 main groups (Fig. 14). Group 1 included sequences of *Ricinus communis, Maihot esculenta, Populus euphratica, Gossypium hirsutum, Vitis vinifera, Malus domestica, Nicotiana tabacum, Ipomea nil* and *Nelumbo nucifera.* Group 2 included the crops *Phalaenopsis equistris, Musa acuminata, Phoenix dactylifera, Elaies guineensis, Oryza sativa, Zea mays, Sorghum bicolor* and *Piper nigrum*. Group 2 was sub divided into 3 sub groups. Subgroup a included *P. equistris* and *P. nigrum*. *M. acuminata* belonged to a separate sub group b. *P. dactylifera, E. guineensis, O. sativa, Z. mays* and *S. bicolor* was clustered in subgroup c.

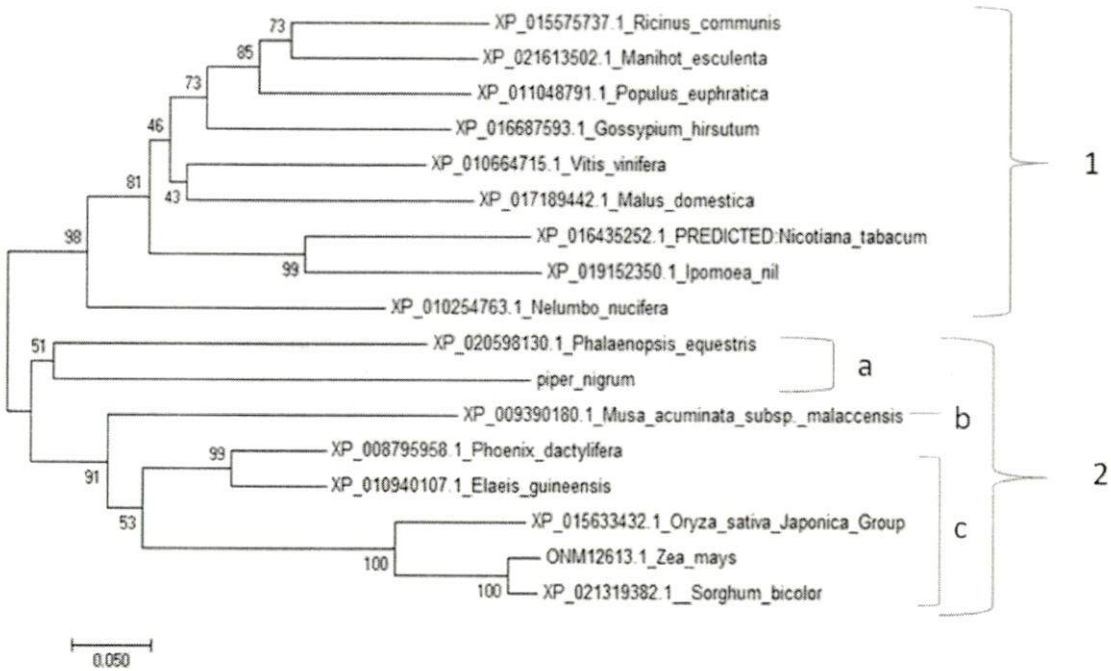


Fig. 14: Phylogram based on the aligned amino acid sequences of selected crops and translated aminoacid sequence of *Pnipk1*

Phylogenetic analysis was performed using MEGA software based on neighour joining method with a bootstrap replication of 2000. The bootstrap value corresponding to each sub group is shown at the nodes. The main groups are denoted as 1 and 2, sub groups are indicated as a, b, c.

4.2 PAGE detection of phytate

The presence of phytic acid was confirmed using PAGE analysis as per the protocol stated by Alimohammadi *et al.*, (2013) for the detection of $InsP_6$. For this, initially different concentrations (1-5nmoles) of standard phytate dipotassium salt were prepared and loaded on the gel for preparing the standard curve with quantity of phytate standards (nmoles) on x- axis and volume (intensity) on y –axis (Fig. 15). Band intensities of different concentrations of phytate standards loaded on the gel were recorded (Table 15).

Table 15. Band intensities of different concentrations of phytate standard

Quantity (nmoles)	Volume (intensity)
1	31376.6
2	41046.0
3	66763.7
4	93452.2
5	125539

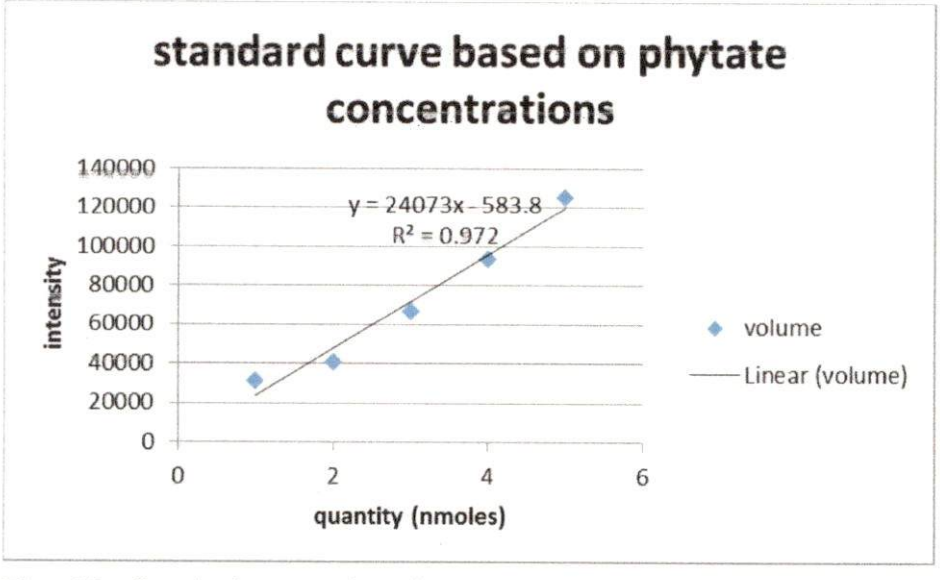


Fig. 15. Standard curve based on band intensities corresponding to different concentrations of phytate standards

X –axis : Concentration/quantity (nmoles) of phytate standards loaded on the gel

Y –axis : Band intensities (volume) corresponding to the quantity of standards loaded on gel.

PAGE analysis of the sample Panniyur 1 gave an acrylamide gel product with a single band corresponding to the band obtained in the standard phytate marker (Plate 7).

Quantity of phytate in sample was calculated using the standard curve (Volume + 583.8 = 24073 * quantity of $InsP_6$) and was expressed as nmoles $InsP_6$ per 20 µL sample. The value was further converted to nmoles/ g of leaf tissue. Quantity of $InsP_6$ in leaf tissue of Panniyur 1 was found to be 622 nmoles/ g.

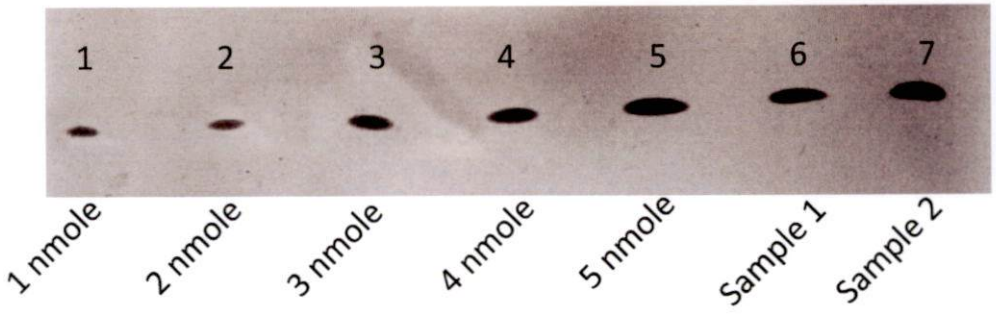


Plate 7: Polyacrylamide gel containing phytate standards and sample from black pepper

Lanes 1 to 5 : different concentrations of phytate standards (1-5 nmoles).

Lane 6 :10 µL phytate sample from Panniyur 1 leaf tissue.

Lane 7 : 20 µL phytate sample from Panniyur 1 leaf tissue.

Discussion

5. DISCUSSION

Black pepper is one of the most important spice crops, known as the king of spices. The crop has got multitude uses as a medicine as well as preservative and contributes to a major share of foreign exchange. Traditional Indian medicinal system uses black pepper for the treatment of chills, muscular pain, rheumatism and indigestion. Piperine, the major alkaloid in black pepper has got an anti-malarial property (Samuel *et al.*, 2016). Recent studies have proved its hepato-protective activity due to antioxidant property; anti-fungal, anti-carcinogenic and anti-diabetic properties of piperine (Gorgani *et al.,* 2017). The crop has therefore been proven as a perfect platform for the plant breeders to conduct genomic and crop developmental studies.

High throughput sequencing platforms used in next generation sequencing have resulted in whole genome sequencing of the model plants as well as most crop plants. This has served as a reference of the genes involved in various functions of the plant systems and has helped in comparison and analysis of the enormous quantity of sequence data under a common platform. The advent of PCR based *in-vitro* cloning techniques in the late 90's along with the sequence information made available through the whole genome sequencing projects have aided in gene discovery and analysis of unraveled genes.

In a previous experiment, a unigene *viz.*, *Pnc135* was obtained from Expressed Sequence Tags (ESTs) generated using Ion Torrent technology (Ion Personal Genome Machine™, Life Technologies). The 995bp long coding sequence prediction showed that the exon size is from 1 to 941 bases. Sequence analysis of this unigene revealed that it is a putative gene coding for the enzyme inositol pentakisphosphate 2-kinase involved in the biosynthetic pathway of Inositol hexakisphosphate ($InsP_6$), commonly called phytic acid (Sujatha, R., unpublished data). The functional properties of phytic acid in plant system are reviewed in chapter 2 'Review of Literature'.

As there are no references about the detection of gene involved in synthesis of phytic acid in black pepper so far, the study was undertaken for the identification of the putative phytic acid gene in black pepper variety Panniyur 1, considering its immense potential metabolic functions in plant systems. The work is therefore expected to open way for future crop developmental studies.

In the present investigation, genome walking was used to decipher the unknown sequences on the upstream and downstream positions of the unigene *Pnc135*. For this, a novel method suggested by Reddy *et al.* (2008) using Phi29 DNA polymerase based rolling circle amplification of the whole genome followed by locus specific amplification was used, instead of the conventional methods like inverse PCR, ligation mediated PCR, randomly primed PCR etc. The advantage of this method is that it is a high–throughput protocol wherein the restriction fragmentation of the genomic DNA is not required and the possibility of amplifying the "off-tragets" is less (Reddy *et al.*, 2008).

In order to detect the presence of phytic acid in black pepper, biochemical analysis was also carried out using crude extract from leaf sample by comparing with the commercially available standard phytate as marker using Polyacrylamide Gel Electrophoresis (PAGE).

5.1 GENOME WALKING

The present study was taken up for deciphering the sequence of the gene for inositol pentakisphosphate 2-kinase (*ipk1*) which is active in both lipid dependent and lipid independent pathways and is involved in the last step of phytic acid metabolism (Gonzalez *et al.*, 2010). The gene *ipk1* has been first described by Biswas *et al.* (1978) and was purified for the first time from immature soybean seeds (Phillipy *et al.*, 1994).

Genomic DNA isolation is the preliminary step involved in any genetic studies. A good quality DNA without any cross contamination is required as a

primary starting material. Genomic DNA was isolated using modified CTAB method (Doyle and Doyle, 1987).

5.1.1 Confirmation of presence of unigene sequence and amplification of flanking regions

The first step of *in vitro* gene cloning using PCR involves primer designing. Specific sequence and optimal concentration of primer is the most essential element determining the efficiency of PCR and there are numerous bioinformatics tools that efficiently help in primer designing and analysis (Abd-Elsalam, 2003). The sequence of unigene *Pnc135* obtained by the contig alignment of ESTs expressed during seed development in black pepper was used as the base sequence in the present investigation. Sequence similarity of the unigene with *ipk1* was confirmed once again before designing the primers by carrying out BLASTn homology search with available sequences in the database and it was showing similarity with *ipk1* gene from around 28 different crops. Hence, primers for the confirmation of sequence information from the genomic DNA of Panniyur 1 were designed based on the unigene *Pnc135*. Simultaneously, primers were designed for the amplification of flanking regions of the unigene by multiple sequence alignment of the selected crops through BLAST search. Primers were designed manually. Software OligoCalc was used for checking the quality requirements of the primers like length of the sequence, GC content, melting temperature and absence of any hairpin loops. In case of the primers designed for the amplification of flanking regions, partial degeneracy was given to the primers designed by multiple sequence alignment. In a similar work conducted by Sun *et al.* (2007), for the amplification of *ipk1* gene sequence of *Zea mays* also the candidate *ipk1* gene sequence was searched against maize EST database and this sequence was used to design primers for the amplification of *Zmipk1* cDNA. Roorkiwal *et al.* (2014) also reported primer designing for amplification of candidate genes using EST data.

The designed primers were used for the amplification of genomic DNA and the products were sequenced to compare the result. The sequence result of PCR product amplified by the primers designed for the confirmation of unigene *Pnc135*

sequence gave an expected product of size 933 bp but did not produce any similarity with the unigene sequence. This may be due to the fact that there might have been some misalignment of the contig data while assembled to produce the unigene *Pnc135*. There is also a chance that the terminal ends of the sequence may belong to the coding regions of some other genes. Reports of Huang and Madan (1999) and review by Pop (2009) also is in agreement with the fact that different assemblers use different algorithms to overcome the errors that may occur during sequence assemblies and the assembled data may often be erroneous either due to limitations of the assembly algorithms or due to the wrong information provided to the assemblers. This may be the possible reason for not getting amplifications at the specified annealing temperatures. Amplification was produced at a lower temperature and therefore the amplification might have been non-specific. The amplification of the flanking regions gave products of unexpected sizes – 238 bp for F3R3 and 272 bp for F4R4a. These products also did not show similarity with the *ipk1* gene sequences of any crops. This confirms that the terminal ends of the unigene *Pnc135* may be part of other gene sequences and as the primer sequences were designed based on the terminal ends, nonspecific amplifications were produced.

5.1.2 Detection of conserved regions of *ipk1* gene sequences by multiple sequence alignment with unigene *Pnc135* and isolation of *ipk1* gene fragment from genomic DNA of Panniyur 1

Since the preliminary sequencing results were not in agreement with the expected results, new primers were designed for the amplification of *ipk1* gene using consensus regions obtained by multiple sequence alignment of *Pnc135* sequence with *ipk1* sequences of 28 different crop species. As none of the crops related to black pepper were sequenced for the same gene so far, multiple sequence alignment was carried out with genetically diverse group of crops including cereals, pulses and vegetables. Three sets of primers - PNC F5/PNC R5, PNC F6/PNC R6 and PNC F7/PNC R7 were designed based on the consensus regions in the multiple sequence alignment/pair wise alignment as detailed in chapter 4. Ovcharenko *et al.* (2005) reported that multiple-sequence alignment analysis is a powerful approach for

detecting functional regulatory elements and understanding phylogenetic relationships and for annotating the genes. Brodin *et al.*, (2013) also pointed out that it is important to locate primers in relatively conserved regions bounding the region of interest, and to design primers.

When different combinations of these three primer sets were used for amplifying the genomic DNA of Panniyur 1, amplifications of varying sizes were generated. Among them, three products – F6R5, F7R6 and F7R7 were selected for sequencing. The amplicons F7R6 and F7R7 were of expected size whereas F6R5, though not of expected size, was a single band of good intensity and was amplified from primers designed based on the consensus regions in all the sequences used for multiple alignment. The sequencing of the 400 bp F6R5 gave a 137 bp sequence showing homology with the *ipk1* gene of various crops. This sequence was further used for designing primers for genome walking of the flanking regions. The other two amplicons (F7R6 and F7R7) were not showing any similarity with *ipk1* gene and hence were not used for further analysis. This might be due to the fact that the primers PNC R6 and PNC R7 that were designed based on the pairwise alignment may be from the terminal regions of the unigene *Pnc135* which was from the region of misalignment of the contig. The primer sequence may also be present elsewhere in the genomic DNA as the primers used for the amplification are short oligonucleotides that can be present in more than one region in the whole genome.

5.1.3 Rolling circle amplification (RCA) of whole genome and amplification of flanking regions of *ipk1* gene using locus specific primers

Genome walking is a method used for the determination of unknown sequences adjacent to a known region and often uses inverse PCR comprising of restriction ligations (Myrick and Gelbart, 2002). There are modifications of genome walking protocols available today, using random primers and rolling circle amplification, avoiding the use of restriction- ligation of genomic DNA. The genome walking protocols are often used for sequencing the upstream regions of a gene including the promoter region. Rishi *et al.* (2004) have reported that in restriction

mediated genome walking protocol, partial digestion of genomic DNA with the restriction enzymes could generate larger DNA fragments. This resulted in larger genome walking results as the restriction sites were far away from the gene specific primers designed and generated specific amplification compared to non specific, smaller amplicons obtained when complete digestion of genomic DNA was done. Adaptor cassettes along with biotinylated specific primers were used in combination with a second round of PCR amplification using nested primers to exclude non-specific amplification by Reddy *et al.* (2002). They could successfully amplify the upstream ends of genes of pea *topoisomerase I* and *II*, *phospholipase C*, tobacco *topoisomerase I* and *II* and *glyaxylase* gene of *Brassica juncea*. Tsaftaris *et al.* (2010) have reported a genome walking method generating circular genomic DNA fragments followed by rolling circle amplification by phi 29 DNA polymerase. They could successfully isolate the promoter regions of MADS box and flowering genes in *Crocus sativus*.

In the present investigation, genome walking by a novel method suggested by Reddy *et al.* (2008) was used. The amplification of Panniyur 1 DNA was done using Phi29 DNA polymerase and four different primers to synthesize overlapping fragments containing walker adapter sites for selective amplification, later using locus specific primers. Among the various primer combinations tried for amplifying *ipk1* gene segment from genomic DNA of Panniyur 1, the best result was obtained for the product F6R5. Hence for genome walking the locus specific primers (LSP) and the nested locus specific primers (NLSP) were designed based on the sequence of F6R5. Genome walking using phi29 DNA polymerase was preferred over other techniques for amplification of flanking regions in the present study due to its ability to produce large overlapping fragments of the whole genome by multiple strand displacement reaction from very low concentrations of template DNA (Reddy *et al.*, 2008). The protocol has also got an advantage that the process is isothermal and therefore does not rely on the use of thermo cyclers and requires a very short time (1-2 h) for the reaction. This has particularly made phi29 mediated strand displacement amplification (SDA) an effective means of whole genome amplification (WGA) of micro-organisms or viruses and for sequencing of different vector constructs (Detter

et al., 2002; Sujayanont *et al.,* 2014). Also, as reported in the protocol by Reddy *et al.* (2008), the two- step PCR involving the use of nested primers for the second round of PCR is expected to eliminate the non specific amplifications.

5.1.4 Assembly and analysis of sequence based on overlapping regions

Sequencing of the product obtained by genome walking gave a 933bp fragment (A3F1) towards the 3' end flanking the sequence of F6R5. The BLAST homology search of the sequence showed an identity of 67 per cent with the *ipk1* sequence of *Vitis vinifera* (query coverage 55%, XM010666413.2), *Carica papaya* (XM022036758.1, query coverage 54%) and *Nelumbo nucifera* (XM010256463.2, query cover 54%) and also with many other crops. The sequence was compared with F6R5 for overlap and was assembled (*Pnipk1*) by joining the overlapping ends and a total sequence length of 1072 bp was obtained. BLAST homology search of the sequence showed similarity with *ipk1* gene sequences of crops like *Vitis vinifera* (XM010666413.2) with a highest query coverage of 48 per cent and identity of 67 per cent, followed by *Nelumbo nucifera* (XM010256463.2), *Phoenix dactylifera* (XM008797736.2), *Musa acuminata* (XM009391905.2) with 47 per cent query cover and 66 per cent identity and so on. As per the protocol by Reddy *et al.* (2008), the walker primer may attach to multiple sites and may therefore result in production of multiple bands. The fragments thus obtained may have the NLSP at one end and WP2 at the other end and the size of the segment may vary depending on the position of attachment of WP2. According to the principle, the fragments should be subsets of the same gene. However the 5' flanking region did not give any sequence similarity with *ipk1* gene. The non-specificity of amplification is expected to be due to the partial degeneracy of the walker primer used in genome walking. The amplification might have started from the walker primer binding region which may be at non-specific or undesired sites. Also, multiple bands were produced in the amplification products. There might have been multiple bands of the same size present in the sequenced products which might have hindered the generation of accurate sequence results. A further extension of the work has to be carried out to sequence the 5' flanking region. A modification of the current protocol, if needed also has to be done to rectify the drawback of non-specific amplification in the present work.

Yu (2013) have reported that *Ricinus communis RcIPK1* gene sequence is 6427 bp long consisting of 11 exons and 10 introns. The cDNA sequence is found 2088 bp long coding 695 amino acids and they have isolated mRNA variants ranging in size from 1665 to 2027 bp and peptides of size 44 and 60 kDa from the embryo. Gene of *Glycine max GmIPK1* is found ~3.3 kb in size and the mRNA is 1541 bp with seven exons and 6 introns (Yuan *et al.*, 2012). *Zea mays ZmIPK1* cDNA clone having a size of 2012 bp and an ORF of 440 amino acids was isolated by Sun *et al.* (2007) and was found to have two mRNA variants. *Phaseolus vulgaris* had an amplified genome sequence of 2,258 bp with an ORF coding for 352 amino acids corresponding to exons 1 to 6 (Fileppi *et al.*, 2010). Reports of various works conducted clearly shows that the genomic size and cDNA construct vary widely in crops belonging to diverse groups and show similarity between the members of related family. Therefore sequence information about the complete gene and its functional cDNA construct is required for the further analysis and structural comparison with gene sequences of other reported crops.

5.1.5 Phylogenetic analysis

The coding region of the nucleotide sequence was found by using the software ORF Finder. The result displayed an ORF with 522 bp encoding 173 amino acids showing similarity to IPK1 of various crops. Amino acid sequence analysis for conserved domains showed the presence of $InsP_5$-2-kinase domain and *Fab1* superfamily domain. Reports show that *Fab1* gene homologues have got a phosphatidylinositol (PtdIns) 3 kinase activity forming polyphospho inositide (PtdIns (3, 5) P_2) in *Arabidopsis*, yeast and *Drosophila* and, *Arabidopsis* is reported to have 4 homologues of FAB1. PtdIns 3 kinase is also reported to have additional functions in vesicle trafficking and maintenance of vacuolar homeostasis (Yamamoto *et al.*, 1995; Odorizzi *et al.*, 1998; Gary *et al.*, 2002; Rusten *et al.*, 2006; Munnik and Vermeer, 2010). Thus the amino acid sequence is confirmed to be a part of the family of proteins involved in phytic acid metabolic pathway.

The nucleotide sequence as well as the amino acid sequence were subjected to phylogenetic analysis to know the evolutionary relationships with other crops. In the present study, *Pnipk1* was subjected to phylogenetic analysis using selected PA gene sequences from BLAST search to understand the evolutionary relationships among the selected crop species. The evolutionary tree analysis showed that *ipk1* gene sequence of pepper was closer to the sequences of *Musa acuminata*, *Phoenix dactylifera, Dendrobium catenatum* and *Phalaenopsis equistris*. Even though the nucleotide sequence of *Pnipk1* showed greater similarity with sequences of *Vitis* and *Nelumbo*, they were clustered in two different groups in the evolutionary tree. The close proximity in the positioning of phylogenetically related crops like *Ricinus* and *Manihot* as well as *Dendrobium* and *Phalaenopsis* in the tree constructed indicate that the clustering was done accurately. Similarly, in the phylogenetic tree based on amino acid sequence also, related crops like *Oryza, Zea* and *Sorghum* as well as *Phoenix* and *Elaeis* were clustered together. This evolutionary analysis has helped in identifying the crops with more similarity to the sequence of *Pnipk1* and therefore provides a closer reference of sequences for further molecular studies. Even though similar works were conducted in the isolation and characterization of *ipk1* gene of different crops, the evolutionary analysis of the gene was not carried out by many. Sweetman *et al.* (2006) have carried out the phylogenetic study of *Arabidopsis thaliana IPK1* with the protein sequences of *Saccharomyces cerevisiae* and *Homo sapiens* and Yu (2013) have studied the phylogenetic relationship of *Ricinus communis IPK1* with the model organisms like *Homo sapiens, Zea mays, Saccharomyces cerevisiae* and *Arabidopsis thaliana* based on the Clustal Omega multiple sequence alignment. Similar phylogenetic studies were carried out in other related genes of the phytate biosynthesis gene family also. Fileppi *et al.* (2010) have reported that in phylogenetic analysis of myoinositol 3-phosphate synthase (*MIPS*) protein of *Phaseolus vulgaris* with related crops *Oryza sativa* (*OsMIPS*) and *Zea mays* (*ZmMIPS*) were clustered to the same sub group. Similar report was obtained for Inositol triskisphosphate 5/6- kinase (*ITP5/6K*) also (Suzuki *et al.*, 2007). Ajay *et al.* (2016) have carried out the phylogenetic analysis of *Arachis hypogea* inositol tetraphosphate 1-kinase (*AhITPK1*) and inositol 1, 4, 5 trisphosphate kinase (*AhIPK2*) proteins with other crop species and was found that the sequences of crops

like *Ricinus, Populous, Vitis* belonged to the same subgroup in the case of *AhIPK2* and *Zea mays* and *Oryza sativa* were clustered together in the case of *AhITPK1*. These results were on par with the analysis carried out in the present study for *PnIPK1* sequence. The results therefore confer the fact that the phytate biosynthesis pathway gene family expresses a similar evolutionary relationship with each other.

5.2 PAGE DETECTION OF PHYTIC ACID

Instead of the most commonly used HPLC and NMR spectroscopy methods, a simple biochemical analysis using polyacrylamide gel electrophoresis (PAGE) was done for the easy and quick detection of $InsP_6$, as reported by Losito *et al.* (2009). The method points out that PAGE analysis can effectively distinguish the various inositol pyrophosphates and its precursors including $InsP_5$, $InsP_6$ in the samples. They had performed both 4, 6-diamidino-2-phenylindole (DAPI) and toluidine blue staining of the gels and results showed that toluidine blue could be used to detect inositol pentkisphosphate ($InsP_5$), inositol hexakisphosphate ($InsP_6$) and inositol pyrophosphate ($InsP_7$) whereas DAPI staining could not identify inositol pyrophosphates like $InsP_7$. Results also showed that DAPI staining was more sensitive compared to toluidine blue but could detect $InsP_5$ and $InsP_6$ up to a limit of 100 nmoles only.

Leaf tissue was used in the present study for detecting the presence of phytic acid in the sample. The quantity of phytic acid in the sample was calculated from the standard curve plotted against the band intensities and quantity of phytate standards loaded on to the gel. In the current work, Panniyur 1 leaf sample was estimated to have around 620 nmoles of PA per gram fresh weight (FW). In a similar work conducted by Alimohammadi *et al.* (2013), 100 nmoles/g FW of phytic acid was detected in tomato, 200 nmoles/g FW of phytic acid in tobacco and 450 nmoles/g FW of phytic acid in rice. The data showed that phytate levels in plants may vary between wide ranges.

Phillipy *et al.* (2015) have reported that HPLC detection of $InsP_6$ in different plants show the presence of considerable amount of $InsP_6$ compared to $InsP_5$. They

have also reported that plants belonging to Malvaceae family had higher levels of $InsP_6$ compared to other crops. Maize and arabidopsis leaf tissues had 100 and 110 µM phytate respectively and cotton, okra and hibiscus sp. belonging to Malvacea family had a higher phytate concentration in the range of 10-170, 90 -340 and 30-500 µM respectively according to their report. They also related the hardiness and non-branching nature of Malvaceae members of plants that may be possibly regulated by the auxins to higher concentrations of $InsP_6$ in those crops. Aberoumand (2011) have reported that estimation of phytate precipitated by ferric chloride in different wild edible plants detected a phytate content of 3.4 mg/g, 8.2 mg/g, and 6.9 mg/g in *Asparagus officinalis*, *Portulaca oleracea* and *Solanum indicum,* respectively.

Even though PAGE detection can be effectively carried out for confirming the presence of InsPs in plant samples, HPLC method of detection is found to be the most accurate and reliable one (Alimohammadi *et al.,* 2013). PAGE analysis and ferric chloride precipitation of phytates may lead to either underestimation or over estimation of the quantity of phytate in the sample and can therefore lead to wrong conclusions. Comparison of the phytate levels estimated by PAGE analysis in Panniyur 1 with the quantity of phytate detected in other reported crops like rice, tobacco and tomato show that black pepper had a higher level of phytate (620 nmoles/g) compared to the 450 nmoles/g in rice, 200 nmoles/g in tobacco and 100nmoles/g in tomato. Also, the phytate contents were estimated as high as 340 µM in okra and 500 µM in hibiscus sp. using HPLC method.

In the present study, phytate estimation was carried out using PAGE analysis as a preliminary confirmatory test for detection of the presence of phytate in the sample. Analysis was done using the leaf tissues of mature vines. Since reports show that the concentration of phytate may vary in different tissues and within the same tissue depending on the developmental stage (Phillipy *et al.*, 2015), further evaluation on the variation in phytate in various tissues at different developmental stages of the plant should be done. The total phosphorus content of the plant tissue

and the percentage of phytate phosphorus in the tissues should also be determined to get an accurate estimate of phytic acid present in the samples.

The present study has provided a preliminary confirmation of the *ipk1* gene and detected the presence of $InsP_6$ in black pepper. Future works have to be extended in order to sequence the 5' flanking region of the *Pnipk1* gene sequence so as to decipher the full length gene of *ipk1*. This would help in the expression studies of genes involved in the biosynthetic pathway of phytic acid which will give an insight of the possible roles played by the biomolecule in plant system, especially in plant defense mechanisms. The detection of higher phosphate derivatives like $InsP_7$ having rather more important roles in protection against biotic as well as abiotic stresses has to be also carried out. The knowledge on the sequence information of the full length gene will help in designing markers that can identify crops with altered levels of phytic acid and can be effectively utilized in the breeding programmes involving marker assisted selection. This will help in breeding of crops with biotic and abiotic stress tolerance. As genetically modified food crops are not permitted to use in the country, genes of $InsP_6$ and its higher derivatives found to provide resistance to plants, can be used in cis-genics involving related crops.

Summary

6. SUMMARY

Black pepper *(Piper nigrum* L.*)* is one of the major spice crops and contributes a major share of country's foreign exchange. The crop has got other diversified uses such as medicinal and as preservative. Being a high value crop, it provides a perfect platform for conducting crop improvement studies.

In a previous work conducted in black pepper, ESTs were generated at the berry development stage and one of the assembled unigenes, *Pnc135* (995bp) showed sequence similarity to the *ipk1* gene sequences of different crops. This gene is coding for the enzyme inositol pentakisphosphate 2-kinase involved in the biosynthetic pathway of Inositol hexakisphosphate ($InsP_6$), commonly called phytic acid which is a natural phosphate storage present in all eukaryotic cells and has got multifaceted roles in the living system. It is found to help in the essential physiological functions of plants like germination, growth and development of seedlings and provide resistance against various biotic as well as abiotic stresses.

Considering the agronomic value of the crop and the physiologically important roles played by the compound, the study on 'Genome walking for putative phytic acid ($InsP_6$) unigene in black pepper (*Piper nigrum* L.)' was proposed with an objective to identify the gene coding for $InsP_6$, validate the gene and detect the presence of $InsP_6$ in black pepper variety Panniyur 1.

1. DNA isolation from Panniyur 1 was carried out for the molecular studies.

- DNA was isolated using CTAB method (Doyle and Doyle, 1987) with a slight modification done in the lab *viz*., precipitation of DNA by using one and half times volume of chilled isopropanol and half volume of 5 M NaCl.
- Quantification of DNA using biophotometer gave an absorbance ($A_{260/280}$) value of 1.9 indicating good quality DNA.
- Concentration of DNA was 5.05 µg/µl and was diluted to 100 ng/µl and used as template for PCR reactions.

2. Confirmation of the unigene *Pnc135* (995 bp) in black pepper and amplification of flanking regions of *Pnc135* unigene was performed.

- Primers (PNC F1/PNC R1 and PNC F2/PNC R2) were designed based on the unigene sequence to re-amplify the sequence and confirm the sequence information.
- Only PNC F1/PNC R1 produced amplification (933 bp) and the sequencing did not produce any similarity with the unigene sequence.
- BLAST search was performed with *Pnc135* as query sequence and found that the sequence belonged to antisense strand. Reverse complement of the unigene sequence was used for designing primers for the amplification of flanking regions.
- Primers PNC R3 and PNC F4 were designed from the 5' and 3' ends of the reverse complement of *Pnc135* respectively. PNC F3 was designed based on the consensus regions of *Nelumbo nucifera* and *Vitis vinifera* based on the multiple sequence alignment of the sequences with *Pnc135* and was used for the amplification of 5' flanking region. PNC R4a and PNC R4b were designed for the amplification of 3' flanking region based on the consensus regions in *Cicer arietinum* and *Sesamum indicum* in multiple sequence alignment.
- The PCR product F3R3 was 300 bp in size and was not of expected size (700 bp). F4R4a was 320 bp in size and was in the expected size range.
- Sequencing of the products F3R3 and F4R4a was carried out but did not show any similarity with the *ipk1* gene sequences of any crop.

3. As the preliminary experiment did not produce any expected results, amplification of *ipk1* gene sequence in black pepper was done using the primers designed based on conserved regions from *Pnc135* and the gene sequence reported from 28 other crops by multiple sequence alignment.

- Three sets of primers were designed. PNC F5 and PNC R6 were based on pairwise alignment of *Pnc135* with *N. nucifera*. PNC F6, PNC F7 and PNC R5 were designed based on the multiple sequence alignment of *Pnc135* with other 28 crop species. PNC R7 was designed based on the pairwise alignment of *Pnc135* with *Elaeis guineensis.*
- PCR amplification was carried out with combinations of forward and reverse primers and the products obtained were of varying sizes.
- PCR products F6R5, F7R6 and F7R7 were selected for sequencing as F7R6 (300 bp) and F7R7 (400 bp) were in the expected size range of 330 bp and 370 bp respectively. F6R5 (400 bp) was selected even though it was not in the expected size (250 bp) as both the primers used for amplification were based on the multiple sequence alignment result.
- Sequencing of F6R5, F7R6 and F7R7 gave results of 395 bp, 193 bp and 297 bp sequences and BLAST search showed that F6R5 was homologous to *ipk1* gene of other crops. F7R6 and F7R7 did not show any similarity with *ipk1* gene.

4. Genome walking for the flanking regions of F6R5 was carried out based on the protocol by Reddy *et al.* (2008).

- Two sets of forward (LSP F1, LSP F2) and reverse primers (LSP R1, LSP R2) as well as their nested primers (NLSP F1, NLSP F2, NLSP R1 and NLSP R2) were designed based on F6R5 sequence. The walker adapter primers (WA1, WA2 WA3 and WA4) and walker primer (WP1) and nested walker primer (WP2) were designed as mentioned in the protocol by Reddy *et al.* (2008).
- Rolling circle amplification (RCA) of black pepper genomic DNA (100 ng) was done using walker adapters (WA1, WA2, WA3, and WA4). The RCA products were further amplified by locus specific PCR (using walker primer1 (WP1) and locus specific primers LSP F1, LSP F2, LSP R1, LSP R2) and the nested PCR (using walker primer 2 (WP2) and nested locus specific primers NLSP F1, NLSP F2, NLSP R1, and NLSP R2).

- Amplification with NLSP F1 gave 1 kb product (A3F1) towards the 3' flanking region of F6R5 in two separate RCA and locus specific PCR reactions and was sequenced.
- Towards the 5' flanking region, PCR amplification using NLSP R1 and NLSP R2 gave multiple bands of varying sizes and 750 bp product of NLSP R1 and 600 bp product of NLSP R2 were sequenced.
- Sequence result of A3F1 gave a similarity to *ipk1* gene of other crops. The BLAST homology search of the sequence showed an identity of 67 percent with the *ipk1* sequence of *Vitis vinifera*, *Carica papaya* and *Nelumbo nucifera* and also with many other crops. None of the other products (A2F1 at 3'end; A4R1 and A4R2 at 5'end) showed any similarity with *ipk1* gene.
- The sequences of F6R5 and A3F1 were assembled to get a total length of 1072 bp (*Pnipk1*).

5. Analysis using ORF Finder was done to find out the coding regions of the sequence.

   - The coding region (Open Reading Frame) was found to be 522 bp long, extending from the base 312 to 833.
   - The ORF coded for 173 amino acids and was found to contain two conserved domains- InsP5 kinase and Fab1 super family.

6. Phylogenetic analysis was conducted to understand the evolutionary relationship of *Pnipk1* with *ipk1* nucleotide and amino acid sequences of other crops.

   - The analysis showed that *Pnipk1* is more closely related to the sequences of *Musa acuminata, Phoenix dactylifera, Phalaenopsis equistris* etc.

7. Biochemical analysis of $InsP_6$ using Polyacrylamide gel electrophoresis (PAGE) was done to detect the presence of $InsP_6$ in Panniyur 1 leaf samples.

- PAGE detection gave positive results showing bands corresponding to the phytate standards loaded on the gel.
- Standard curve was drawn plotting the band intensities corresponding to different concentrations of phytate standard recorded using the software GelQuant.NET and was used for quantifying phytate in the sample.
- Quantity of phytate was estimated to be 620 nmoles/g fresh weight of leaf sample.

References

Abd-Elsalam, K. A. 2003. Bioinformatic tools and guideline for PCR primer design. *Afr. J. Biotech.* **2**(5): 91-95.

Abe, T. K., and Taniguchi, M. 2014. Identification of myo-inositol hexakisphosphate (IP6) as a β-secretase 1 (BACE1) inhibitory molecule in rice grain extract and digest. *FEBS Open Bio*. **4**(1): 162-167.

Aberoumand, A. 2011. Protein, fat, calories, minerals, phytic acid and phenolic in some plant foods based diet. *J. Food Processing Technol.* **2**(3): 1-4.

Adachi, E., Shimamura, K., Wakamatsu, S., and Kodama, H. 2004. Amplification of plant genomic DNA by Phi29 DNA polymerase for use in physical mapping of the hypermethylated genomic region. *Plant Cell Rep.* **23**(3): 144-147.

Aggarwal, S., Shukla, V., Bhati, K. K., Kaur, M., Sharma, S., Singh, A., Mantri, S., and Pandey, A. K. 2015. Hormonal regulation and expression profiles of wheat genes involved during phytic acid biosynthesis pathway. *Plants* **4**(2): 298-319.

Ahmad, N., Fazal, H., Abbasi, B. H., Farooq, S., Ali, M., and Khan, M. A. 2012. Biological role of *Piper nigrum L.*(Black pepper): A review. *Asian Pacific J. Trop Biomed.* **2**(3): 1945-1953.

Ahmed, R. R., and Ebaid, H. 2015. Possible causes of ileal injury in two models of microbial sepsis and protective effect of phytic acid. *Iranian J. Med. Sci.* **35**(1): 40-47.

Ajay, B. C., Bera, S. K., Devaiah, K., James, O., Violetka, T., and Anthony, A. 2016. Isolation and characterization of inositol tetraphosphate 1-kinase

(AhITPK1) and Iinositol 1, 4, 5-tris-phosphate kinase (AhIPK2) gene in peanut. *Cloning Transgenesis* **5**(1): 1-9.

Alimohammadi, M., Ali, N., and Khodakovskaya, M. 2013. Quantitative detection of inositol hexakisphosphate (InsP6) in crop plants using polyacrylamide gel electrophoresis (PAGE). *Am. J. Plant Sci.* **4**: 1-6.

Ali, N., Paul, S., Gayen, D., Sarkar, S. N., Datta, K., Datta, S. K. 2013. Development of low phytate rice by RNAi mediated seed-specific silencing of inositol 1,3,4,5,6-pentakisphosphate 2-kinase gene (*IPK1*). *PLoS ONE* **8**(7): 68161.

Anekonda, T. S., Wadsworth, T. L., Sabin, R., Frahler, K., Harris, C., Petriko, B., Ralle, M., Woltjer, R. and Quinn, J.F. 2011. Phytic acid as a potential treatment for alzheimer's pathology: evidence from animal and in vitro models. *J. Alzheimer's Dis.* **23**(1): 21-35.

Bhattacharya, S., and Bandyopadhyay, A. 2017. Performance of bush pepper under new alluvial zone of west bengal, India. *Int. J. Curr. Microbiol. App. Sci.* **6**(7): 2546-2550.

Bauman, A. T., Chateauneuf, G. M., Boyd, B. R., Brown, R. E., and Murthy, P. P. 1999. Conformational inversion processes in phytic acid: NMR spectroscopic and molecular modeling studies. *Tetrahedron Letters* **40**(24): 4489-4492.

Berman, A. J., Kamtekar, S., Goodman, J. L., Lázaro, J. M., de Vega, M., Blanco, L., Salas, M., and Steitz, T.A. 2007. Structures of phi29 DNA polymerase complexed with substrate: the mechanism of translocation in B-family polymerases. *Eur. Mol. Biol. Organ. J.* **26**(14): 3494-3505.

Bhandari, R., Juluri, K. R., Resnick, A. C., and Snyder, S. H. 2008. Gene deletion of inositol hexakisphosphate kinase 1 reveals inositol pyrophosphate regulation of insulin secretion, growth, and spermiogenesis. *Proc. Natl. Acad. Sci.* **105**(7): 2349-2353.

Bhati, K. K., Aggarwal, S., Sharma, S., Mantri, S., Singh, S. P., Bhalla, S., Kaur, J., Tiwari, S., Roy, J. K., Tuli, R., and Pandey, A. K. 2014. Differential expression of structural genes for the late phase of phytic acid biosynthesis in developing seeds of wheat (*Triticum aestivum* L.). *Plant Sci.* **224**: 74-85.

Biswas, S., Maity, I. B., Chakrabarti, S., and Biswas, B. B. 1978. Purification and characterization of myo-inositol hexaphosphate-adenosine diphosphate phosphotransferase from *Phaseolus aureus*. *Arch. Biochem. Biophys.* **185**(2): 557-566.

Bohn, L., Meyer, A. S., and Rasmussen, S. K. 2008. Phytate: impact on environment and human nutrition. A challenge for molecular breeding. *J. Zhejiang Univ. Sci. B* **9**(3): 165-191.

Brehm, M. A., Wundenberg, T., Williams, J., Mayr, G. W., and Shears, S. B. 2013. A non-catalytic role for inositol 1, 3, 4, 5, 6-pentakisphosphate 2-kinase in the synthesis of ribosomal RNA. *J. Cell Sci.* **126**(2): 437-444.

Brodin, J., Krishnamoorthy, M., Athreya, G., Fischer, W., Hraber, P., Gleasner, C., Green, L., Korber, B., and Leitner, T. 2013. A multiple-alignment based primer design algorithm for genetically highly variable DNA targets. *BMC Bioinforma.* **14**(1): p.255.

Brown, E. C., Heit, M. L., and Ryan, D. E. 1961. Phytic acid: an analytical investigation. *Can. J. Chem.* **39**(6): 1290-1297.

Butt, M. S., Pasha, I., Sultan, M. T., Randhawa, M. A., Saeed, F., and Ahmed, W. 2013. Black pepper and health claims: a comprehensive treatise. *Crit. Reviews Food Sci. Nutr.* **53**(9): 875-886.

Caddick, S. E., Harrison, C. J., Stavridou, I., Mitchell, J. M L., Hemmings, A. M., and Brearley, C. A. 2008. A *Solanum tuberosum* inositol phosphate kinase (StITPK1) displaying inositol phosphate–inositol phosphate and inositol phosphate–ADP phosphotransferase activities. *Fed. Eur. Biochem. Soc. Letters* **582**(12): 1731-1737.

Campion, B., Sparvoli, F., Doria, E., Tagliabue, G., Galasso, I., Fileppi, M., Bollini, R., and Nielsen, E. 2009. Isolation and characterisation of an lpa (low phytic acid) mutant in common bean (*Phaseolus vulgaris* L.). *Theor. Appl. Genet.* **118**(6): 1211-1221.

Chawla, R., and De Mason, D. 2003. Genome walking in pea: an approach to clone unknown flanking sequences. *Pisum Genet.* **35**: 3-5.

Detter, J. C., Jett, J. M., Lucas, S. M., Dalin, E., Arellano, A. R., Wang, M., Nelson, J. R., Chapman, J., Lou, Y., Rokhsar, D., and Hawkins, T. L. 2002. Isothermal strand-displacement amplification applications for high-throughput genomics. *Genomics* **80**(6): 691-698.

Dhole, V. J., and Reddy, K. S. 2016. Association of phytic acid content with biotic stress tolerance in mungbean *(Vigna radiate). Phytoparasitica* **44**(2): 261-267.

Doria, E., Galleschi, L., Calucci, L., Pinzino, C., Pilu, R., Cassani, E., and Nielsen, E. 2009. Phytic acid prevents oxidative stress in seeds: evidence from a maize (*Zea mays L.*) low phytic acid mutant.*J. Exp. Bot.* **60**(3): 967–978.

Dorsch, J. A., Cook, A., Young, K. A., Anderson, J. M., Bauman, A. T., Volkmann, C. J., Murthy, P. P., and Raboy, V. 2003. Seed phosphorus and inositol phosphate phenotype of barley low phytic acid genotypes. *Phytochemistry* **62**(5): 691-706.

Doyle, J. J. and Doyle, J. L. 1987. A rapid procedure for DNA purification from small quantities of fresh leaf tissue. *Phytochemical Bull.* **19:** 11-15.

Fakruddin, M., Mannan, K. S. B., Chowdhury, A., Mazumdar, R. M., Hossain, M. N., Islam, S., and Chowdhury, M. A. 2013. Nucleic acid amplification: alternative methods of polymerase chain reaction. *J. Pharmacy Bioallied Sci.* **5**(4): 245.

Fileppi, M., Galasso, I., Tagliabue, G., Daminati, M. G., Campion, B., Doria, E., and Sparvoli, F. 2010. Characterisation of structural genes involved in phytic acid biosynthesis in common bean (*Phaseolus vulgaris L.*). *Mol. Breed.* **25**(3): 453-470.

Gadkar, V. J., and Filion, M. 2011. A novel method to perform genomic walks using a combination of single strand DNA circularization and rolling circle amplification. *J. Microbiological Methods* **87**(1): 38-43.

Gadkar, V., and Rillig, M. C. 2005. Application of Phi29 DNA polymerase mediated whole genome amplification on single spores of arbuscular mycorrhizal (AM) fungi. *Fed. Eur. Microbiological Soc. Microbiol. Letters* **242**(1): 65-71.

Gary, J. D., Sato, T. K., Stefan, C. J., Bonangelino, C. J., Weisman, L. S., and Emr, S. D. 2002. Regulation of Fab1 phosphatidylinositol 3-phosphate 5-kinase

pathway by Vac7 protein and Fig4, a polyphosphoinositide phosphatase family member. *Mol. Biol. Cell* **13**(4): 1238-1251.

Gonzalez, B., Banos-Sanz, J. I., Villate, M., Brearley, C. A., and Sanz-Aparicio, J. 2010. Inositol 1, 3, 4, 5, 6-pentakisphosphate 2-kinase is a distant IPK member with a singular inositide binding site for axial 2-OH recognition. *Proc. Natl. Acad. Sci.* **107**(21): 9608-9613.

Good, L. L. 2001. Isolation and characterization of d-myo-inositol-3-phosphate synthase gene family members in soybean. M.Sc thesis, Virginia Polytechnic Institute and State University, Blacksburg, Virginia. 133p.

Gorgani, L., Mohammadi, M., Najafpour, G. D., and Nikzad, M. 2017. Piperine—The bioactive compound of black pepper: from isolation to medicinal formulations. *Comprehensive Rev. Food Sci. Food Safety* **16**(1): 124-140.

Graf, E., Mahoney, J. R., Bryant, R. G., and Eaton, J. W. 1984. Iron-catalyzed hydroxyl radical formation. stringent requirement for free iron coordination site. *J. Biol. Chem.* **259**(6): 3620-3624.

Greiner, R., Konietzny, U., and Jany, K. D. 2006. Phytate-an undesirable constituent of plant-based foods?. *J. Ernahrungsmedizin* **8**(3): 18-28.

Guil-Guerrero, J. L., Ramos, L., Moreno, C., Zuniga-Paredes, J. C., Carlosama-Yepez, M., and Ruales, P. 2016. Plant foods by-products as sources of health-promoting agents for animal production: a review focusing on the tropics. *Agron. J.* **108**(5): 1759-1774.

Guo, H. and Xiong, J. 2006. A specific and versatile genome walking technique. *Gene* **381**: 18-23.

Guttieri, M., Bowen, D., Dorsch, J. A., Raboy, V., and Souza, E. 2004. Identification and characterization of a low phytic acid wheat. *Crop Sci.* **44**(2): 418-424.

Han, I. K., Lee, J. H., Piao, X. S., and Li, D. 2001. Feeding and management system to reduce environmental pollution in swine production. *Asian Australasian J. Animal Sci.* **14**(3): 432-444.

Hanakahi, L. A., Bartlet-Jones, M., Chappell, C., Pappin, D., and West, S. C. 2000. Binding of inositol phosphate to DNA-PK and stimulation of double-strand break repair. *Cell* **102**(6): 721-729

Hawkins, P. T., Poyner, D. R., Jackson, T. R., Letcher, A. J., Lander, D. A., and Irvine, R. F. 1993. Inhibition of iron-catalysed hydroxyl radical formation by inositol polyphosphates: a possible physiological function for myo-inositol hexakisphosphate. *Biochemical J.* **294**(3): 929-934.

Hitz, W. D., Carlson, T. J., Kerr, P. S., and Sebastian, S. A. 2002. Biochemical and molecular characterization of a mutation that confers a decreased raffinosaccharide and phytic acid phenotype on soybean seeds. *Plant Physiol.* **128**(2): 650-660.

Huang, X. and Madan, A. 1999. CAP3: A DNA sequence assembly program. *Genome Res.* **9**(9): 868-877.

Huang, C., Ma, W. Y., Hecht, S. S., and Dong, Z. 1997. Inositol hexaphosphate inhibits cell transformation and activator protein 1 activation by targeting phosphatidylinositol-3′ kinase. *Cancer Res.* **57**(14): 2873-2878.

Johne, R., Müller, H., Rector, A., Van Ranst, M., and Stevens, H., 2009. Rolling-circle amplification of viral DNA genomes using phi29 polymerase. *Trends Microbiol.* **17**(5): 205-211.

Kanani and Poonam. 2015. Biotechnological interventions of phytase gene in crop plants. *Readers Shelf* **12**(02): 22-24.

Kuo, H. F., Chang, T. Y., Chiang, S. F., Wang, W. D., Charng, Y. Y., and Chiou, T. J. 2014. Arabidopsis inositol pentakisphosphate 2-kinase, AtIPK1, is required for growth and modulates phosphate homeostasis at the transcriptional level. *Plant J.* **80**(3): 503-515.

Kuwano, M., Mimura, T., Takaiwa, F., and Yoshida, K.T. 2009. Generation of stable 'low phytic acid'transgenic rice through antisense repression of the 1d-myo-inositol 3-phosphate synthase gene (RINO1) using the 18-kDa oleosin promoter. *Plant biotechnol. J.* **7**(1): 96-105.

Larson, S. R., Young, K. A., Cook, A., Blake, T. K., and Raboy, V. 1998. Linkage mapping of two mutations that reduce phytic acid content of barley grain. *Theor. Appl. Genet.* **97**(1): 141-146.

Larson, S. R., Rutger, J. N., Young, K. A., and Raboy, V. 2000. Isolation and genetic mapping of a non-lethal rice (L.) mutation. *Crop Sci.* **40**(5): 1397-1405.

Larsson, O., Barker, C. J., Sjoholm, A., Carlqvist, H., Michell, R. H., Bertorello, A., Nilsson, T., Honkanen, R. E., Mayr, G. W., Zwiller, J., and Berggren, P. O. 1997. Inhibition of phosphatases and increased Ca2+ channel activity by inositol hexakisphosphate. *Science* **278**(5337): 471-474.

Lasken, R. S. and Egholm, M. 2003. Whole genome amplification: abundant supplies of DNA from precious samples or clinical specimens. *Trends Biotechnol.* **21**(12): 531-535.

Lemtiri- Chlieh, F., MacRobbie, E. A., Webb, A. A., Manison, N. F., Brownlee, C., Skepper, J. N., Chen, J., Prestwich, G. D., and Brearley, C. A. 2003. Inositol hexakisphosphate mobilizes an endomembrane store of calcium in guard cells. *Proc. Natl. Acad. Sci.* **100**(17): 10091-10095.

Leoni, C., Gallerani, R., and Ceci, L. R. 2008. A genome walking strategy for the identification of eukaryotic nucleotide sequences adjacent to known regions. *BioTechniques* **44**(2): 229.

Leoni, C., Volpicella, M., De Leo, F., Gallerani, R., and Ceci, L. R., 2011. Genome walking in eukaryotes. *Febs J.* **278**(21): 3953-3977.

Losito, O., Szijgyarto, Z., Resnick, A. C., and Saiardi, A. 2009. Inositol pyrophosphates and their unique metabolic complexity: analysis by gel electrophoresis. *PLoS One* **4**(5) p5580.

Lott, J., N., A. 1984. Accumulation of seed reserves of phosphorus and other minerals. In: Murray D., R (ed.). *Seed Physiology.* Academic Press, New York, 139-166.

Lott, J. N., Ockenden, I., Raboy, V., and Batten, G. D., 2000. Phytic acid and phosphorus in crop seeds and fruits: a global estimate. *Seed Sci. Res.* **10**(01): 11-33.

Lv, Y., Zhang, Z., Hou, L., Zhang, L., Zhang, J., Wang, Y., Liu, C., Xu, P., Liu, L., Gai, X., and Lu, T., 2015. Phytic acid attenuates inflammatory responses and the levels of NF-κB and p-ERK in MPTP-induced Parkinson's disease model of mice. *Neurosci. Letters* **597**: 132-136.

Ma, D., Zuo, Y., Sun, D., Wang, C., and Guo, T., 2013. Characterization of the TaMIPS gene from winter wheat (*Triticum aestivum* L.) and changes in its expression pattern with phytic acid accumulation in seeds during grain filling. *J. Cereal Sci.* **57**(3): 437-443.

Macbeth, M. R., Schubert, H. L., VanDemark, A. P., Lingam, A. T., Hill, C. P., and Bass, B. L. 2005. Inositol hexakisphosphate is bound in the ADAR2 core and required for RNA editing. *Science* **309**(5740): 1534-1539.

Matsuno, K. and Fujimura, T. 2014. Induction of phytic acid synthesis by abscisic acid in suspension-cultured cells of rice. *Plant Sci.* **217**: 152-157.

Morton, R. K. and Raison, J. K. 1963. A complete intracellular unit for incorporation of amino-acid into storage protein utilizing adenosine triphosphate generated from phytate. *Nature* **200:** 429-433.

Munnik, T. and Vermeer, J. E. 2010. Osmotic stress-induced phosphoinositide and inositol phosphate signalling in plants. *Plant, Cell Environ.* **33**(4): 655-669.

Murphy, A. M., Otto, B., Brearley, C. A., Carr, J. P., and Hanke, D. E. 2008. A role for inositol hexakisphosphate in the maintenance of basal resistance to plant pathogens. *Plant J.* **56**(4): 638-652.

Myrick, K. V. and Gelbart, W. M. 2002. Universal fast walking for direct and versatile determination of flanking sequence. *Gene* **284**(1): 125-131.

Navarro, L., Dunoyer, P., Jay, F., Arnold, B., Dharmasiri, N., Estelle, M., Voinnet, O., and Jones, J. D. 2006. A plant miRNA contributes to antibacterial resistance by repressing auxin signaling. *Science* **312**(5772): 436-439.

Odorizzi, G., Babst, M., and Emr, S. D. 1998. Fab1p PtdIns (3) P 5-kinase function essential for protein sorting in the multivesicular body. *Cell* **95**(6): 847-858.

Organ, M. G., Greenwood, J. S., and Bewley, J. D. 1988. Phytin is synthesized in the cotyledons of germinated castor-bean seeds in response to exogenously supplied phosphate. *Planta* **174**(4): 513-517.

Ovcharenko, I., Loots, G. G., Giardine, B. M., Hou, M., Ma, J., Hardison, R. C., Stubbs, L. and Miller, W. 2005. Mulan: multiple-sequence local alignment and visualization for studying function and evolution. *Genome Res.* **15**(1): 184-194.

Paez, J. G., Lin, M., Beroukhim, R., Lee, J. C., Zhao, X., Richter, D. J., Gabriel, S., Herman, P., Sasaki, H., Altshuler, D., and Li, C. 2004. Genome coverage and sequence fidelity of ϕ29 polymerase-based multiple strand displacement whole genome amplification. *Nucleic Acids Res.* **32**(9): 71

Phillippy, B. Q., Ullah, A. H., and Ehrlich, K. C. 1994. Purification and some properties of inositol 1, 3, 4, 5, 6-Pentakisphosphate 2-kinase from immature soybean seeds. *J. Biol. Chem.* **269**(45): 28393-28399.

Phillippy, B. Q., Perera, I. Y., Donahue, J. L., and Gillaspy, G. E. 2015. Certain Malvaceae plants have a unique accumulation of myo-inositol 1, 2, 4, 5, 6-pentakisphosphate. *Plants* **4**(2): 267-283.

Pop, M. 2009. Genome assembly reborn: recent computational challenges. *Briefings Bioinforma.* **10**(4): 354-366.

Raboy, V. 1997. Accumulation and storage of phosphate and minerals. In: *Cellular and Molecular Biology of Plant Seed Development*. Springer Netherlands, 441-477

Raboy, V. 2001. Seeds for a better future:'low phytate'grains help to overcome malnutrition and reduce pollution. *Trends Plant Sci.* **6**(10): 458-462.

Raboy, V., Young, K. A., Dorsch, J. A., and Cook, A. 2001. Genetics and breeding of seed phosphorus and phytic acid. *J. Plant Physiol.* **158**: 489-497.

Raboy, V. 2003. myo-Inositol-1, 2, 3, 4, 5, 6-hexakisphosphate. *Phytochemistry* **64**(6): 1033-1043.

Raghunathan, A., Ferguson, H. R., Bornarth, C. J., Song, W., Driscoll, M., and Lasken, R. S., 2005. Genomic DNA amplification from a single bacterium. *Applied environ. Microbiol.* **71**(6): 3342-3347.

Reddy, N. R., Sathe, S. K., and Salunkhe, D. K. 1982. Phytates in legumes and cereals. *Adv. Food Res.* **28**: 1-92.

Reddy, M. K., Nair, S., and Sopory, S. K. 2002. A new approach for efficient directional genome walking using polymerase chain reaction. *Anal. Biochem.* **306**(1): 154-158.

Reddy, P. S., Mahanty, S., Kaul, T., Nair, S., Sopory, S. K., and Reddy, M. K., 2008. A high-throughput genome-walking method and its use for cloning unknown flanking sequences. *Anal. Biochem.* **381**(2): 248-253.

Rill, D. R., Santana, V. M., Roberts, W. M., Nilson, T., Bowman, L. C., Krance, R. A., Heslop, H. E., Moen, R. C., Ihle, J. N., and Brenner, M. K. 1994. Direct demonstration that autologous bone marrow transplantation for solid tumors can return a multiplicity of tumorigenic cells. *Blood* **84**(2): 380-383.

Rishi, A. S., Nelson, N. D., and Goyal, A. 2004. Genome walking of large fragments: an improved method. *J. Biotechnol.* **111**(1): 9-15.

Roorkiwal, M., Nayak, S. N., Thudi, M., Upadhyaya, H. D., Brunel, D., Mournet, P., This, D., Sharma, P. C., and Varshney, R. K. 2014. Allele diversity for abiotic stress responsive candidate genes in chickpea reference set using gene based SNP markers. *Front. Plant Sci.* 5.

Rusten, T. E., Rodahl, L. M., Pattni, K., Englund, C., Samakovlis, C., Dove, S., Brech, A., and Stenmark, H. 2006. Fab1 phosphatidylinositol 3-phosphate 5-kinase controls trafficking but not silencing of endocytosed receptors. *Mol. Biol. Cell* **17**(9): 3989-4001.

Samuel, M., Oliver, S. V., Coetzee, M., and Brooke, B. D. 2016. The larvicidal effects of black pepper (*Piper nigrum* L.) and piperine against insecticide resistant and susceptible strains of Anopheles malaria vector mosquitoes. *Parasites Vectors* **9**(1): 238.

Sasakawa, N., Sharif, M., and Hanley, M. R. 1995. Metabolism and biological activities of inositol pentakisphosphate and inositol hexakisphosphate.*Biochemical Pharmacol*.**50**(2): 137-146.

Shamsuddin, A. M. and Vucenik, I. 1999. Mammary tumor inhibition by IP6: a review. *Anticancer Res.* **19**(5A): 3671-3674.

Sparvoli, F. and Cominelli, E. 2015. Seed biofortification and phytic acid reduction: a conflict of interest for the plant?. *Plants* **4**(4): 728-755.

Shen, X., Xiao, H., Ranallo, R., Wu, W. H. and Wu, C. 2003. Modulation of ATP-dependent chromatin-remodeling complexes by inositol polyphosphates. *Science* **299**(5603): 112-114.

Shi, J., Wang, H., Wu, Y., Hazebroek, J., Meeley, R. B., and Ertl, D. S. 2003. The maize low-phytic acid mutant lpa2 is caused by mutation in an inositol phosphate kinase gene. *Plant Physiol.* **131**(2): 507-515.

Silva, E. O. and Bracarense, A. P. F. 2016. Phytic acid: from antinutritional to multiple protection factor of organic systems. *J. Food Sci.* **81**(6): 1357-1362.

Sivaraman, K., Kandiannan, K., Peter, K. V. and Thankamani, C. K. 1999. Agronomy of black pepper (*Piper nigrum* L.)-a review. *J. Spices Aromat. Crop* **8**(1): 1-18

Smernik, R. J. and Dougherty, W. J. 2007. Identification of phytate in phosphorus-31 nuclear magnetic resonance spectra: The need for spiking. *Soil Sci. Soc. Am. J.* **71**(3): 1045-1050.

Srinivasan, K. 2007. Black pepper and its pungent principle-piperine: a review of diverse physiological effects. *Crit. Rev. Food Sci. Nutr.* **47**(8): 735-748.

Stevenson-Paulik, J., Bastidas, R. J., Chiou, S. T., Frye, R. A., and York, J.D. 2005. Generation of phytate-free seeds in Arabidopsis through disruption of inositol polyphosphate kinases. *Proc. Natl. Acad. Sci. of U. S. Am.* **102**(35): 12612-12617.

Sujayanont, P., Chininmanu, K., Tassaneetrithep, B., Tangthawornchaikul, N., Malasit, P., and Suriyaphol, P. 2014. Comparison of phi29-based whole genome amplification and whole transcriptome amplification in dengue virus. *J. Virol. Methods.* **195**: 141-147.

Sun, Y., Thompson, M., Lin, G., Butler, H., Gao, Z., Thornburgh, S., Yau, K., Smith, D. A., and Shukla, V. K. 2007. Inositol 1, 3, 4, 5, 6-pentakisphosphate 2-kinase from maize: molecular and biochemical characterization. *Plant Physiol.* **144**(3): 1278-1291.

Suzuki, M., Tanaka, K., Kuwano, M., and Yoshida, K. T. 2007. Expression pattern of inositol phosphate-related enzymes in rice (*Oryza sativa* L.): implications for the phytic acid biosynthetic pathway. *Gene* **405**(1): 55-64.

Sweetman, D., Johnson, S., Caddick, S. E., Hanke, D. E., and Brearley, C. A. 2006. Characterization of an Arabidopsis inositol 1, 3, 4, 5, 6-pentakisphosphate 2-kinase (AtIPK1). *Biochemical J.* **394**(1): 95-103.

Tan, X., Calderon-Villalobos, L. I. A., Sharon, M., Zheng, C., Robinson, C. V., Estelle, M., and Zheng, N. 2007. Mechanism of auxin perception by the TIR1 ubiquitin ligase. *Nature* **446**(7136): 640-645.

Tsaftaris, A., Pasentzis, K., and Argiriou, A. 2010. Rolling circle amplification of genomic templates for inverse PCR (RCA–GIP): a method for 5′-and 3′-genome walking without anchoring. *Biotechnol. Letters* **32**(1): 157.

Valluru, R. and Van den Ende, W. 2011. Myo-inositol and beyond–emerging networks under stress. *Plant Sci.* **181**(4): 387-400.

Vats, P., Bhattacharyya, M. S., and Banerjee, U. C. 2005. Use of phytases (myo-inositolhexakisphosphate phosphohydrolases) for combatting environmental pollution: A biological approach. *Crit. Reviews Environ. Scl. Technol.* **35**(5): 469-486.

Verbsky, J. W., Wilson, M. P., Kisseleva, M. V., Majerus, P. W., and Wente, S. R. 2002. The synthesis of inositol hexakisphosphate Characterization of

137

human inositol 1, 3, 4, 5, 6-pentakisphosphate 2-kinase. *J. Biol. Chem.* **277**(35): 31857-31862.

Verbsky, J. and Majerus, P. W., 2005. Increased levels of inositol hexakisphosphate (InsP6) protect HEK293 cells from tumor necrosis factor α-and Fas-induced apoptosis. *J. Biol. Chem.* ***280***(32): 29263-29268.

Volpicella, M., Leoni, C., Costanza, A., Fanizza, I., Placido, A., and Ceci, L.R. 2012. Genome walking by next generation sequencing approaches. *Biol.* **1**(3): 495-507.

Volpicella, M., Leoni, C., Fanizza, I., Rius, S., Gallerani, R., and Ceci, L. R. 2012. Genome walking by klenow polymerase. *Anal.Biochem.* **430**(2): 200-202.

Vucenik, I., Podczasy, J. J. and Shamsuddin, A. M., 1999. Antiplatelet activity of inositol hexaphosphate (IP6). *Anticancer Res.* **19**(5A): 3689-3693.

Williams, S. P., Gillaspy, G. E., and Perera, I. Y., 2015. Biosynthesis and possible functions of inositol pyrophosphates in plants. *Frontiers Plant Sci.* **6**(67): 1-12.

Xu, J., Brearley, C. A., Lin, W. H., Wang, Y., Ye, R., Mueller-Roeber, B., Xu, Z. H., and Xue, H.W. 2005. A role of Arabidopsis inositol polyphosphate kinase, AtIPK2α, in pollen germination and root growth. *Plant Physiol.* **137**(1): 94-103.

Yamamoto, A., DeWald, D. B., Boronenkov, I. V., Anderson, R. A., Emr, S. D., and Koshland, D. 1995. Novel PI (4) P 5-kinase homologue, Fab1p, essential for normal vacuole function and morphology in yeast. *Mol. Biol. Cell* **6**(5): 525-539.

136

York, J. D., Odom, A. R., Murphy, R., Ives, E. B., and Wente, S. R., 1999. A phospholipase C-dependent inositol polyphosphate kinase pathway required for efficient messenger RNA export. *Science* **285**(5424): 96-100.

Yu, J. 2013.Castor bean *(Ricinus communis L.)* genes involved in phytic acid biosynthetic pathways: expression analysis in response to phosphate and characterization of inositol 1, 3, 4, 5, 6-pentakisphosphate 2-kinase. Doctoral dissertation, University of Guelph, 156p.

Yu, J., Saiardi, A., Greenwood, J. S., and Bewley, J. D. 2014. Molecular and biochemical identification of inositol 1, 3, 4, 5, 6-pentakisphosphate 2-kinase encoding mRNA variants in castor bean (*Ricinus communis* L.) seeds. *Planta*, **239**(5): 965-977.

Yuan, F. J., Zhu, D. H., Tan, Y. Y., Dong, D. K., Fu, X. J., Zhu, S. L., Li, B. Q., and Shu, Q. Y. 2012. Identification and characterization of the soybean IPK1 ortholog of a low phytic acid mutant reveals an exon-excluding splice-site mutation. *Theor. Appl. Genet.* **125**(7): 1413-1423.

Zhan, H., Zhong, Y., Yang, Z., and Xia, H. 2015. Enzyme activities of Arabidopsis inositol polyphosphate kinases AtIPK2α and AtIPK2β are involved in pollen development, pollen tube guidance and embryogenesis. *Plant J.* **82**(5): 758-771.

GENOME WALKING FOR PUTATIVE PHYTIC ACID ($InsP_6$) UNIGENE IN BLACK PEPPER (*Piper nigrum* L.)

by

ANANDUCHANDRA GIRIDHARI
(2015-11-061)

ABSTRACT

Submitted in partial fulfillment of the requirement for the degree of

Master of Science in Agriculture
(PLANT BIOTECHNOLOGY)
Faculty of Agriculture
Kerala Agricultural University

CENTRE FOR PLANT BIOTECHNOLOGY AND MOLECULAR BIOLOGY
COLLEGE OF HORTICULTURE
VELLANIKKARA, THRISSUR – 680 656
KERALA, INDIA
2017

Abstract

ABSTRACT

Black pepper (*Piper nigrum* L.), is an important spice crop contributing a major share of country's foreign exchange. However the genomic information of this crop is limited. Expressed sequence tags (EST) data obtained by next generation sequencing technology in a previous study to understand the genes expressed during berry development in black pepper showed a candidate unigene (*Pnc135*- 995 bp) coding for inositol pentakisphosphate 2-kinase *(ipk1)*. This is an enzyme involved in the biosynthesis of Myo-inositol hexakisphosphate ($InsP_6$) or commonly called as phytic acid. The reports of *ipk1* gene sequences from other crops showed that the size of *ipk1* gene varied widely between crops in the range of 3- 6 kb. Apart from helping in phosphate storage, phytic acid plays diversified role in plants such as a signalling molecule during drought stress, key role in plant defence mechanism, cofactor in auxin mediated gene expression of the auxin receptor, safe binding site for Fe during Fe transport in cytosol etc (Graf *et al.*, 1984; Lemtiri- Chlien *et al.*, 2003; Murphy *et al.*, 2008).

Being a commercially important crop and considering the important biological roles played by phytic acid in the plant systems, a study on 'Genome walking for putative phytic acid ($InsP_6$) unigene in black pepper (*Piper nigrum* L.)' was taken up to sequence the flanking regions of unigene (*Pnc135*) and to detect the presence of $InsP_6$ using polyacrylamide gel electrophoresis (PAGE) in black pepper.

Genomic DNA was isolated from var.Panniyur1 and used for the PCR amplification of the unigene sequence *Pnc135*. As a preliminary step of the study, two sets of primers (PNC F1/PNC R1, PNC F2/PNC R2) were designed based on the unigene sequence for the confirmation of sequence information and two primer sets (PNC F3/PNC R3, PNC F4/PNC R4a, PNCR4b) were designed based on the pairwise alignment of selected sequences from BLAST search using unigene for the amplification of flanking regions of unigene. Three sets of primers (PNC

142

F5/PNCR5, PNC F6/PNC R6, PNC F7/PNC R7) were designed based on the multiple sequence alignment of unigene with the selected sequences of 28 different crops and PCR amplification was carried out with the different primer combinations. Among the six PCR products sequenced (F1R1, F3R3, F4R4a, F6R5, F7R6 and F7R7), the product F6R5 obtained by primer set PNC F6/ PNC R5 designed based on the consensus regions of *Pnc135* with *ipk1* gene sequence reported in 28 other crops using multiple sequence alignment gave an amplicon of 394 bp. This PCR product (F6R5) showing similarity to *ipk1* gene was used for performing genome walking for the flanking regions.

Genome walking is a method to sequence the flanking regions of a known sequence. The genome walking protocol by Reddy *et al.* (2008) which make use of the strand displacement amplification of phi29 DNA poiymerase was used in the present study. The walker adapter primers (WA1, WA2, WA3, WA4) and the walker primer (WP1) and nested walker primer (WP2) used in the study were same as described in the protocol by Reddy *et al.* (2008). The forward and reverse locus specific primers (LSP F1, LSP F2, LSP R1, LSP R2) and the corresponding nested primers (NLSP F1, NLSP F2, NLSP R1, NLSP R2) were designed based on the sequence of F6R5 for genome walking. From the nested PCR amplification, four PCR products A3F1 and A2F1 (flanking regions obtained towards the 3'end of F6R5) and A4R1 and A4R2 (flanking regions obtained towards the 5'end of F6R5) were sequenced and homology search was done for similarity with *ipk1* gene. Among these, only one PCR product *viz.*, A3F1 of 933 bp showed similarity to *ipk1*. Based on overlapping regions, this was assembled with the F6R5 sequence (394bp) to get a total length of 1072 bp (*Pnipk1*).

The assembled sequence (*Pnipk1*) was analysed for the coding region (Open Reading Frame) and was found to have 522 bp coding for 173 amino acids. The nucleotide sequence and amino acid sequence were used for the phylogenetic analysis using *ipk1* sequences of other crops showing similarity in BLAST search.

The analysis showed that *Pnipk1* was evolutionarily closer to sequences from plants like *Phaelenopsis equistris, Musa acuminata,* and *Phoenix dactylifera.*

Polyacrylamide gel electrophoresis (PAGE) analysis was carried out to detect the presence of $InsP_6$ in Panniyur 1 leaf sample. The leaf extract as well as phytate standards were loaded onto gels and the corresponding band intensities of each concentration of phytate standard and samples were recorded using the software GelQuant.NET. The quantity of phytate in the sample was calculated from the standard curve drawn by plotting band intensities of phytate standards on the y-axis against quantity of phytate standards (nmoles) on x-axis. Quantity of phytate in the sample was estimated to be 620 nmoles/g fresh weight in the leaf tissues.

The study has given the confirmation of the presence of *ipk1* gene in black pepper and was able to decipher a total sequence length of 1072 bp by genome walking towards the 3'end of the 394 bp amplicon obtained from var. Panniyur1. Identification of the sequence towards the 5'end to get the full length gene will help to understand the role played by the gene in biotic and abiotic stress resistance in black pepper.

174235

KERALA AGRICULTURAL UNIVERSITY
CENTRAL LIBRARY
K.A.U. P.O. THRISSUR

144