Genome-wide identification of Auxin Response Factor (ARF) genes and expression analysis in *Capsicum* fruits

M.sc Thesis

2018

Submitted to

CENTRAL DEPARTMENT OF BIOTECHNOLOGY

Tribhuvan University

Kirtipur, Kathmandu, Nepal

A

Dissertation

For partial fulfillment of the requirement for the

Master of Science in Biotechnology

By

Bidyakala Sunuwar

Roll No: BT 301/072

T.U. Regd. No: 5-2-0038-0139-2011

GENOME-WIDE IDENTIFICATION OF AUXIN RESPONSE FACTOR (ARF) GENES AND EXPRESSION ANALYSIS IN *Capsicum* FRUITS

M.Sc. Thesis

2018

Submitted to

Central Department of Biotechnology

Tribhuvan University

Kirtipur, Kathmandu, Nepal

Submitted by

Bidyakala Sunuwar

Supervisor

Prof. Dr. Rajani Malla

Central Department of Biotechnology

Tribhuvan University

Kirtipur, Kathmandu, Nepal

Tel. No.: +977 1 4336221

Date………………………

Recommendation

This is to certify that the research work entitled **"Genome-wide identification of Auxin Response Factor (ARF) genes and expression analysis in *Capsicum* fruits"** has been carried out by **Ms. Bidyakala Sunuwar** under my supervision.

This thesis work was performed for the partial fulfillment of the Master of Science in Biotechnology under the course code BT 621. The result presented here is her original findings. I, hereby, recommend this thesis for final evaluation.

Prof. Dr. Rajani Malla
(Supervisor)
Former- Head of Department
Central Department of Biotechnology
Tribhuvan University
Kirtipur, Kathmandu, Nepal

School of Life Sciences
Jawaharlal Nehru University
New Delhi 110067, India

Nirala Ramchiary, Ph.D.
Assistant Professor
Email : nramchiary@mail.jnu.ac.in

Tel. : 91-11-26704531
Fax : 91-11-26742558
nrudsc@gmail.com

Date: 04-06-2018

Certificate

This is to certify that this dissertation entitled **"Genome-wide identification of Auxin Response Factor (ARF) genes and expression analysis in *Capsicum* fruits"** is a bona fide work of **Ms. Bidyakala Sunuwar**, towards completion of her dissertation in the partial fulfilment for the requirement of the degree of Master of Science in Biotechnology.

This research work is original and has not been submitted so far, in part or in full, for any degree or diploma in this or any other university or institute.

Bidyakala Sunuwar
(Candidate)

Dr. Nirala Ramchiary
(Supervisor)

Central Department of Biotechnology

Tribhuvan University

Kirtipur, Kathmandu, Nepal

Tel. No.: +977 1 4336221

Date………………………

Certificate of Evaluation

This is to certify that this thesis entitled **"Genome-wide identification of Auxin Response Factor (ARF) genes and expression analysis in *Capsicum* fruits"** presented to evaluation committee by Ms. **Bidyakala Sunuwar** is found satisfactory for the partial fulfillment of Master of Science in Biotechnology.

……………………………………

Dr. Krishna Das Manandhar

(Head of Department)

Central Department of Biotechnology
Tribhuvan University
Kirtipur, Kathmandu, Nepal

……………………………………

(External examiner)

……………………………………

Dr. Rajani Malla

(Supervisor)

Former Head of Department

Central Department of Biotechnology
Tribhuvan University
Kirtipur, Kathmandu, Nepal

……………………………………

(Internal Supervisor)

Central Department of Biotechnology
Tribhuvan University
Kirtipur, Kathmandu, Nepal

ACKNOWLEDGEMENT

I would first like to pay my sincere gratitude to my supervisor ***Dr. Rajani Malla****, Professor, Central Department of Biotechnology, Tribhuvan University for providing me continuous moral support and encouragement throughout my thesis as well as during course time. I am in-debt for her precious suggestions and recommendations to execute this important work in Jawaharlal Nehru University.*

Likewise, I would like to express my sincere regards to **Dr. Nirala Ramchiary***, Assistant Professor, Jawaharlal Nehru University for providing me a golden opportunity to conduct my thesis in his lab. He provided me all mandatory trainings and guidance require in order to improve my research skill. I would also like to thank him a lot for being so careful and cheerful all the time with me. Along with him, I would like to remember all the lab members* of ***Laboratory of Translational and Evolutionary Genomics*** *for their helpful hands.*

I gratefully acknowledge **Dr. Krishna Das Manandhar***, Head of Department, Central Department of Biotechnology,* **Tribhuvan University** *for granting me permission to carry out my research in JNU and for encouraging me with continual support.*

My sincere vote of thanks goes to all the ***faculty, friends, seniors, juniors*** *and* ***staffs*** *of CDBT who have been very supportive during my master's study. I want to express my deepest gratitude to all the people who helped me directly or indirectly throughout my graduation journey.*

Last but not the least, I am grateful towards the almighty **God** *and* ***my family****: my parents, my siblings for their immense love, support and blessing showered throughout this period and my life in general.*

Bidyakala Sunuwar

ACRONYMS

ARF	-	Auxin Response Factor
cDNA	-	complementary DNA
cm	-	centimeter
Ct	-	Cycle threshold
DDW	-	Double distilled water
DEPC	-	Diethyl pyrocarbonate
DNA	-	Deoxyribonucleic Acid
dNTPs	-	dinucleotide Phosphates
EDTA	-	Ethylene diamine tetra acetic acid
EtBr	-	Ethiduim bromide
h	-	hour
kb	-	kilobase
M	-	Molar
mL	-	mili liter
Mm	-	mili molar
MOPS	-	3-(N-Morpholino)propanesulfonic acid
MQ	-	Milli Q
mRNA	-	messenger RNA
qRT-PCR	-	Quantitative Real Time Polymerase Chain Reaction
RT-PCR	-	Reverse Transcriptase Polymerase Chain Reaction
RNase	-	Ribonuclease
SSR	-	Simple Sequence Repeats

Taq - *Thermus aquaticus*

TBE - Tris-Borate-EDTA

TF - Transcription Factor

PCR - Polymerase Chain Reaction

pmol - pico-mole

rpm - Revolutions per minute

RNA - Ribonucleic Acid

RT - Room Temperature

˚C - Degree Celsius

% - Percentage

µL - micro-liter

µm - micro-meter

µM - micro-molar

ng - nano-gram

µg - micro-gram

CONTENTS

List of Figures

Figure 31: Semi-quantitative PCR products on 1.2% agarose gel visualized under UV captured by gel doc system

Figure 32: Semi-quantitative PCR products on 1.2% agarose gel visualized under UV captured by gel doc system

Figure 33: Semi-quantitative PCR products on 1.2% agarose gel visualized under UV captured by gel doc system

Figure 34: Semi-quantitative PCR products on 1.2% agarose gel visualized under UV captured by gel doc system

Figure 35: Semi-quantitative PCR products on 1.2% agarose gel visualized under UV captured by gel doc system

Figure 36: Semi-quantitative PCR products on 1.2% agarose gel visualized under UV captured by gel doc system

Figure 37: Semi-quantitative PCR products on 1.2% agarose gel visualized under UV captured by gel doc system

Figure 38: qRT-PCR analyses for relative expression of ARF genes in *C. annuum* and *C. chinense*.

List of Tables

ABSTRACT

One of the important plant transcription factors is Auxin Response Factor (ARF), which has vital role in regulation of growth and development of the plants, and stress tolerance. ARFs key regulator is the phyto-hormone Auxin. Very less investigation has been done about ARF family of pepper (*Capsicum annuum*) and their roles in biological processes. From the available updated genome database, 28 ARF genes were identified in pepper genome and 24 of them are distributed on 12 different chromosome and 4 are found to be pseudogenes. A phylogenetic tree was constructed depending on ARFs sequences derived from Pepper, Tomato, Arabidopsis, Potato and Rice and Simple Sequence Repeats (SSRs) were identified present in Capsicum ARF. In this study, expression of selected 12 ARF genes in fruits (developmental stage: early, breaker and mature) of two distinct *Capsicum* species (*Capsicum annuum* and *Capsicum chinense*) were analyzed in transcriptome level. Confirmation of expression analysis was done by quantitative real time PCR (qRT-PCR). The preliminary results of our experiment elucidates, ARF plays significant role in fruit development of pepper. This result will provide base for identifying proper function and molecular breeding studies of ARF genes in pepper in forthcoming days.

Keywords: Transcription factor, plant hormone, transcriptome, quantitative real time PCR (qRT-PCR), molecular breeding

CHAPTER 1
INTRODUCTION

1.1 Background

Genus *Capsicum*, commonly called as chili pepper, bell pepper, paprika or just *Capsicum* is a vegetable. *Capsicum* is indigenous genus of South and Central America. This crop is most important members of the family Solanaceae after tomato and potato, consisting more than 38 species. Among them only six species are cultivated namely *Capsicum annuum, Capsicum chinense, Capsicum frutescence, Capsicum pubescens, Capsicum baccatum and Capsicum assamicum* in different parts of world (Ramchiary *et al*. 2014). Chilies are used as vegetables, spices, flavorants, colorants and condiments all over the world in fresh or dried forms. Wide variations of intra as well interspecies are seen for fruit morphology, plant types and other agronomic traits in *Capsicum* species. The fruits of *Capsicum*s are consumed in any three stages namely green, breaker and mature freshly and in dried red form also. One of the unique and important properties of *Capsicum* species is its pungency due to Capsaicin apart from other Solanaceae members. Pungency of *Capsicum* is limited to its fruit. Hence, detailed genetic study of fruit development in *Capsicum* is necessary for improving its economic value.

Chilies fruit are consumed all over the world as vegetables and spices. These are one of the oldest vegetable known to be originating in the New World. An archeological study of pepper starch fossils shows that pepper domestication was done 6,000 years ago in southern Mexico (Perry *et al*., 2007). *Capsicum* species flowers are mostly stellate to rotate corollas which shows different patterns of pigmentation, and fleshy, mostly globose berries of different sizes and colors (Carrizo García *et al*., 2016).

Due to climate change pepper production has been dropped. As it is susceptible to various biotic factors (bacteria, aphids and viruses) and abiotic stresses (high temperature, salinity and drought) its production is decreasing (Diao *et al*., 2016). Many modern agricultural technologies have been introduced combined with the conventional method to improve the *Capsicum* regeneration and its quality. But its further improvement is limited due to its restricted gene pool. Because of its recalcitrance nature, *in vitro* studies have been always a problem (Hegde, Partap, and Yadav, 2017).

1.2 Current studies

C. annuum and *C. chinense* are two different *Capsicum* species with distinct developmental pattern, morphology and pungency contents. Most of the *C. annuum* fruits are found to be elongated varying the size from 12.5 mm to 0.3 m in length (The Editors of Encyclopaedia Britannica, 2018). *C. chinense* fruits are bigger in size ranging its size from 6-150 mm long, with smooth or wrinkled surface (Bagchi and Srivastava, 2003). However, till date only few genes governing fruit size/shape have been reported. Hence, the fruit development study in the *Capsicum* is very important to identify genes which could be later used for genetic improvement program in *Capsicum.*

The phenotypic characteristic of living being is characterized by the expression of the genes encoded by their genomic DNA (gDNA). The sequential molecular reactions occur inside the cell with the help of many transcription factors for proper expression of a particular gene at specific conditions. Auxin has critical role in fruit development starting from formation of flower to patterning of gynoecium through setting of fruit, growth and ripening (Pattison, Csukasi, and Catalá, 2014). The presence of auxin also determines the final size of the fruit, it is reviewed that auxin regulates the fruit maturation (Pattison *et al*., 2014).

However, available literature lacks report on involvement of ARF genes in fruit development. Hence, this study has been carried out to understand the fruit morphology differences because of Auxin Response factor genes (ARF). The detailed study of ARF genes would also help in deciphering the inter-connection of Auxin hormone with other hormones and its roles in stress tolerance in combination with other transcription factors in future days.

1.3 Hypothesis

Null Hypothesis H_0: The expression pattern reveals the role of ARF gene in fruit development of *Capsicum.*

Alternative Hypothesis H_1: The expression pattern only doesn't reveal the role of ARF gene in fruit development of *Capsicum.*

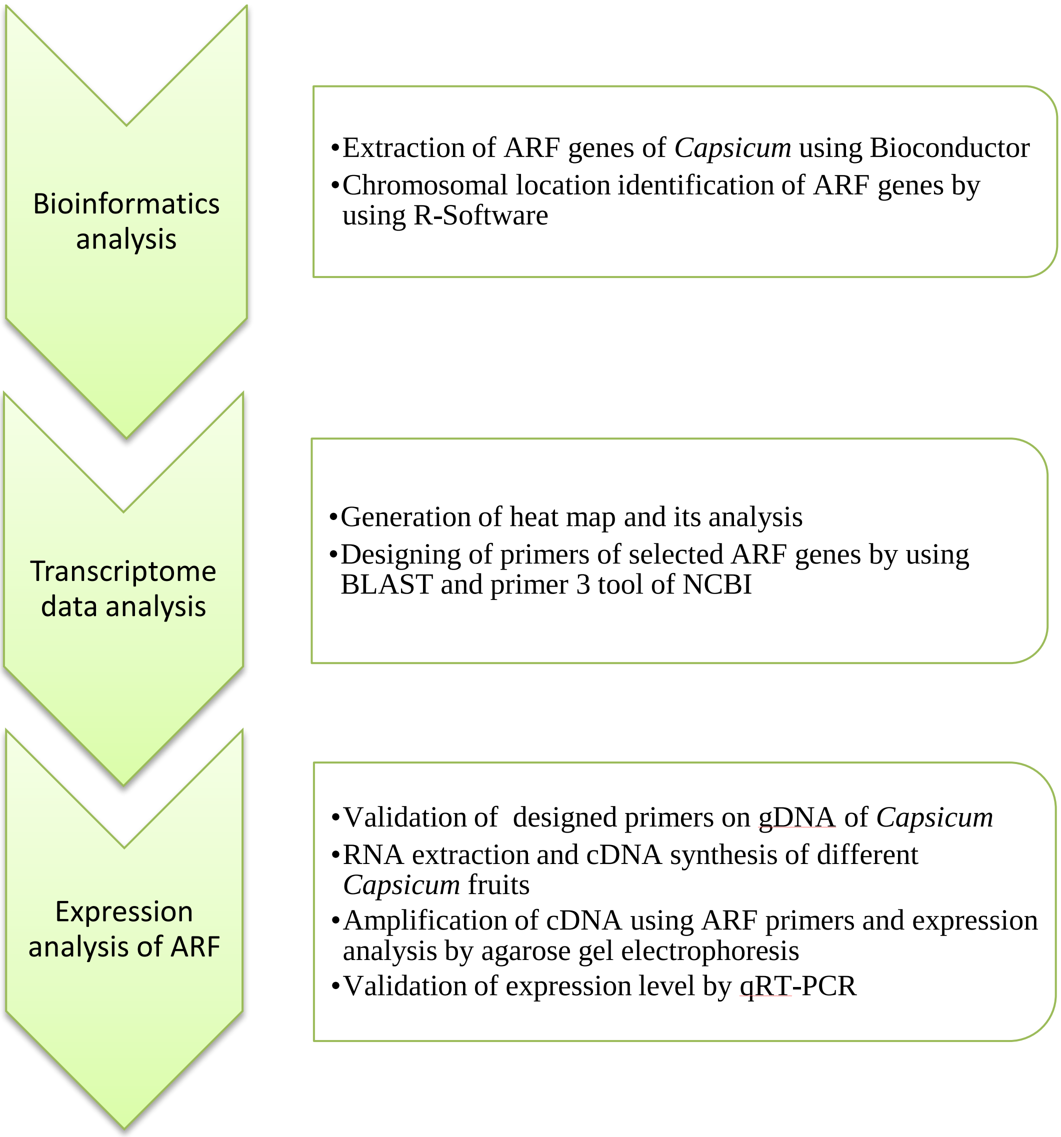
1.4 Objectives

Therefore, the present study is designed with the objective of Genome-wide identification of Auxin Response Factor Genes (ARF) and their expression analysis in *Capsicum* fruits to identify their role in fruit size/shapes. The work is stepwise carried out under the following specific objectives.

1. Genome-wide identification of Auxin Response factor gene in *Capsicum* genome and their chromosomal distribution.
2. Analysis of phylogenetic relationship in between ARF genes from *Capsicum annuum, Arabidopsis thaliana, Solanum tuberosum, Solanum lycopersicum* and *Oryza sativa.*
3. Expression analysis of ARF genes in different fruit developmental stages of *Capsicum annuum, Capsicum chinense* and *Capsicum frutescens* using already existing transcriptome data.
4. Validation of gene expression of ARF genes by semi-quantitative PCR followed by qRT-PCR.

1.5 Research plan

In order to meet the objectives mentioned above, work-flow given below in figure 1 has been designed.



Bioinformatics analysis

- Extraction of ARF genes of *Capsicum* using Bioconductor
- Chromosomal location identification of ARF genes by using R-Software

Transcriptome data analysis

- Generation of heat map and its analysis
- Designing of primers of selected ARF genes by using BLAST and primer 3 tool of NCBI

Expression analysis of ARF

- Validation of designed primers on gDNA of *Capsicum*
- RNA extraction and cDNA synthesis of different *Capsicum* fruits
- Amplification of cDNA using ARF primers and expression analysis by agarose gel electrophoresis
- Validation of expression level by qRT-PCR

Figure 1 : Research plan outline

1.6 Rationale of the study

The structures and functions of every living being are controlled solely by exon genes even after being encountered with adverse environment condition. From the production of important metabolites to its phenotypic characters is determined by the genetic codes. The expression of these genes is regulated by different transcription factors. Hence, transcription factors and their interaction with others chemical factors, environmental factors changes the expression pattern of the genes. *Capsicum* vegetable is economically and medically important. There are various species of *Capsicum* each possessing different characteristic property. Hence, the study of different transcription factors will help to identify the expression pattern of certain genes and their interaction with others also. This study preliminarily results in determining the level of expression of different ARF gene in different *Capsicum* fruit species through RNA expression analysis.

CHAPTER 2
LITERATURE REVIEW

2.1 Importance of *Capsicum*

Capsicum is considered as one of the important crop, which can be used as vegetable as well as spices. They are found to be highly nutritious and rich in phytochemicals.

2.1.1 Nutritional and medicinal value of *Capsicum*

Capsicum fruits are rich in nutrients which contain disease preventing and health promoting benefits. The principle component found in *Capsicum* is Capsaicin, an alkaloid which is a secondary metabolite. Beside capsaicin, *Capsicum* is rich in vitamin C, vitamin E, pro-vitamin A i.e. carotenoids, flavonoids, essential oils and essentials minerals like iron, copper, zinc, manganese, potassium, selenium and magnesium (FAOSTAT-2011). Vitamin C, has major function in maintaining collagen, and non-pungent *Capsicum* is an important source of Vitamin C. It also reduces risk of cardiovascular disease and cancers. It also lowers the production of reactive oxygen species (ROS) if consumed on daily basis recommended by US-FDA (Perla *et al*., 2016).

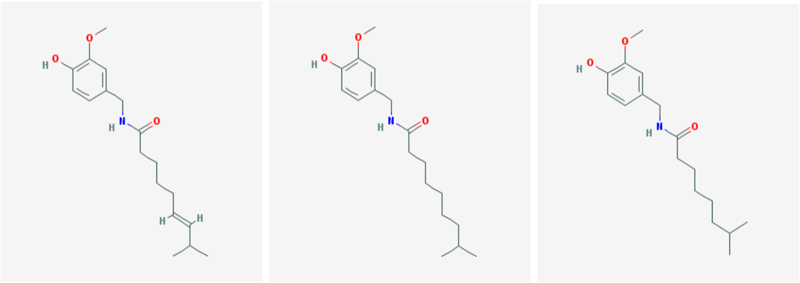
In ethno-medicinal science, *Capsicum* is regarded as prestigious source. *Capsicum* are rich in secondary metabolites, their extracts are used for treatment of various diseases. It has anti-inflammatory, anti-cancerous, anti-microbial, anti-fungal, anti-obesity, anti-aging, anti– mutagenic, anti-oxidant properties. It is used in treatment of diseases like arthritis, cancer. The secondary metabolites found in *Capsicum* are bioactive compounds which help in alleviation of such diseases. Capsaicin found in *Capsicum* acts as an analgesic compound and hence used as topical agent in treatment of post-operative neuralgia, osteoarthritis and painful diabetic neuropath(Anand and Bley, 2011; Park, Yoon, Lee, and Choi, 2016). Along with its benefits, it is irritant and toxic to many living organisms along with plants and animals as well as humans. Hence, it should be used in very small quantity.

Not only this, *Capsicum* extracts are able to show inhibitory effects against these bacteria which commonly causes food poisoning and present in environment namely

Bacillus cereus, Bacillus subtilis, Clostridium sporogenes, Clostridium tetani, and *Streptococcus pyogenes* (Cichewicz and Thorpe, 1996).

2.1.2 Capsaicin in *Capsicum*

Pungency, unique characteristic of the genus *Capsicum* is due to chemical group called Capsaicinoids, which is an alkaloid. Different studies shows that capsaicinoids accumulation in chili is increased as it gets matured from its green stage (Manikharda *et al*., 2018). So, older fruits are more pungent. Primary capsaicinoids present in *Capsicum* are Capsaicin, dihydrocapsaicin and nordihydrocapsaicin (Figure 2) which are produced only inside the placenta of the *Capsicum* fruit (Kim *et al*., 2014). But non placental expression in pericarp tissues was also reported and demonstrated to have accessorial capsaicinoids secreting vesicles in pericarp of some “super-hot” chilies (Bosland, Coon, and Cooke, 2015). These are the major compounds responsible for bringing hot sensation after consuming the chilies. Hotness of chilies is measured in Schoville Heat Unit (SHU) and High Performance Liquid Chromatography (HPLC) technique is used for its measurement. Its value ranges from less than hundred to more than 3 million SHU in different *Capsicum* species. Sweet bell peppers have lower SHU, regarded as non-pungent and used mostly as vegetable whereas hot peppers possess higher SHU, mostly used as spices and in pharmaceuticals.



Capsaicin Dihydrocapsaicin Nordihydrocapsaicin

Figure 2 : Chemical structures of different capsaicinoids found in Capsicum (adapted from PubChem)

Another group of secondary metabolites present in *Capsicum* is capsinoids. These are majorly present in non-pungent species of *Capsicum* and do not cause sensation of burning. Capsinoids found in non-pungent are capsiate, dihydrocapsiate and nordihydrocapsiate, and they also have anti-oxidative property, anti-inflammatory property, inhibit angiogenesis and shown to induce apoptosis (Aza-González, Núñez-Palenius, and Ochoa-Alejo, 2011).

Various physiological effects of Capsaicin are listed below,

1. Beneficial effects on lipid homeostasis
    - Hypocholesteromic influence
    - Hypolipidemic influence
    - Anti-lithogenic influence
    - Protective effect on RBC integrity
2. Beneficial influence on Gastro-intestinal system
    - Beneficial modulation of small intestinal ultra-structure
    - Digestive stimulant action
    - Enhanced absorption of micronutrients
3. Antioxidant effect
4. Anti-inflammatory effect
5. Anti-diabetic effect
6. Cancer preventive potential
7. Thermogenic effect
8. Pain relief
9. Anti-ulcer activity etc. (Srinivasan, 2016).

2.1.3 Capsaicin biosynthesis pathway

Capsaicin being major and unique secondary metabolite of *Capsicum species*, makes *Capsicum* an important plant with high medicinal value and important ingredient of world cuisine. Capsaicinoid is synthesized in placental epidermal cells and finally it gets secreted towards outer cell wall and accumulates in structures named “blisters”. Most of the *Capsicum* species starts to synthesize capsaicinoids approximately after 20 days post-anthesis (DPA) (Aza-González *et al*., 2011). Though many studies have been done till the date but synthesis of capsaicinoids is not fully understood in molecular levels due to presence of limited material and dynamics of capsaicin synthesis which complicate gene discovery and genetics (Zhang *et al*., 2016). It is reported that Capsaicinoids is biosynthesized by the combination of two biosynthetic pathways, namely Phenylpropanoid pathway and Branched fatty acid pathway.

2.1.4 *Capsicum* production

According to data published in year 2016, 4 million tons of *Capsicum* is produced and distributed all over the world. Among them, India is the largest producer of pepper in the world. The data in figure 3 represents that the *Capsicum* is abundantly cultivated all over the world.

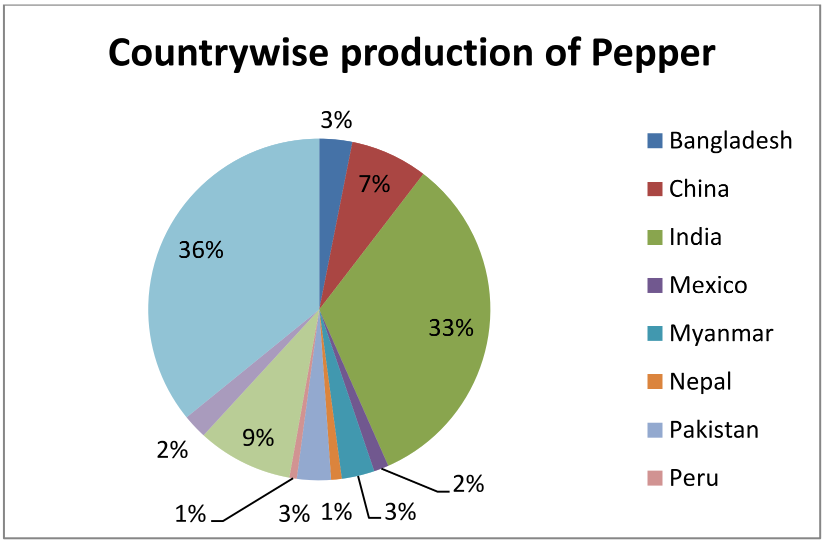


Figure 3 : Capsicum production according to FAO-STAT, 2013

2.2 Transcription Factors

Genetic information of all living organisms is stored in either Deoxyribonucleic Acid (DNA) or Ribonucleic Acid (RNA) in case of RNA viruses (Dragon *et al*., 2002). This stored information is expressed in the series of mechanisms and finally transform into functional proteins, which has diversified functions in the living cell. The steps involved are Replication, Transcription and Translation. This sequential information transfer from DNA/RNA to protein is known as Central dogma of molecular biology of life, which omitted the reverse transfer of information from protein to nucleic acid and also from protein to protein (Crick, 1970).

Transcription is initial and critical stage for expression of gene. As a result, sense strand of DNA ultimately provides primary RNA transcript of specific gene. For the expression of gene, numbers of complex processes are proceeded after formation of RNA transcript namely RNA splicing and translation which results a functional protein. The transcription process is regulated by special proteins, named Transcription Factors (TFs). TFs bind to specific regulatory DNA regions for activation of transcription (Latchman, 1993). TFs are characteristic proteins having extraordinary abilities and qualities. These class specific of proteins' functions are interlinked to modulate different pathways of regulation. TFs are specific to DNA sequence that it gets binds to target gene and achieve specificity in function by communicating with other TFs also (Yusuf *et al*., 2012). TFs can either act as promoting factor or a repressing factor during complete expression of genes. These factors solely decide the activity of the RNA polymerase enzyme during the Transcription and are regarded essential elements for initiating gene transcription. For the natural development of any living organism TFs plays major roles, aids in determining designation of the individual cells (Cooper, 2013). There are many numbers of TFs present in eukaryotes which are assorted into different transcription factor families. The architecture of DNA binding domain of each DNA binding proteins that is TFs is the basis of their assortment (Gonzalez, 2016).

A DNA binding domain, an oligomerization site, a nuclear localization signal and a transcription regulation area are the regions in veritable TFs of most plants excluding fewer plants and this has been extrapolated by comparing the amino acid sequences derived from cDNA from plants with their animal counterparts. The DNA binding domain of TFs comprises specific residues of amino acid which binds to bases of DNA in cis-acting elements. These ascertain the protein specificity. Oligomerization of TFs results in the higher specificity and affinity for promoter elements and nuclear localization and their sequences are also majorly conserved. Nuclear localization signal of plant TFs differs in number, sequence and organization and helps in translocating the proteins in the nucleus. Regulation of transcription process is controlled by transcription regulation domain which either activates or represses gene transcription. The possible mechanism for preventing gene expression are prohibiting activators to binds to target promoters, TFs interaction with repression proteins and dimerizing the TFs to block regulation domain (Liu, White, and Macrae, 1999; Shiu, Shih, and Li, 2005).

2.2.1 Transcription Factors in Plants

Roles of different plants TFs are identified in various biological processes, from development of the plants to defense mechanism against different pathogens (Seo, Choi, and Choi, 2015). In plants, organ development process is perpetual throughout its life cycle. Being sessile organisms, plants have to confront various environmental conditions in order to survive as compared to animals. Thus, their TFs are precisely regulated through different signaling cascades which control development and defense mechanism at molecular levels. According to different studies, 6-10% of the whole plant genome codes for TFs of plants (Franco-Zorrilla *et al.*, 2014).

AP2, ARF,ARR-B, B3, BBR-BPC, BES1, C2H2, C3H, CAMTA, CO-like, CPP, DBB, Dof, E2F/DP, EIL, ERF, FAR1, G2-GATA, GRAS, GRF, GeBP, HB-PHD, HB-other, HD-ZIP, HRT-like, HSF, LBD, LFY, LSD, M-type MADS, MIKC MADS ,MYB, MYB related, NAC, NF-X1, NF-YA, NF-YB, NF-YC, NZZ/SPL, Nin-like, RAV, S1Fa-like, SAP, SBP, SRS, STAT, TALE,TCP, Trihelix, VOZ, WOX, WRKY, Whirly, YABBY, ZF-HD, bHLH and bZIP are different TFs families which are enlisted in plant TF database. Among these TFs, bHLH consists of 28698 i.e. highest number of TFs and lowest one is for SAP i.e. 164 genes (PlantTFDB - Plant Transcription Factor Database @ CBI, PKU, 2017).

The complex process regulating cascades of genes is governed by TFs which results in changes to physiology, morphology and metabolism of the plants. AP2/EREB i.e. APETALA2/Ethylene Response Element Binding protein, WRKY that contains highly conserved WRKYGQK amino-acid sequence, ARF i.e. Auxin Response Factor, Aux/IAA i.e. Auxin/Indole-3-Acetic Acid, Dof represents DNA binding with one finger and NAC i.e. no apical meristem ATAF cup- shaped cotyledon are TFs that are only identified in plants and are responsible for differentiation of tissues, development of root, shoot, flower, fruit, leaves and also provides resistance against different diseases(fungal and bacterial) and adverse environmental conditions (drought, cold, high salinity, and mechanical wounding) (Montiel, 2004). NAC is the largest and widely studied TFs in plants as they imparts in essential physiological processes like cell wall biosynthesis, grain nutrient remobilization, synthesis of flavonoids. Apart from these diversified function NAC also participates in different hormone signaling pathways of auxin, cytokinin, ethylene, abscisic acid, methyl jasmonate, salicylic acid, gibberlic acid etc. NOR is domain of TF NAC, which displays non-ripening characteristics in normal condition, if mutated then changes its gene expression pattern resulting ethylene synthesis, accumulation of carotenoid, softening of the fruits, increased respiration, volatile aroma production etc. (Shan *et al*., 2012).

Similarly, another most important family is WRKY, which is also responsible for protecting plants against biotic and abiotic stresses. 61WRKY genes were known in *Capsicum annuum* genome according to study, it has roles in ripening of fruits, growth and development of the plants. ARF, AP2/ERF, NAC and WRKY TFs are reported to perform specific defense function in Solanaceae family (Cheng *et al*., 2016; Kim *et al*., 2014). Other important TF of plant is YABBY which are responsible for fruit primordia development especially in carpel and nectar growth.

Similarly, MADS box TFs which derives from the initial letter of the genes MCM1, AGAMOUS, DEFICIENS and SRF is involved in governing flower organ identification according to ABC model. In *Arabidopsis thaliana,* AGAMOUS and SHATTERPROOF 1 and SHATTERPROOF 2 are responsible for carpel identification. Whereas for fruit ripening regulation the TFs CNR (Colorless Non Ripening), SBP (SQUAMOSA promoter binding protein) and MADS BOX protein RIN (Ripening Inhibitor) are responsible. Mutation in these genes inhibit the ripening of fruits (Karlova *et al*., 2014).

bZIP, abbreviates for basic leucine zipper TFs family which are highly responsible for overcoming various abiotic stresses in the plants. bZIP comprises of conserved bZIP domain containing a highly basic nuclear localization and a region for DNA binding at N-terminus. At C-terminus a motif rich in leucine is present for dimerization. Some reported genes from bZIP TF family responsible for drought tolerance are *OsbZIP23, GmbZIP1, ZmbZIP17, OsAREB1, OsbZIP16, DgZFP* and *HvDRF1.* Salinity tolerance is also offered by *OsbZIP16, OsbZIP23, GmbZIP44, GmbZIP62.* Heat and cold stress resistance responsible bZIP genes are *OsAREB1, ZmbZIP17* and *DgZFP, GmbZIP132, GmbZIP78* respectively (Joshi *et al*., 2016).

2.3 Auxin Response Factor (ARF)

Auxin is an important phyto-hormone which plays vital roles in plant growth and engineers many developmental processes. Understanding regulation of auxin and its interactions mechanisms can be fruitful for deciphering numerous aspects of plant growth and development. Auxin helps in determining root and shoot structure, organ patterning, vascular structure development, growth in tissue culture. Auxin in combination other hormones modulates various processes. In plant tissue culture, in vitro propagation of root and shoot development is induced by auxin in combination with cytokinin. In in-vivo, the relation of these two hormones is inversed as auxin treatment inhibits the synthesis of cytokinin (WOODWARD AND BARTEL, 2005).

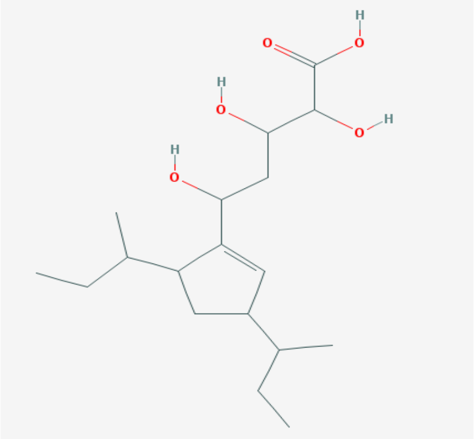


Figure 4 : Chemical structure of Auxin (extracted from https://pubchem.ncbi.nlm.nih.gov)

The most commonly found form of auxin is the Indole-3-acetic acid (IAA) which generates most potent auxin effects natively and functions in plants. Along with this, 4-chloroindole-3-acetic acid (4-Cl-IAA), indole-3-butyric acid (IBA) and phenyl acetic acid (PAA) are other endogenous auxins present in plants (Simon and Petrášek, 2011).

2.3.1 Structure of ARFs proteins

The auxin and IAA act by binding to auxin response factors (ARF) and then controls the expression of genes which confers to regulate auxin hormone dependent growth and developmental processes of cells and tissues. The general structure of ARFs consists of 3 main domains, namely DNA binding domain (DBD), middle region (MR) and C-terminal carboxyl interaction domain (CTD). Most of ARFs get binds to generic auxin response element (AuxRE). The basic AuxRE identified is TGTCTC in promoter region. MR designates whether the ARF will activate or repress the targeted genes (Boer *et al.*, 2014). The ARFs activator domains (ADs) are rich in amino acids glutamine, serine and leucine residues whereas repressor domains (RDs) have serine, proline, leucine and glycine residues. These domains are present in between DBD and CTD i.e. in MR (Guilfoyle and Hagen, 2007).

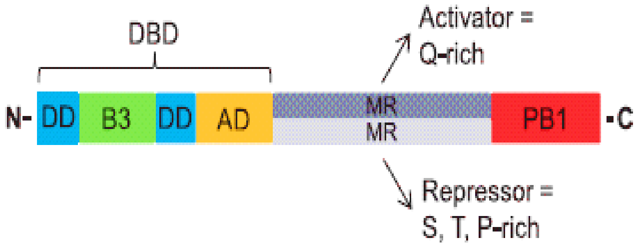


Figure 5 : Schematic modular structure of ARF protein (Chandler, 2016)

N-terminal DBD domain of ARF is of B3-type, which is also found in other plant TFs and function outside the Auxin response pathway too. DD is the dimerization domain that mediates homo-dimerization of the tandem repeats of AuxREs and activates ARFs. The PB1 (previously known as domain III and domain IV) domain present at CTD stands for Phox and Bem1 domain. This domain is responsible for heterotypic protein – protein interactions in between ARFs and IAA/Aux proteins through electrostatic contacts. These domains are also known in facilitating homo-dimerization and oligomerization in between IAA proteins/Auxin with ARF TFs (Chandler, 2016a).

2.3.2 Regulatory mechanism of ARFs in plants

Regulation and activation of ARF is demonstrated in *Arabidopsis thaliana* ARF. The transcription of ARF is dependent on the level of Auxin present in the cell.

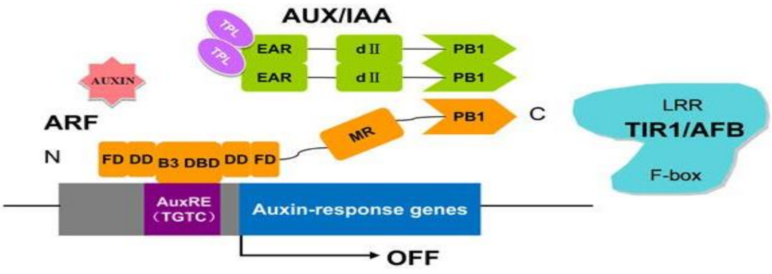


Figure 6 : Regulatory mechanism of ARF in presence of low/no auxin concentration (adapted from Li et. al. 2016)

In presence of low amount or absence of auxin, ARFs activity is inhibited as they form dimer with the Aux/IAA proteins along with enrollment of TOPLESS (TPL), a co-repressor.

As a result, auxin responsive genes are repressed and no any gene expression is initiated.

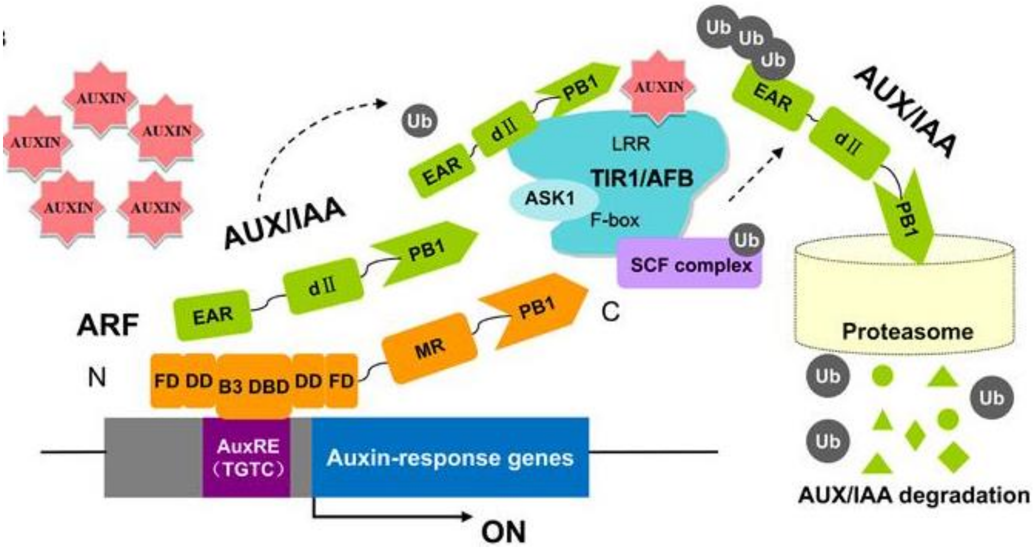


Figure 7 : Regulatory mechanism of ARF in presence of high auxin concentration (Li *et al*., 2016)

In the presence of sufficient amount of Auxin hormones, degradation of Aux/IAA proteins occurs through ubiquitination by SKP-Cullin-F-box$^{TIR1/AFB}$ (SCF$^{TIR1/AFBs}$) E3 ubiquitin ligase complex and finally complete degradation by 26S proteasome. These complexes consist of auxin receptor TRANSPORT INHIBITOR RESPONSE (TIR1)/AUXIN RECEPTOR F-BOX PROTEINS (AFBs). Then after, ARF freed and hence, regulates the transcription of its target auxin response genes (Li, Xie, Hu, and Zhang, 2016a; Rademacher *et al*., 2011).

2.3.3 Roles of ARF transcription factor in plants

The auxin plays an important role as plant hormone initiating from regulating the plant growth and development to defense against the pathogens and abiotic stresses for survival of the plant. The auxin mediated gene transcription has significant roles in determining the patterning of root apical meristem (RAM), shoot apical meristem (SAM) to senescence of the plant cells. Auxin can either act as repressor or activator in gene expression of the ARFs. In *Arabidopsis,* PIN FORMED (PIN) family membrane efflux facilitators established auxin maxima concentration for positioning and growth of new organs. The mutation In PIN and ARF TFs family leads to formation of abnormalities in hypocotyls and roots (Carey and Krogan, 2017; Krogan, Berleth, Marcos, and Weiner, 2016).

AtARF2 is known for development of floral organs and seedlings grown in presence and absence of light whereas SlARF2 is responsible for fruit ripening in tomato in combination with ethylene. AtARF3, 4 helps in developing reproductive and vegetative tissues whereas AtARF5 and 6 imparts its roles in embryo and flower development respectively (Guilfoyle and Hagen, 2007; Hao *et al*., 2015). Mutational studies of ARF genes helps in identifying functions of ARF genes in *Arabidopsis*. Discovered arf3/ett affects gynoecium and floral patterning, arf5/monopteros (mp) affects cotyledon, vasculature and root pole development, arf2 regulates responses to both developmental and environmental factors (Chandler, 2016).

Another important role of auxin is in normal development of fruits. After pollination and fertilization, signaling cascade of auxin and gibberellin are up-regulated within ovary. Those fruits which are independent of fertilization can set either naturally or by application of phyto-hormones like auxin and gibberellins exogenously. In fruit initiation cascade, auxin signaling is identified as one of the early events. For example, ARF8 and Aux/IAA results a complex of repressor protein which prohibits gene expression of auxin responsive genes. After pollination, auxin concentration increases and activates ubiquitination process continuing proteolytic deterioration of Aux/IAA resulting expression of fruit specific genes. Fruit development is the initiated along with the initiation of transcription. In case of mutant of ARF8 i.e. arf8 non-functional protein is formed which couldn’t activate the ARF genes properly. But fruit operating genes are activated without pollination and fertilization resulting into formation of parthenocarpic fruit (Wani and Lattoo, 2017).

CHAPTER 3
MATERIALS AND METHODS

3.1 Reagents

The chemicals used in this work were purchased from Merck, Sigma-Aldrich, Hi-Media, New England Biolab (NEB) and Sisco Research Laboratory (SRL) companies. High graded chloroform for molecular biology work was purchased Fischer Scientific. Tris buffer was obtained from SRL which is found to be free from DNase, RNase, proteases. The agarose for fine separation and resolution of small nucleic acids was purchased from Cambrex Bio Science, Rockland, USA. The petri-plates and falcons were from Tarson. For real time PCR, SYBR Green Master Mix was purchased from Applied Biosystem.

The Eppendorf tubes used during the work was from Axygen. Glasswares used were from the company Duran, Borosil. The sterile nitrile powder free gloves used during the experiments were from Kimberly-Clark, USA. The machine used for PCR is Eppendorf master cycler, nexus gradient and centrifuge 5418 R also from Eppendorf.

3.2 Plant Materials

The seeds were provided from the Laboratory of Translational and Evolutionary Genomics, School of Life Sciences, JNU, New Delhi, India. The seeds were prepared for germination after proper sterilization. The seeds were first treated with 0.1% Bavestein for 5 min. and washed with distilled water. Again the seeds were treated with 2% sodium hypochlorite solution for 1 min. and washed properly with distilled water 3 times. The seeds were then ready. After that the seeds were sown in soil containing vermiculites.

The plants were grown in 18h light and 8h dark condition at room temperature 24˚C about a period of 1 month with proper supplement of water. The plants were then ready after 30-40 days for the collection of samples.

The plants were grown until fruit setting and fruits were harvested for study at different developmental stages, young immature/green, breaker and mature stages of *Capsicum chinense* and *Capsicum annuum.*

3.3 Retrieval of ARF family genes of *Capsicum annuum, Arabidopsis thaliana, Solanum tuberosum, Solanum lycopersicum and Oryza sativa* R-programming software

The genes of Auxin Response Factor (ARF) which is a transcription factor (TF) present in different plants were extracted using the KEGGREST package of R-Bioconductor. The standalone version of R was first installed if not available in the system. Then the *KEGGREST* package was installed into R bioconductor library. The KEGGREST package mainly provide a query platform to user to facilitate query from KEGG (Kyoto Encyclopedia of Genes and Genomes) database which is main resource for understanding high-level functions and utilities of the biological system, such as the cell, the organism and the ecosystem, from molecular-level information, especially large-scale molecular datasets generated by genome sequencing and other high-throughput experimental technologies. Further, to retrieve the ARF family gene following R-script was used.

```
> library(KEGGREST)  ## calling package into current session
> ## Using keggFing function gene information was retrieved
> ## Below is usage of keggFind function
> ##keggFind(databaseID, query, option = c("search terminology"))
> keggFind (“T04646”,c(“auxin response factor”))  ## Finds entries with matching query keywords in a given database
> ## Lists of ARF genes will appear on the screen.
> ## Now, saving previous command output in user defined object called “arf_cann”
> arf_cann= keggFind (“T04646”,c(“auxin response factor”))
> write.table(arf,file=“arf.txt.”,quote=F,sep=“\t”) ## writing results in tabular text format in current working directory
```

Then go to your working directory and open file containing your result generated from above script. Retrieved gene symbol/LocusID from the file and retrieved nucleotide sequence for corresponding genes. In next step, *keggGet()* fucntion was used to retrieve nucleotide sequences for given genes. Maximum 10 gene symbols at a time can be supplied as query for retrieving nucleotide sequence.

Following script was used to fetch sequences.

```
> ## Below is usage of keggGet functions
> ## keggGet(dbentries, c("aaseq", "ntseq"))
> arf.fasta_Cann =keggGet(c(“"LOC107864089", "LOC107848333",
"LOC107857426", "LOC107859326","LOC107862124"”),“ntseq”)
> write.table(arf.fasta_Cann, file= “Capsicum_arffasta.txt”, quote=F, sep=“\n”)
## saving results in a text file separated with new line character
```

All the retrived genes are now available in fasta format in a new notepad. After this the further work is preceded.

List of KEGG database IDs for each species used for extracting gene sequences are:

a.	*Capsicum annuum*	T04646
b.	*Arabidopsis thaliana*	T00041
c.	*Solanum tuberosum*	T02981
d.	*Solanum lycopersicum*	T02665
e.	*Oryza sativa*	T01015.

3.3.1 Chromosomal distribution of ARF genes in *Capsicum*

The genomic distribution of ARF genes on *Capsicum* chromosome were visualized using R *karyoplotR* package. *Capsicum* genome was forged in BSgenome R package library to create a *BSgenome.Cannuum.zunla.refv1* pacakge which could facilitate genome wide query platform for *Capsicum* along with its annotation information in R. This *BSgenome.Cannuum.zunla.refv1* was already created in our laboratory. The ARF gene chromosomal cordinates (start and end position on chromosome) was used by karyoplotR package and following R script was used:

```
> library(karyoploteR)    ## loading the package in current session
> library(BSgenome.Cannuum.zunla.refv1)  ## loading Capsicum BSgenome package
> seqlen=seqlengths(Cannuum) ## checking each chromosome size
```

```
> chr.len=seqlen[grep("NC_018552.1|NC_024624.1",names(seqlen),invert = T)]

 ## excluding chloroplast and mitochondiral chromosome for visualization

### making object for zunla chromosomes

> myIdeo=GRanges(seqnames=names(chr.len),ranges=IRanges(start=1,width = chr.len))

### making data frame and then GRanges object for ARF genes to be displayed on
chromosomes

> ARF_genes<- toGRanges(data.frame(chromosome = c("NC_029977.1", "NC_029977.1",
"NC_029977.1", "NC_029978.1", "NC_029979.1", "NC_029979.1", "NC_029979.1",
"NC_029980.1", "NC_029980.1", "NC_029980.1", "NC_029981.1", "NC_029982.1",
"NC_029983.1", "NC_029983.1", "NC_029983.1", "NC_029984.1", "NC_029984.1",
"NC_029985.1", "NC_029985.1", "NC_029985.1", "NC_029986.1", "NC_029987.1",
"NC_029987.1", "NC_029988.1"), start = c(1624661, 271574440, 300853606, 134474092,
9129637, 17927460, 25036537, 3949684, 104618950, 204804436, 161756942, 7658096,
11509023, 106323324, 134522301, 124280278, 136288859, 15536545, 122791286, 232571312,
158010081, 3097640, 2240074, 82597475),  end = c(1629988, 271581458, 300854004,
134481308, 9135438, 17928402, 25037942, 3953336, 104633632, 204813550, 161763811,
7663513, 11514365, 106333863, 134533619, 124285243, 136295018, 15555036, 122795329,
232575885, 158026307, 3106354, 2245104, 82603999), gene = c("LOC107864089;ARF9-like",
"LOC107848333;ARF6-like", "LOC107857426;ARF9-like", "LOC107859326;ARF3",
"LOC107862124;ARF2-like", "LOC107862483;ARF9-like", "LOC107862705;ARF19-like",
"LOC107866931;ARF5", "LOC107867903;ARF17", "LOC107868273;ARF8-like",
"LOC107870938;ARF19-like", "LOC107875585;ARF18-like", "LOC107877667;ARF1-like",
"LOC107877949;ARF19-like", "LOC107878007;ARF19-like", "LOC107839300;ARF9-like",
"LOC107839829;ARF1", "LOC107841891;ARF6-like", "LOC107842580;ARF1-like",
"LOC107843081;ARF18-like", "LOC107843346;ARF18-like", "LOC107847819;ODE2;ARF4",
"LOC107848100;ARF18", "LOC107850591;ARF2-like"), stringsAsFactors=FALSE))

> ARF_genes=sort(ARF_genes)  ## sorting genes based on chromosomes

## plotting chromosomes wise gene location

> kp=plotKaryotype(genome = myIdeo) ## plotting first chromosomes

> kpAddBaseNumbers(kp,cex=0.9)  ## adding base pair position on chromosome

####plotting genes on chromosomes

> kpPlotMarkers(kp, data=ARF_genes, labels=ARF_genes$gene, text.orientation = "horizontal",
r1=0.7, cex=0.8, marker.parts = c(0.3, 0.7, 0.2),label.color = c("blue"))
```

3.3.2 Phylogenetic analysis

After extracting the ARF genes of different plant species multiple sequence alignment (MSA) was performed using Clustal Omega (www.ebi.ac.uk/Tools/msa/clustalo) with default parameter. The Clustal Omega program uses a seeded guide trees and HMM profile-profile techniques to generate alignments between **three or more** sequences. Based on MSA result, a phylogenetic tree was constructed representing the phylogeny among the different species based on ARF family gene homology. *The combined DNA sequences in fasta format of 5 species all together 127 sequences were given as input file and ClustalW was chosen as output format along with default parameter and then submitted. Aligned sequences appear on the screen and phylogeny was observed. Data of phylogenetic tree which is* Neighbour-joining tree without distance corrections was exported in phylogeny format. Then, for constructing circular phylogenetic tree the phylogeny format file was imported in iTOL (INTERACTIVE TREE OF LIFE- https://itol.embl.de/upload.cgi.), which is a web based tool (Letunic and Bork, 2016).

3.3.3 In-silico identification of ARF specific SSR

For SSR identification, 3kb upstream and 3kb downstream genes of ARF were downloaded; SSR were identified using Gramene SSR tool (http://archive.gramene.org/db/markers/ssrtool) using cutoff value of atleast 3 repeats 8for each SSR in di, tri tetra, penta and hexa category. This Simple Sequence repeats identification tool (SSRIT) uses perl regular expression to find perfect sequence repeats within a sequence. It detects simple repeats in 2 to 10 bases length excluding mono nucleotide repeats. It generates output in a tabular format and provides following details about sequence repeats Dinucleotide, trinucleotide, tetranucleotide and pentanucleotide SSR in different chromosome were represented in circular diagram.

3.4 Heat–Map generation

Overall 28 ARF genes from *Capsicum* were retrieved and heatmap was generated using ggplot2 R-package based on the transcriptome data of fruits of three different *Capsicum* species i.e. *Capsicum annuum, Capsicum frutescens* and *Capsicum chinense* which already existed in the Laboratory of Translational and Evolutionary Genomics, School of Life Sciences, JNU. The transcript expression was described as in term of FPKM

(Fragments Per Kilobase of transcript per Million mapped reads) values which were generated using aligning transcriptome data against the available reference genome using TopHat, followed by assembling and calculating transcript abundance as well as their differential expression using cufflink, cuffmerge and cuffDiff.

3.5 Primer designing

The primers for the selected ARF genes were first designed by using the NCBI website. The gene sequences and their annotation were downloaded and viewed using SNAP Gene viewer software. The exon region was selected, copied and pasted in the query box of blast 3 primer designing tool. . The primers used in this work are designed from (https://www.ncbi.nlm.nih.gov/tools/primer-blast) Primer 3 and BLAST of NCBI. Then parameters were changed as per requirement and numbers of possible primers were displayed on the screen. After that, best primer for the Genome was chosen among the displayed one by analyzing its Tm in between 55˚C-60˚C, GC content of 50% to 60% and product length in between 70bp to 200bp. Self-complementary of the primers less than 3 or 0 is the best one.

Table 1: List of ARF genes primers used in this study

S.N.	Locus ID	Primer name		Sequence	Size	Tm	%GC	Product size
1.	LOC107856233	Ca_ARF3FP	FP	AGCAGGGGAGAAGAAATGCT	20	59.00	50.00	118
		Ca_ARF3RP	RP	CTGCTCACTATGACCTTGAGGA	22	59.24	50.00	
2.	LOC107857083	Ca_ARF4FP	FP	TTCTGAGCTGTGGCATGCTT	20	60.25	50.00	85
		Ca_ARF4RP	RP	GTGCCCCTGTGGGAAGTAAT	20	59.67	55.00	
3.	LOC107848333	Ca_ARF6FP	FP	TAGAGTTGTGGCATGCGTGT	20	59.97	50.00	78
		Ca_ARF6RP	RP	TGACCTTGAGGTACAGCACTC	21	59.38	52.38	
4.	LOC107859326	Ca_ARF8FP	FP	GGAAAGCAATCTTCTGCCGC	20	59.82	55.00	158
		Ca_ARF8RP	RP	GGAAAGCAATCTTCTGCCGC	20	60.18	55.00	
5.	LOC107862124	Ca_ARF9FP	FP	ACCATTCCAAGGCCAGATCG	20	60.11	55.00	107
		Ca_ARF9RP	RP	CTAGACCTTGGCCTTTTTGGC	21	59.45	52.38	
6.	LOC107866931	Ca_ARF12FP	FP	TTATTGCTGGGAGTGAGGCG	20	60.11	55.00	118
		Ca_ARF12RP	RP	TATTAGCAGCAGCATGGGCA	20	59.82	50.00	
7.	LOC107868273	Ca_ARF14FP	FP	TGGGTATTCCTAGCAGGCAG	20	58.87	55.00	81

		Ca_ARF14RP	RP	AAAAACCGCCATGCGTACTG	20	59.76	50.00	
8.	LOC107870938	Ca_ARF15FP	FP	GCAGGGGAGAGGATTAACCG	20	59.89	60.00	98
		Ca_ARF15RP	RP	GGACGCTTGGAGCTGAAGAA	20	60.32	55.00	
9.	LOC107877949	Ca_ARF18FP	FP	CCAGGTTCTACCACCGACAA	20	59.32	55.00	94
		Ca_ARF18RP	RP	ATCCGTAACCACTGACTGGC	20	59.75	55.00	
10.	LOC107878007	Ca_ARF19FP	FP	AGTCACAGCTTCTCTTGGGT	20	58.57	50.00	137
		Ca_ARF19RP	RP	AAACGGGCTGTTGTTTGCAG	20	60.18	50.00	
11.	LOC107841891	Ca_ARF22FP	FP	CTAGGCTTGCAGCCTGACAT	20	60.04	50.00	128
		Ca_ARF22RP	RP	TTCCGGGAAACCTTGAAGCA	20	61.48	60.00	
12.	LOC107842580	Ca_ARF23FP	FP	AAACTGCCGATTTCCAAGACT	21	58.14	42.86	122
		Ca_ARF23RP	RP	TCCTCGCCCACAAATCTCAT	20	59.09	50.00	
13.	LOC107843081	Ca_ARF24FP	FP	CAGCATTGCAGATCCGTTGG	20	59.90	55.00	124
		Ca_ARF24RP	RP	CGAATGGGATCGGAGACCTG	20	59.97	60.00	
14.	LOC107843346	Ca_ARF25FP	FP	AAGAGACGAAGAGGAGAAGTGC	22	59.77	50.00	121
		Ca_ARF25RP	RP	ATGCTCTGCGTGTCCTTGA	19	59.33	52.63	

The selected primers of all 14 genes were ordered and synthesized by Eurofins Genomics India Pvt. Ltd. companies. Only 14 out of 28 genes are selected on the basis of their differential expression pattern observed in the Heat-map of ARF gene (Figure:11). The primers were obtained in lyophilized form hence it was dissolved in aliquoted amount of autoclaved MQ to make stock concentration of 100pmol/µl. The amount of water was mentioned in the Oligonucleotide Synthesis Report provided by the company along with the primers. The primers of working concentration 10pmol/µl were prepared from the stock.

3.6 DNA extraction of *Capsicum* leaves

The leaves sample were collected from glass-house and stored in liquid nitrogen. The leaves were crushed with sterile mortar and pestle in liquid nitrogen. Powdered sample of leaves were then transferred in 50 ml sterile falcon containing 15 ml of pre-warmed cetyl trimethyl ammonium bromide (CTAB) DNA extraction buffer, a cationic detergent which facilitates separation of polysaccharides of the plant tissues. Along with this 30 µl of 0.2% β-mercapto-ethanol and pinch of polyvinyl pyrolidone (PVP) was added and mixed well, and incubated at 65˚C water bath with occasional stirring for 1 hour. After sufficient incubation, tubes were allowed to cool down at room temperature, then added 15ml of Chloroform:Isoamyl alcohol in the ratio 24:1. The contents were mixed by inverting tubes for about 5 minutes and centrifuged at 13000 rpm at 20˚C for phase separation. The separated layers contain DNA in the upper-most aqueous layer, and other organic soluble molecules in the lower-most layer. The aqueous upper-most layer was transferred into new falcon tube, added 10ml of ice-cold isopropanol and kept at -20˚C for 2 hours. This storage step can be prolonged for enhanced yield by keeping at 4˚C overnight. Next day, DNA was pelleted by centrifugation at 4˚C with speed of 13000 rpm for 15 min. The supernatant was discarded and pellet was washed with 5ml of 70% ethanol at 13000 rpm for 10 min at 4˚C. The pellet was air dried by keeping at room temperature with the tubes open. Then 1 ml of autoclaved MQ water was added for dissolving DNA pellet and kept at 4˚C overnight.

3.6.1 DNA quality confirmation

Finally the extracted DNA samples were quantified by gel electrophoresis and nano-drop reading. 1% of agarose gel was prepared in 50 ml of 1X Tris-Borate-EDTA (TBE) buffer and boiled enough to dissolve all agarose particles properly. Then after, allowed to cool and 10 µl of 20ng/µl ethidium bromide (EtBr) was added, mixed properly and gel was casted in tray having required comb. Finally the agarose gel was prepared and gel was placed in gel electrophoresis tank having 1X TBE buffer enough to dip the gel. Then 2 µl samples were loaded after mixing with 2 µl of loading dye. 1kb ladder was also loaded. Then gel was run at 100 V for 45 min. After that, gel was exposed to UV in gel-doc system and picture was acquired.

Nano-drop reading was taken in 2000 ND nano-drop. The software was opened and program was set for DNA. The arm was cleaned before use with a clean tissue paper. 1 µl of autoclaved MQ was placed and set as blank. Then the place was cleaned and again 1 µl of sample was placed on and click on measure. The results obtained were noted as concentration of DNA and ratio of 260/280.

3.6.2 Standardization of primers on gDNA by Temperature Gradient Polymerase Chain Reaction (PCR)

The obtained DNA having good integrity and quality was used later on for checking the activity of primers. For the proper amplification of the target genes the primers should work very well. In different conditions of PCR the amplification may varies so we need to standardize the conditions like annealing temperature, number of PCR cycle, $MgCl_2$ are concentration few to mention.

For this temperature gradient PCR reaction were set up to standardize the annealing temperature of each primer. The master mix for the PCR was prepared as following.

Table 2: PCR reaction Mix

S.N.	Reagents	Reaction mix volume (µl)	Final concentration
1.	*i* - Taq^{TM} 10X PCR Buffer	1	10 mM
2.	10 mM dNTPs mix	0.5	
3.	Forward primer	0.5	10 pmol
4.	Reverse primer	0.5	10pmol
5.	*i* - Taq^{TM} DNA polymerase	0.2	500 U/µl
6.	DNA template	1	100 ng
7.	Nuclease free water (NFW)	6.3	
	Total volume	10	

After preparing the mixture the PCR was set in following conditions in the PCR machine of Eppendorf master cycler.

Table 3: PCR condition

S.N.	Step	Cycle	Temperature	Time
1.	Initial denaturation	1	94˚C	4 min.
2.	Denaturation		94˚C	2 min.

3.	Primer annealing	30	52˚C-62˚C	45 sec.
4.	Extension		72˚C	30 sec.
5.	Final extension	1	72˚C	2 min.
6.	Hold	-	4˚C	-

The amplified products were performed agarose gel electrophoresis in 1.2% agarose gel in TBE buffer and observed under UV and captured using ALPHA IMAGE GEL DOC System software.

3.7 Ribonucleic acid (RNA) isolation of *Capsicum* fruit samples

The RNA is isolated from the fruit sample of different stages *Capsicum* by using RNAiso Plus reagent provided from TaKaRa Company of Japan. Using Trizol method for isolating RNA from plant cells gives the high quality and better amount of RNA as compared to Lithium Chloride (LiCl) method around the time period of 1 hour. Trizol is monophasic solution of guanidine isothiocyanate, 38% phenol. The RNA was isolated from fruit sample of all 3 stages i.e. immature, breaker and mature stages of *Capsicum annuum* L. Isc jammu and *Capsicum chinnensce* L. Peren. The samples for isolation of RNA were collected and frozen immediately in liquid nitrogen. Then 50 to 100 mg of sample was crushed in 0.1% di-ethyl pyro-carbonate (DEPC) treated and autoclaved mortar and pestle with liquid nitrogen. DEPC treatment is done to inhibit the activity of RNase enzyme during the isolation of RNA and it also degrades the DNA. Finely powdered sample was transferred to micro centrifuge tube containing 1 mL of trizol solution and homogenized properly by slight vortex following incubation of 5 minutes at room temperature. Chloroform (200 µl) was added after incubation and shaken tube vigorously for 15 seconds for proper mixing and again left for incubation at room temperature for 5 minutes. Then tube was centrifuged at 13000 rpm at 4˚C for 15 minutes. After centrifugation 3 layers were obtained in the tube, a red lower organic layer containing proteins, carbohydrates, cell debris, an interphase consisting DNA and uppermost clear layer having RNA. Hence, uppermost layer was carefully transferred to new tube along with 0.5 mL of isopropanol, mixed well and incubated at room temperature for 10 minutes for precipitation of RNA. Again tube was centrifuged for 10 minutes at 4˚C with speed of 13000 rpm. The supernatant was removed and pellet was washed with 1 mL of 75% ethanol and again centrifuged at 1000 rpm for 5 min again at 4˚C. Temperature is maintained at 4˚C throughout the procedure in order to prevent the degradation of RNA as it is heat sensitive. The wash solution i.e. 75% was discarded and white pellet was air dried at room temperature. Finally, isolated RNA was dissolved in 30 µl nuclease free water from SRL, BioLit^{TM}. Tube was then incubated in heat block for 15 minutes maintaining temperature 55˚C for complete dissolution of RNA. The isolated RNA is now ready for proceeding further steps and if not used then stored at -80˚C.

Quantification of isolated RNA

The concentration of RNA is quantified using the spectrophotometric analysis using Nano-drop ND 2000. The RNA concentration was measured in the unit ng/µl and purity of RNA is checked by calculating absorbance ratio OD_{260}/OD_{280} and OD_{260}/OD_{230} . For the better quality RNA the ratio should be in between 1.8 to 2.1. First of all, blank was set by using nuclease free water and optical densities of sample were measured taking 1 µl of each sample.

3.7.1 Observation of RNA quality by denaturing agarose gel electrophoresis

Denaturing gel electrophoresis is done for checking integrity of the extracted RNA samples. Denaturing gel is prepared using formaldehyde and 10X MOPS (3-(N-morpholino) propanesulfonic acid) buffer in agarose. Agarose (1g) was weighed and and melted adding 72 mL of DEPC treated autoclaved water. 10X MOPS (10mL) was added after cooling gel around temperature 55˚C followed by addition of 18 mL 37% formaldehyde. While adding the MOPS and formaldehyde care should be taken and performed inside the fume hood in order to minimize the spreading toxic fume in the laboratory. Finally, 20 µL of 10µg/µL EtBr was added and swirled properly to even mixing of all contents, and poured for solidification. Then, sample was loaded along with RNA loading dye, run in 1X MOPS buffer at 70V for 1 hour and observed under UV in ALPHA IMAGER GEL DOC system.

3.7.2 Synthesis of complementary DNA (cDNA)

The cDNA abbreviates complementary DNA that is synthesized by taking mRNA as a template and the reaction is catalyzed by special called reverse transcriptase. Hence, the process for synthesizing cDNA is also renowned as Reverse transcription polymerase chain reaction i.e. RT-PCR. For the synthesis of cDNA for this project, PrimeScript^{TM} 1st strand cDNA Synthesis Kit from TaKaRa was used. The kit contains a M-MLV (Moloney Murine Leukemia Virus)-derived reverse transcriptase which synthesize first strand cDNA from total or polyA^{+} RNA. The kit constitutes all the necessary components for cDNA synthesis and the reaction mixture was prepared as follows.

Table 4: Reaction mixture for preparation of Template RNA Primer Mixture

S.N.	Reagents	Reaction mix volume (µl)	Final concentration
1.	50 µM Oligo dT primer	1	5 µM
2.	10 mM dNTP mixture	1	1 mM
3.	Template RNA	Variable (depends on RNA concentration of each sample)	1 µg/µl
4.	RNase free dH_2O	Calculated to make final volume 10 µl	
	Total volume	10 µl	

This mixture was incubated at 65˚C for 5 minutes time duration and immediately cooled on ice. Again, another reaction mixture was prepared according to the following table.

Table 5: Reaction mixture for RT-PCR

S.N.	Reagents	Total Reaction mix volume (20µl)
1.	Template RNA primer mix	10 µl
2.	5X Prime Script Buffer	4 µl
3.	RNase Inhibitor (40U/µl)	0.5 µl
4.	Prime Script RTase (200U/µl)	1 µl

After mixing all components, the synthesis of cDNA was performed by setting the reaction mixture for incubation at 42°C for 35 minutes followed by 5 min at 95°C to inactivate enzyme. Then, finally hold at 4°C for 10 minutes and placed on ice after completing the synthesis process.

3.7.3 Confirmation of synthesized cDNA by PCR

For the confirmation of prepared cDNA PCR was performed using the primers of a constitutive gene named elongation factor-1 alpha (EF-1α). This gene is constitutively expressed in the eukaryotic system. The amplification were performed starting with a 5 min denaturation step at 94°C, followed by 25 cycle of 94°C for 45 seconds, 52°C for 30 seconds and 72°C for 30 seconds. The final extension step was performed for 7 min and hold at 4°C. The reaction mixture for this PCR was prepared as shown in the table.

Table 6: Reaction mix preparation for PCR

S.N.	Reagents	Reaction mix volume (µl)	Final concentration
1.	*i* - Taq™ 10X PCR Buffer	2	10 Mm
2.	10 mM dNTPs mix	1	
3.	Forward primer	1	10 pmol
4.	Reverse primer	1	10pmol
5.	*i* - Taq™ DNA polymerase	0.2	500 U/µl
6.	DNA template	1	100 ng
7.	Nuclease free water (NFW)	13.8	
	Total volume	20	

cDNA from different fruit staged *C. annuum* and *C. chinense* were taken as template for this PCR and performed in Eppendorf master cycler machine. The amplified product was observed by 1.2% agarose gel electrophoresis under UV.

3.7.4 Semi-quantitative expression analysis of ARF genes

For the semi quantitation of ARF genes in fruits of *Capsicum species* at different stages during fruit development, cDNA of the fruits were amplified using the primers of Ca_ARF genes. The expression levels of genes are observed by amplifying the cDNA which is complementary to mRNA (Walker, Worst, and Vrana, 2003) of the fruit samples in our experiment. The amplification was done again by general PCR.

The PCR reaction conditions was set as initial single cycle of 94°C for 10 min and then 30 cycles of: 94°C, 30 seconds; standard annealing Tm for each primer in °C, 45 seconds; 72°C for 30 seconds and final extension of one cycle 72°C for 5 min making hold at 4°C. For the amplification, the reactions were prepared in 0.2 mL PCR tubes from AXYGEN, China and the contents added were as follows;

Table 7: PCR reaction mixture for semi-quantitative PCR

S.N.	Reagents	Reaction mix volume (µl)	Final concentration
1.	*i* - Taq^{TM} 10X PCR Buffer	2	10 mM
2.	10 mM dNTPs mix	1	
3.	Forward primer	1	10 pmol
4.	Reverse primer	1	10pmol
5.	*i* - Taq^{TM} DNA polymerase	0.2	500 U/µl
6.	DNA template	1	100 ng
7.	Nuclease free water (NFW)	13.8	
	Total volume	20	

The amplified PCR products were confirmed by agarose gel electrophoresis. 1.2% Agarose gel containing 0.5µg/ml ethidium bromide was prepared in 0.5X TBE buffer. DNA Gel loading dye was added to the PCR products and mixed well and loaded to the respective well in the gel along with DNA ladder of 100 bp and run for 1 hrs at 70V. The gel was visualized under UV in alpha image gel documentation system.

3.7.5 Quantitative Real Time PCR (qRT-PCR) analysis

The qRT-PCR was performed with PowerUp SYBR Green Master Mix (Thermo Fischer SCIENTIFIC, USA) as per the manufacturer's protocol. The qRT-PCR was conducted in AB17500 Fast system (Applied Biosystems) with the following thermal protocol: 50˚C for 2 minutes, 95˚C for 2 minutes of hold and 30 cycles of 95˚C for 15 sec., 56˚C for 45 sec. The reactions were carried out in duplicates for each sample. Negative control mixture was prepared without addition of template. Then, data generated was calculated using comparative $2^{-(\Delta Ct)}$ to calculate relative expression.

CHAPTER 4
RESULT AND DISCUSSION

4.1 Distribution of ARF genes chromosome of *Capsicum annuum*

Chromosomal distribution of all 28 ARF genes was depicted on *Capsicum annuum* genome, 24 genes were distributed across 12 chromosomes and only 4 genes are identified as pseudogene. The highest number of exons present in the identified genes is 15 and lowest number of exon is 2. The largest genome size among different ARF is 21 kbp and smallest one is 518 bp which belongs to gene having gene ID 107848100 and 107857426 respectively. At least 3 ARF genes are distributed on chromosome 1, 2, 3, 7 and 9 while only a single ARF gene is located on chromosome 2, 5, 6, 10 and 12.

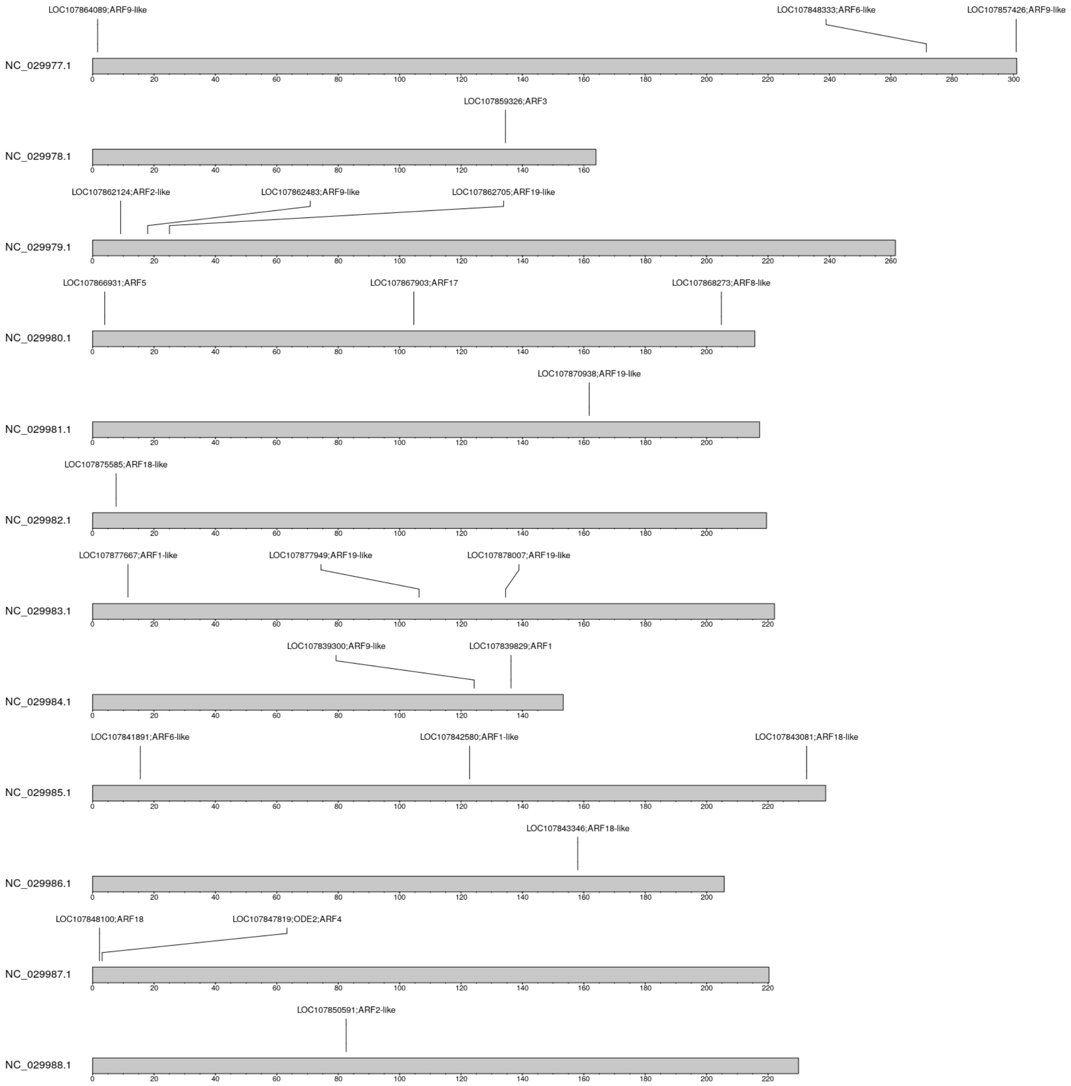
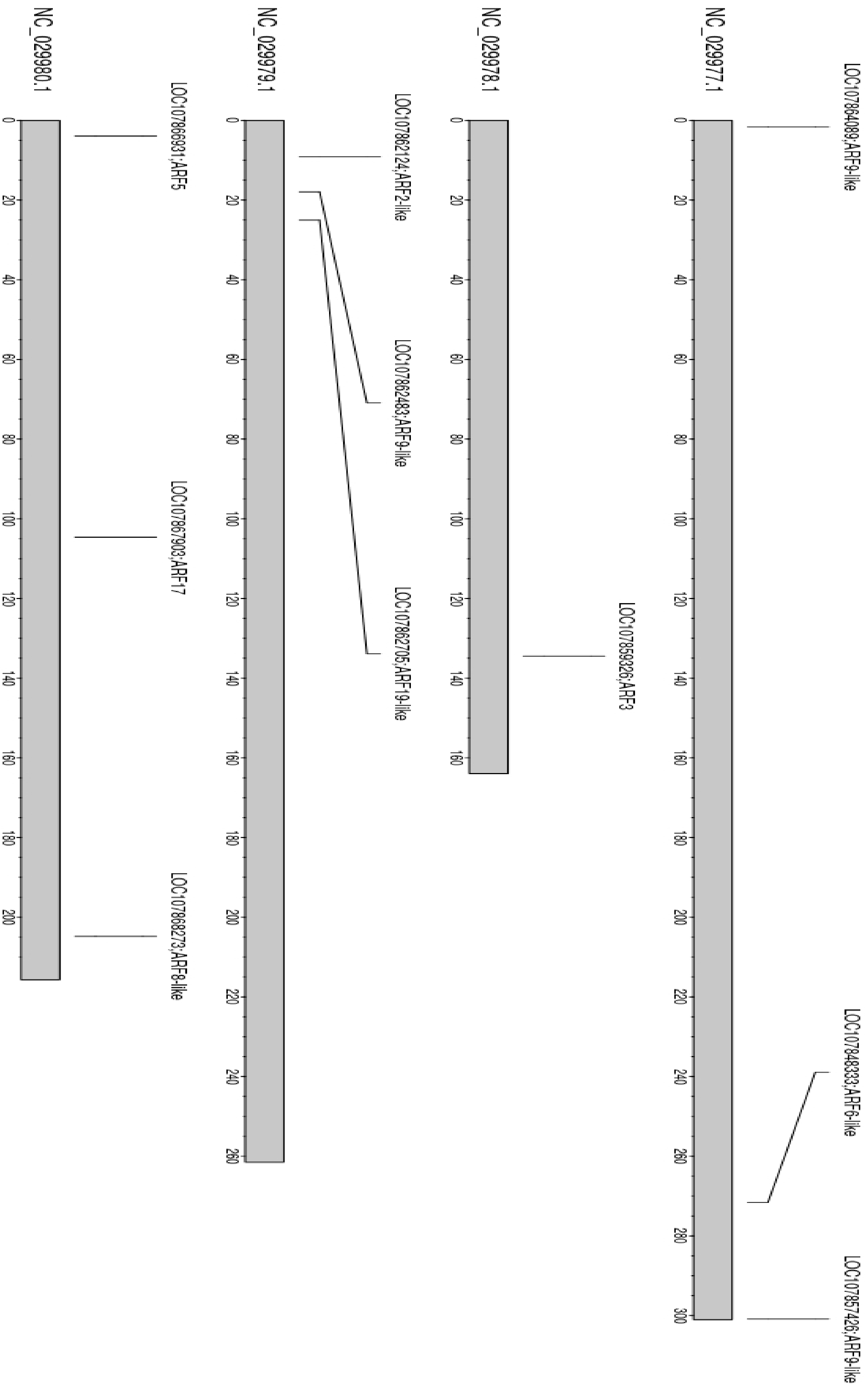
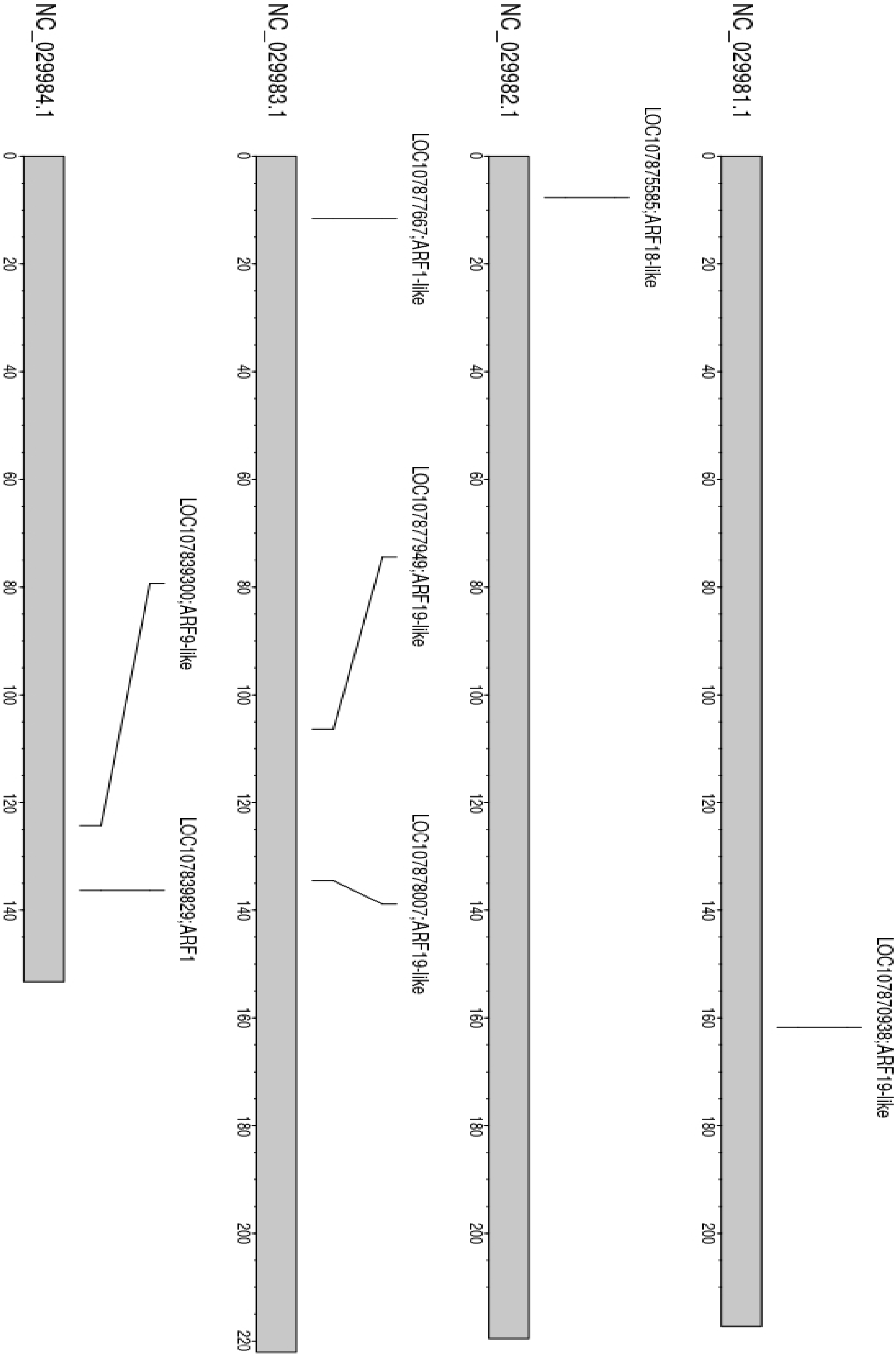


Figure 8: Distribution of ARF genes in different chromosome of *C. annuum* (The enlarged form of the Figure 8 is in following page).


NC_029977.1
LOC107864089;ARF9-like
LOC107848333;ARF6-like
LOC107857426;ARF9-like
0
20
40
60
80
100
120
140
160
180
200
220
240
260
280
300
NC_029978.1
LOC107859326;ARF3
0
20
40
60
80
100
120
140
160
NC_029979.1
LOC107862124;ARF2-like
LOC107862483;ARF9-like
LOC107862705;ARF19-like
0
20
40
60
80
100
120
140
160
180
200
220
240
260
NC_029980.1
LOC107866931;ARF5
LOC107867903;ARF17
LOC107868273;ARF8-like
0
20
40
60
80
100
120
140
160
180
200


NC_029981.1
LOC107870938;ARF19-like
NC_029982.1
LOC107875585;ARF18-like
NC_029983.1
LOC107877667;ARF1-like
LOC107877949;ARF19-like
LOC107878007;ARF19-like
NC_029984.1
LOC107839300;ARF9-like
LOC107839829;ARF1
0
20
40
60
80
100
120
140
160
180
200
220

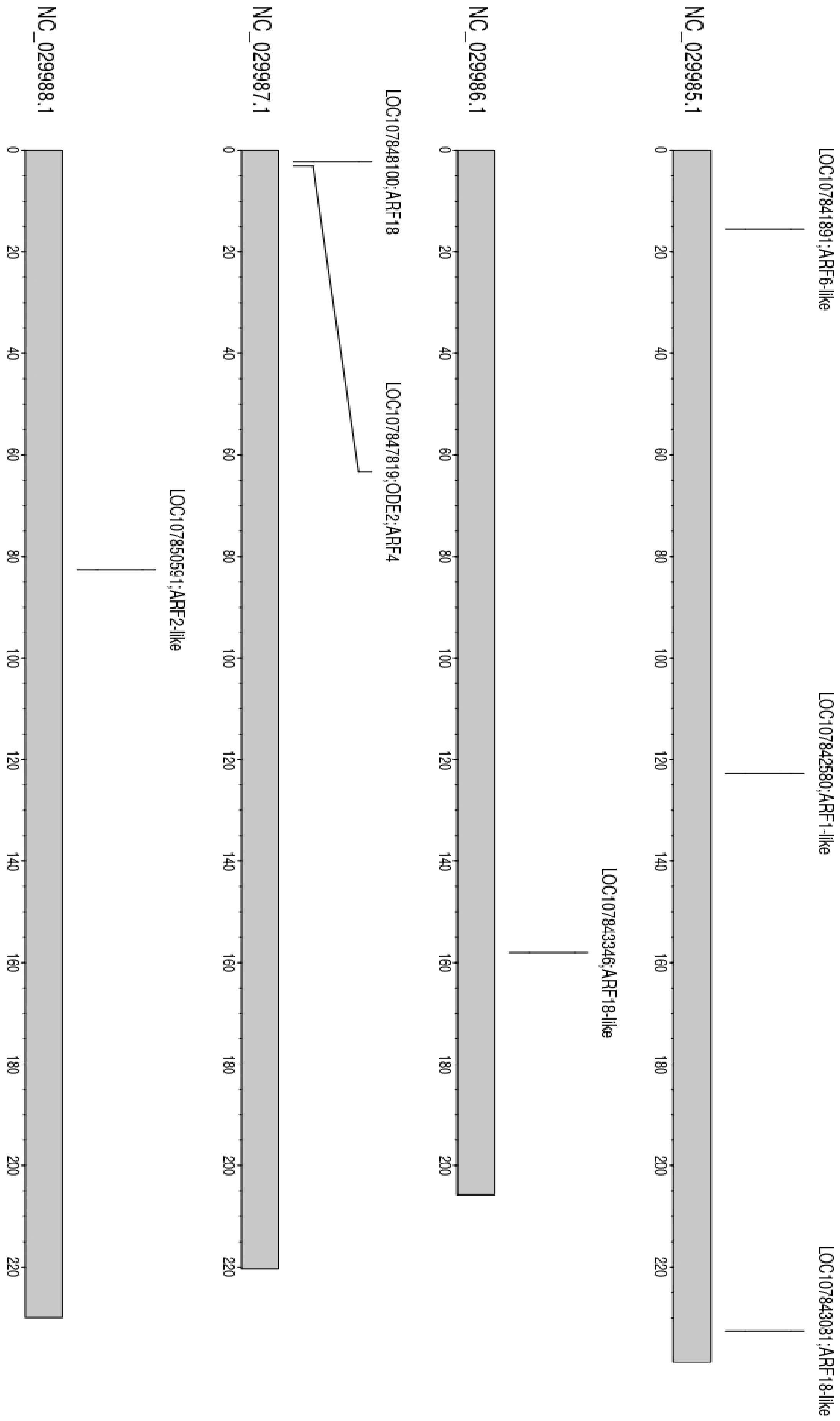

NC_029985.1
LOC107841891;ARF6-like
LOC107842580;ARF1-like
LOC107843081;ARF18-like
NC_029986.1
LOC107843346;ARF18-like
NC_029987.1
LOC107848100;ARF18
LOC107847819;ODE2;ARF4
NC_029988.1
LOC107850591;ARF2-like
0
20
40
60
80
100
120
140
160
180
200
220

Table 8: List of *C. annuum* ARF genes, its gene IDs with chromosome location

S.N.	Gene Name	Gene ID	Location on chromosome Start site..end site	Size (kbp)	Exons
1.	auxin response factor 9-like	107851954	Unplaced scaffold, 25..2732	8.2	7
2.	auxin response factor 8-like	107853231	Unplaced scaffold, 855041..866637	15	14
3.	auxin response factor 6-like	107856233	Unplaced scaffold, 115233..122611	9.6	14
4.	auxin response factor 5-like	107857083	Unplaced scaffold, 7..2038	2.0	6
5.	auxin response factor 9-like	107864089	1, 1624662..1629984	6.9	14
6.	auxin response factor 6-like	107848333	1, 271574499..271581435	9.0	7
7.	auxin response factor 9-like	107857426	1, 300853607..300854004	0.5	2
8.	auxin response factor 3	107859326	2, 134474093..134481204	9.2	12
9.	auxin response factor 2-like	107862124	3, 9129638..9135438	7.5	15
10.	auxin response factor 9-like	107862483	3, 17927461..17928402	1.2	3
11.	auxin response factor 19-like	107862705	3, 25036538..25037942	1.8	4
12.	auxin response factor 5	107866931	4, 3949721..3953336	4.7	9
13.	auxin response factor 17	107867903	4, 104618951..104633632	19	5
14.	auxin response factor 8-like	107868273	4, 204804437..204813550	12	14
15.	auxin response factor 19-like	107870938	5, 161756943..161763811	8.9	13
16.	auxin response factor 18-like	107875585	6, 7658097..7663513	7.0	5
17.	auxin response factor 1-like	107877667	7, 11509171..11514186	6.5	12
18.	auxin response factor 19-like	107877949	7, 106323325..106333857	14	14
19.	auxin response factor 19-like	107878007	7, 134522364..134533618	15	14
20.	auxin response factor 9-like	107839300	8, 124280289..124285203	6.4	15
21.	auxin response factor 1	107839829	8, 136288860..136295018	8.0	15
22.	auxin response factor 6-like	107841891	9, 15536701..15545465	11	14
23.	auxin response factor 1-like	107842580	9, 122791307..122795329	5.2	12
24.	auxin response factor 18-like	107843081	9, 232571960..232575815	5.0	4
25.	auxin response factor 18-like	107843346	10, 158010082..158026291	21	6
26.	ODE2; auxin response factor 4	107847819	11, 3097641..3106354	11	12
27.	auxin response factor 18	107848100	11, 2240075..2245104	6.5	5
28.	auxin response factor 2-like	107850591	12, 82597952..82603893	7.7	15

4.2 Phylogeny analysis of Auxin Response Factor (ARF) TFs of *Capsicum annuum, Arabidopsis thaliana, Solanum tuberosum, Solanum lycopersicum and Oryza sativa*

Using ARF gene sequences from already characterized species such as *Solanum tuberosum, Solanum lycopersicum and Oryza sativa* against *Capsicum annuum* genome database, we identified homologous ARF genes from *Capsicum*s. A total of 28 ARF genes was identified. The reported number of ARF genes reported *in C. annuum* is the highest compared to *Solanum tuberosum, Solanum lycopersicum* and *Oryza sativa* which have equal number of ARF genes in each species I.e. 25. The lowest number of ARF genes is found in *Arabidopsis thaliana* which is 24 as compared to other 4 species. . Common evolutionary ancestry is observed for *Capsicum* ARF genes.

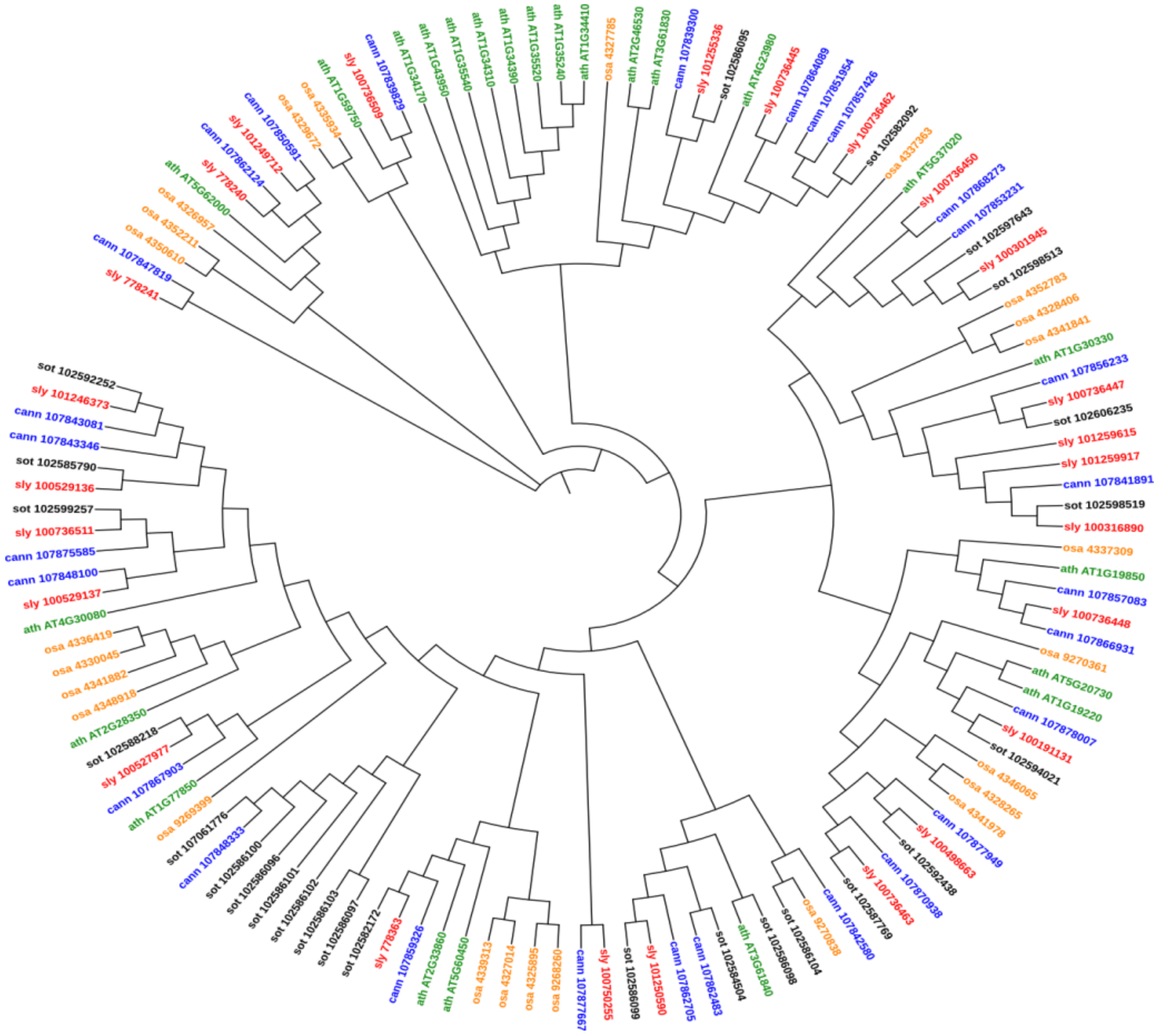


Figure 9 : Circular phylogeny tree of ARF genes of *Capsicum annuum* (blue), *Solanum lycopersicum* (red), *Solanum tuberosum* (black), *Arabidopsis thaliana* (green), and *Oryza sativa* (orange), tree is generated using iTOL.

4.3 SSR distribution in *Capsicum* ARF genes

Simple sequence repeats (SSRs), are DNA tracts in which short base-pair motif is repeated from few to many times in tandem (like ATCATCATCATCATC or TGTGTGTGTGTG). They are also known as Microsatellites and considered as informative molecular genetic markers in many crop species (Saha *et al*., 2003). The SSR present in all 28 Ca_ARF genes have been identified and calculated. The number of dinucleotide SSRs are found in the highest number which is 768 and lowest is of hexanucleotide i.e. 6. Whereas trinucleotides, tetranucleotides and pentanucleotide number counts to 167, 29 and 13 respectively.The graphical representation of SSR present in all genes and their distribution in 12 chromosomes is shown in figure 10.

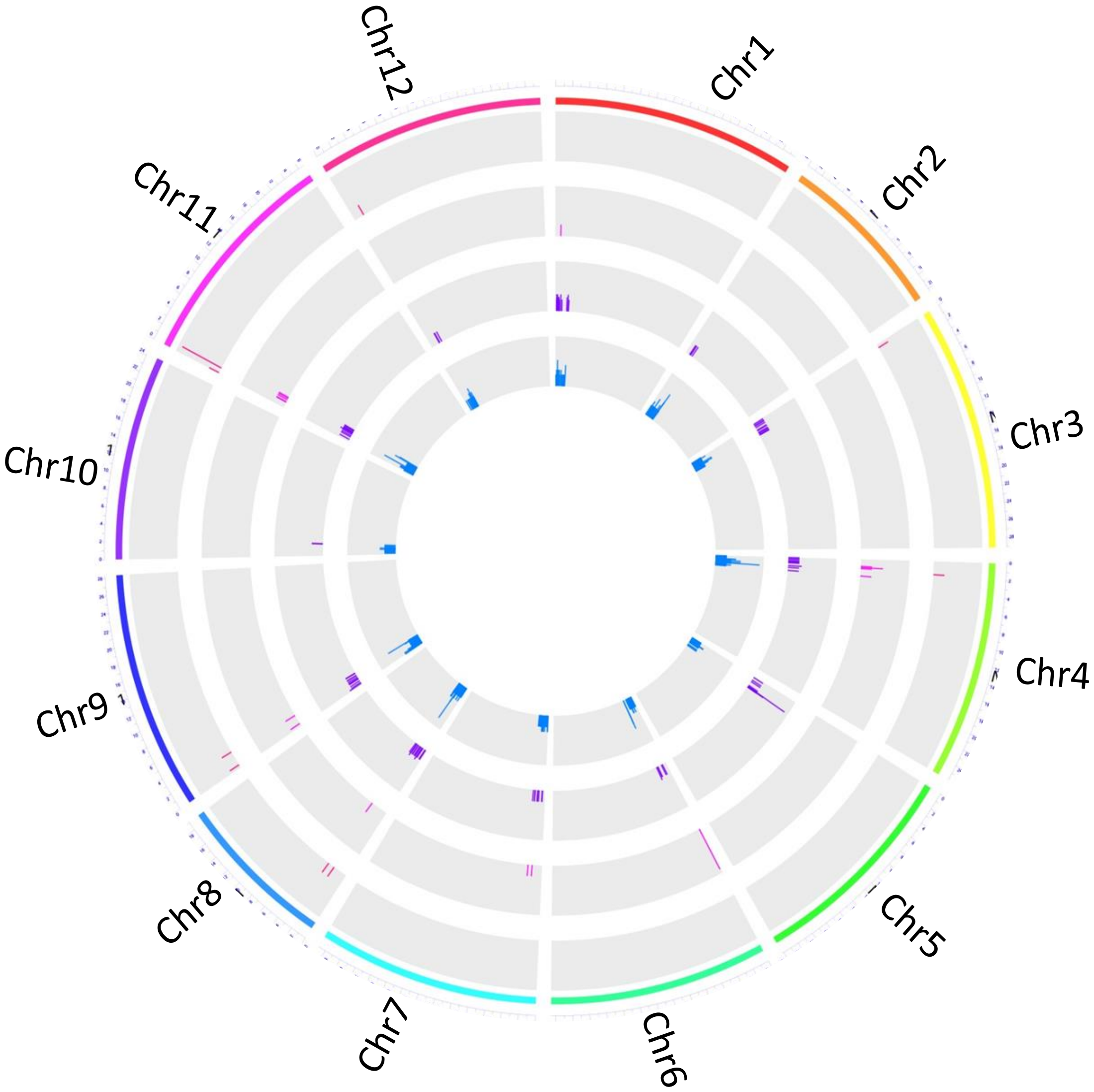


Figure 10 : SSR distribution of ARF genes represented in Map Density. 1st circle represents dinucleotide, 2nd represents trinucleotide, 3rd represents tetranucleotide and 4th represents pentanucleotide and hexanucleotide SSRs.

4.4 Expression analysis of CaARF genes

The expression of ARF genes in flowers and fruits of different species are analyzed using the transcriptome data already available in the laboratory.

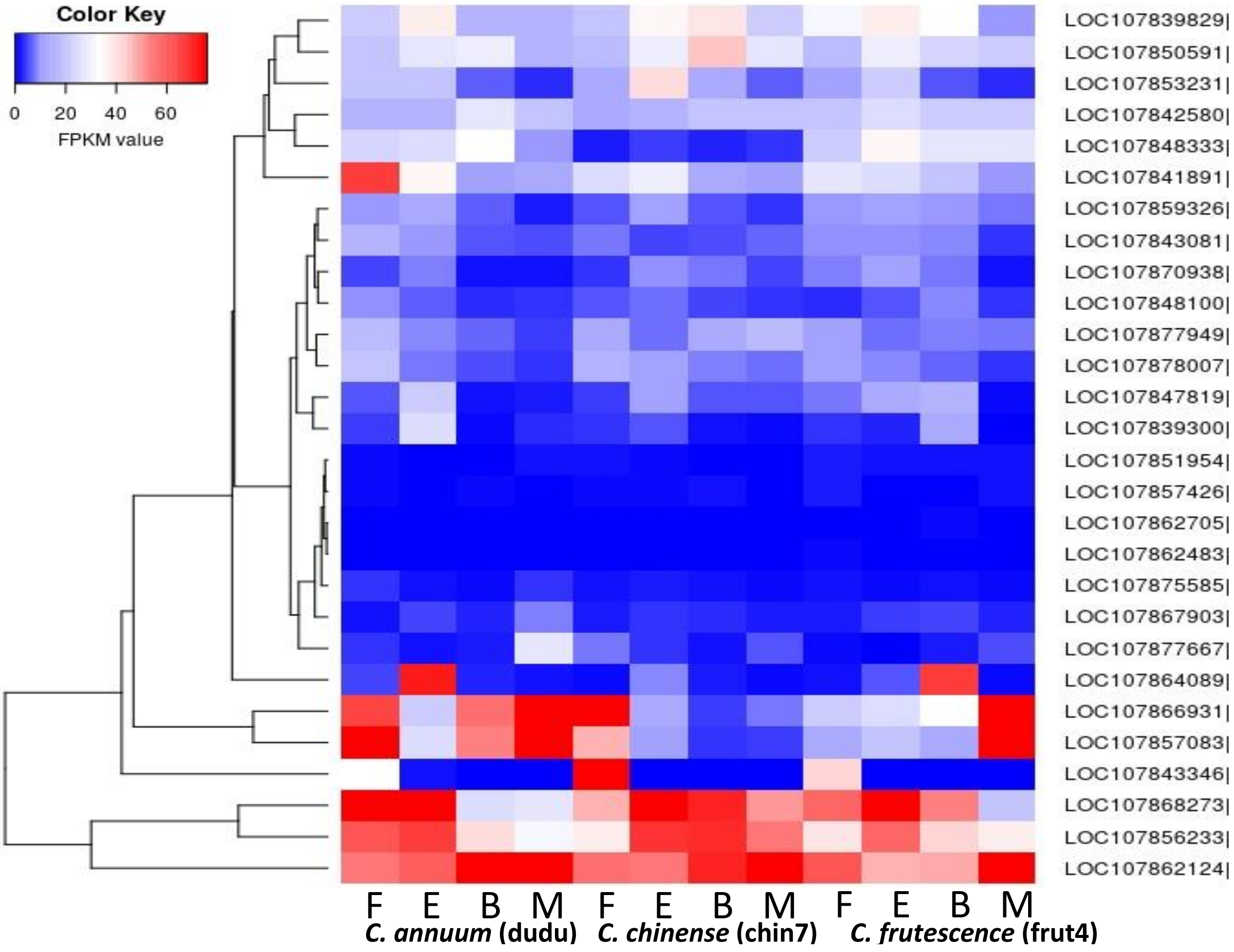


Figure 11 : Heat-map representation of expression profiles for ARF genes from C*apsicum* species. 28 ARF genes are represented as rows and different stages of fruit development in column matrix (F=flower, E=early/green fruit, B=breaker fruit, M=mature/red fruit, FPKM-Fragments Per Kilobase of transcript per Million)

Left portion of the heat map represents clustering pattern of the ARF genes. The color key represents the expression level of the ARF genes in term of FPKM value. The red color indicates high expression of the genes whereas blue color indicates low or no expression of genes. The whitish color signifies moderate expression of ARF gene. The expression pattern of different ARF genes clearly divides them in two major cluster; the upper cluster with low and moderate expression pattern which covering most of the genes while the lower cluster representing the highly expressed genes throughout the samples during fruit development. Somehow, expressions of few ARF genes are not observed during fruit development.

4.5 Standardization of annealing temperature of designed primers on genomic DNA (gDNA) by temperature gradient polymerase chain reaction

The primers were first tested on genomic DNA (gDNA) of the *Capsicum annuum* in order to standardize the annealing temperature of the individual primers. For this temperature gradient PCR was set up in between the temperature of 52-62˚C.

a. Standardization of Ca_ARF3 primer

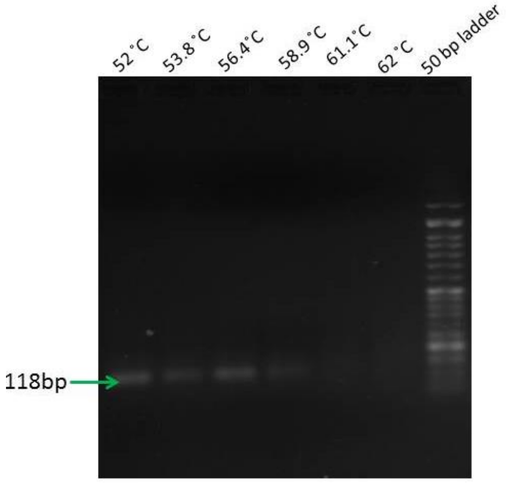


Figure 12 : Standardization of annealing temperature of Ca_ARF3 Primer on *Capsicum annuum*. PCR reaction was performed with gDNA from *Capsicum annuum* for 30 cycles at various temperatures ranging from 52˚C to 62˚C. The PCR product were run on 1.2% agarose gel and visualized under ALPHA IMAGE GEL DOCUMENTATION SYSTEM as shown in figure.

b. Standardization of Ca_ARF4 primer

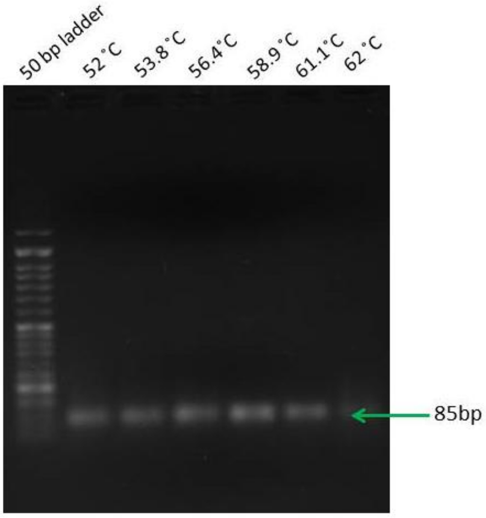


Figure 13 : Standardization of annealing temperature of Ca_ARF4 Primer on *Capsicum annuum*. PCR reaction was performed with gDNA from *Capsicum annuum* for 30 cycles at various temperatures ranging from 52˚C to 62˚C. The PCR product were run on 1.2% agarose gel and visualized under ALPHA IMAGE GEL DOCUMENTATION SYSTEM as shown in figure.

c. Standardization of Ca_ARF6 primer

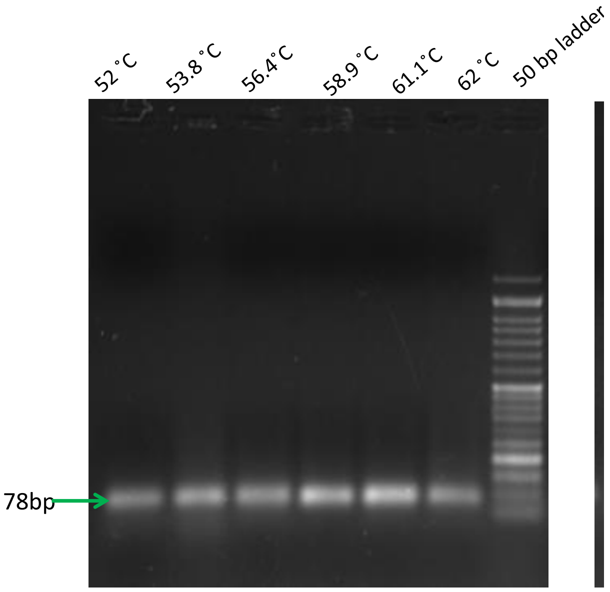


Figure 14 : Standardization of annealing temperature of Ca_ARF4 Primer on *Capsicum annuum*. PCR reaction was performed with gDNA from *Capsicum annuum* for 30 cycles at various temperatures ranging from 52˚C to 62˚C. The PCR product were run on 1.2% agarose gel and visualized under ALPHA IMAGE GEL DOCUMENTATION SYSTEM as shown in figure.

d. Standardization of Ca_ARF9 primer

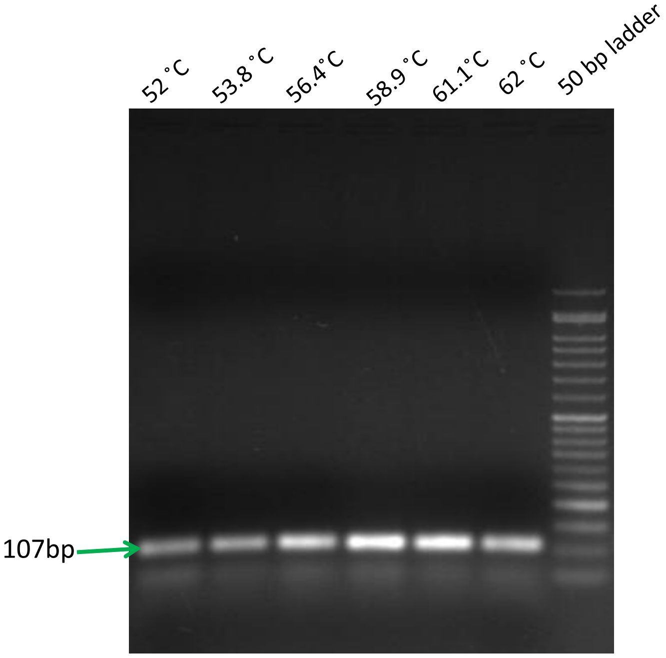


Figure 15: Standardization of annealing temperature of Ca_ARF9 Primer on *Capsicum annuum*. PCR reaction was performed with gDNA from *Capsicum annuum* for 30 cycles at various temperatures ranging from 52˚C to 62˚C. The PCR product were run on 1.2% agarose gel and visualized under ALPHA IMAGE GEL DOCUMENTATION SYSTEM as shown in figure.

e. Standardization of Ca-ARF12 primer

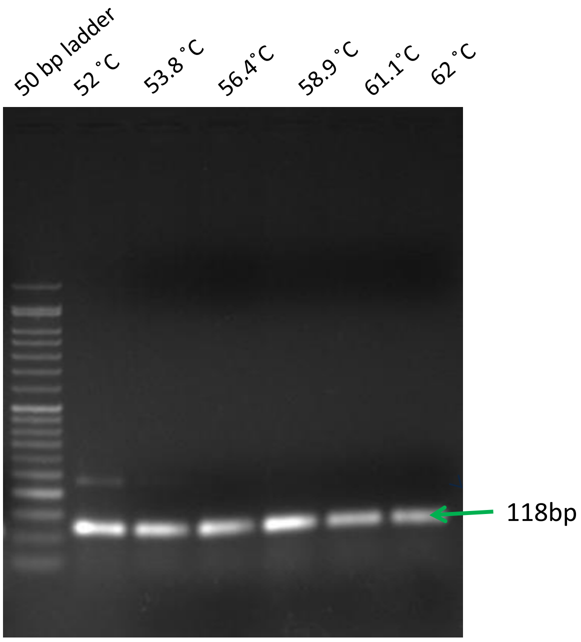


Figure 16 : Standardization of annealing temperature of Ca_ARF12 Primer on *Capsicum annuum*. PCR reaction was performed with gDNA from *Capsicum annuum* for 30 cycles at various temperatures ranging from 52˚C to 62˚C. The PCR product were run on 1.2% agarose gel and visualized under ALPHA IMAGE GEL DOCUMENTATION SYSTEM as shown in figure.

f. Standardization of Ca_ARF14 primer

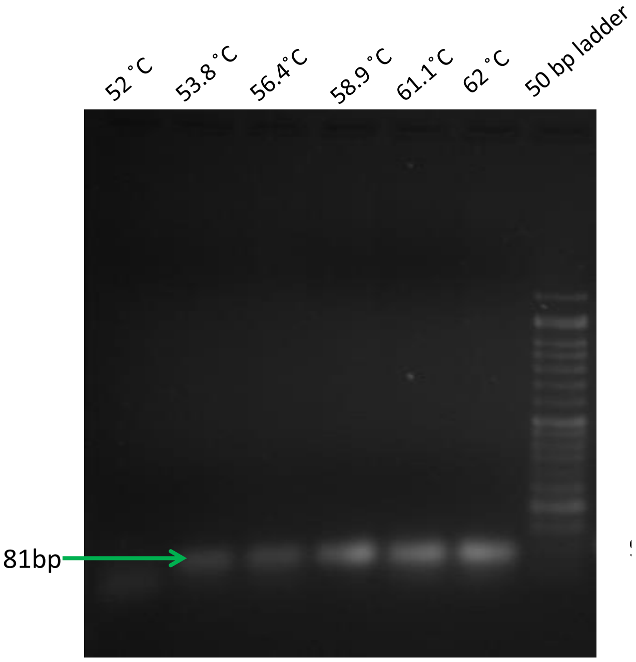


Figure 17: Standardization of annealing temperature of Ca_ARF14 Primer on *Capsicum annuum*. PCR reaction was performed with gDNA from *Capsicum annuum* for 30 cycles at various temperatures ranging from 52˚C to 62˚C. The PCR product were run on 1.2% agarose gel and visualized under ALPHA IMAGE GEL DOCUMENTATION SYSTEM as shown in figure.

g. Standardization of Ca_ARF15 primer

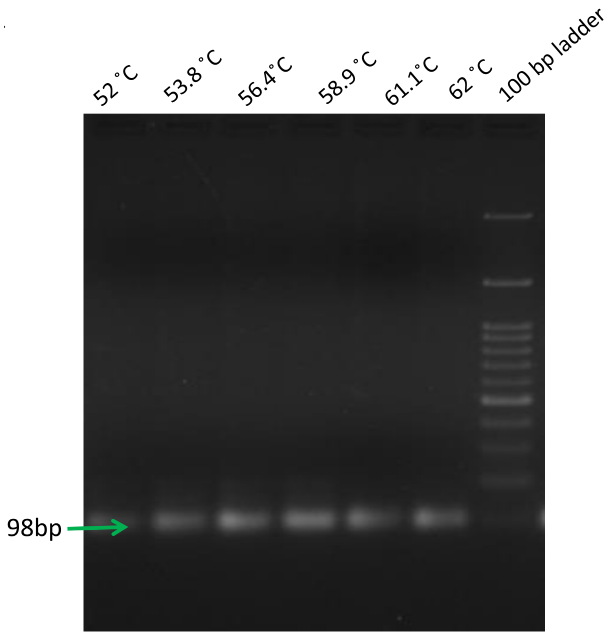


Figure 18: Standardization of annealing temperature of Ca_ARF15 Primer on *Capsicum annuum*. PCR reaction was performed with gDNA from *Capsicum annuum* for 30 cycles at various temperatures ranging from 52˚C to 62˚C. The PCR product were run on 1.2% agarose gel and visualized under ALPHA IMAGE GEL DOCUMENTATION SYSTEM as shown in figure.

h. Standardization of Ca_ARF18 primer

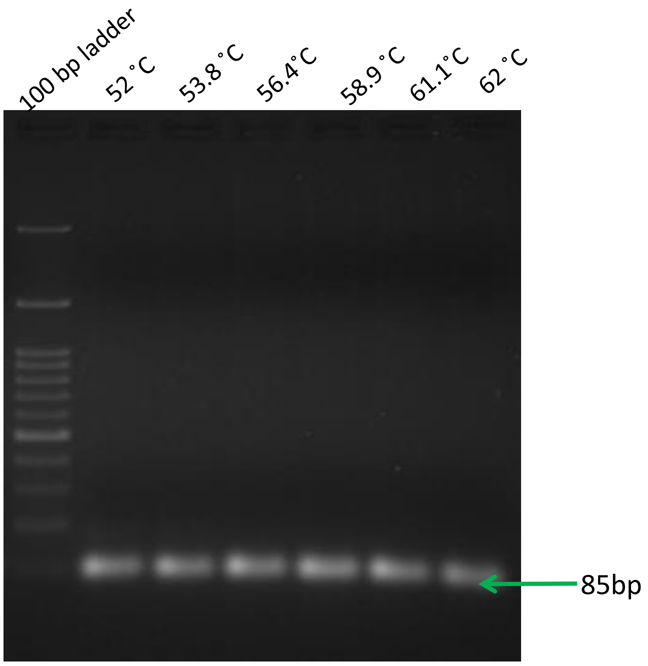


Figure 19: Standardization of annealing temperature of Ca_ARF18 Primer on *Capsicum annuum*. PCR reaction was performed with gDNA from *Capsicum annuum* for 30 cycles at various temperatures ranging from 52˚C to 62˚C. The PCR product were run on 1.2% agarose gel and visualized under ALPHA IMAGE GEL DOCUMENTATION SYSTEM as shown in figure.

i. Standardization of Ca_ARF19 primer

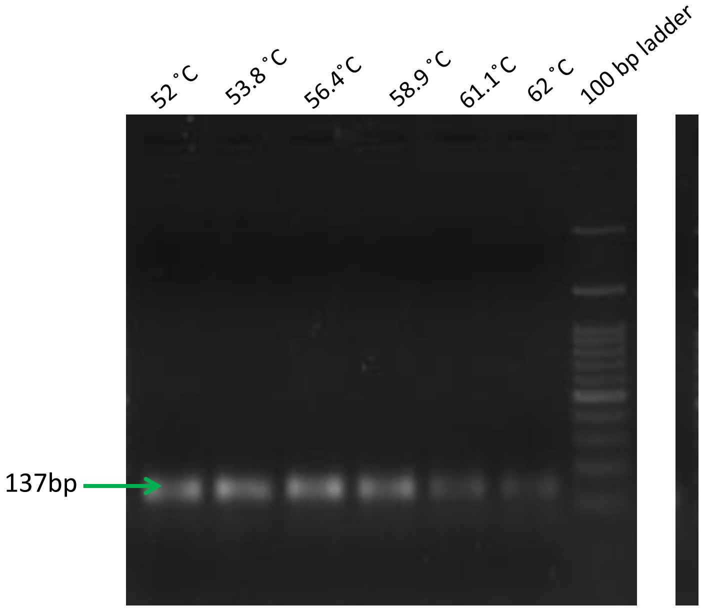


Figure 20: Standardization of annealing temperature of Ca_ARF19 Primer on *Capsicum annuum*. PCR reaction was performed with gDNA from *Capsicum annuum* for 30 cycles at various temperatures ranging from 52˚C to 62˚C. The PCR product were run on 1.2% agarose gel and visualized under ALPHA IMAGE GEL DOCUMENTATION SYSTEM as shown in figure.

j. Standardization of Ca_ARF23 primer

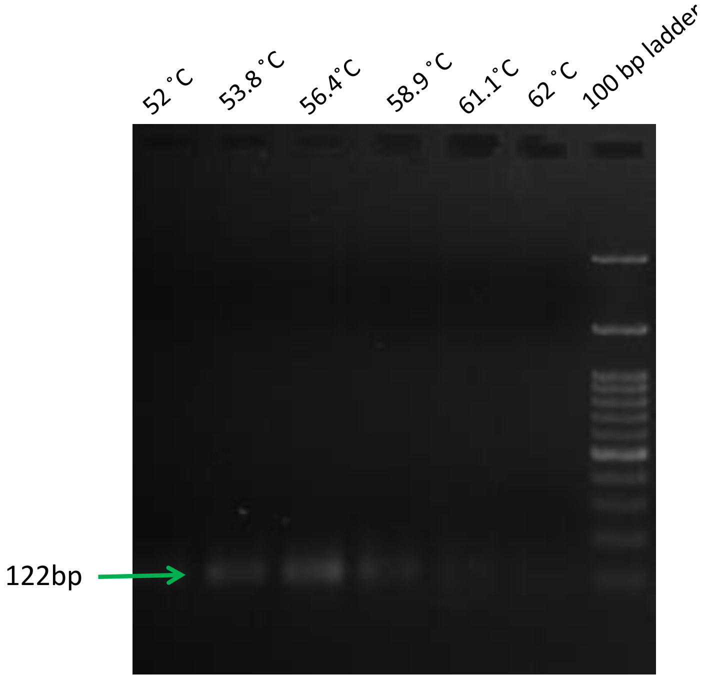


Figure 21: Standardization of annealing temperature of Ca_ARF23 Primer on *Capsicum annuum*. PCR reaction was performed with gDNA from *Capsicum annuum* for 30 cycles at various temperatures ranging from 52˚C to 62˚C. The PCR product were run on 1.2% agarose gel and visualized under ALPHA IMAGE GEL DOCUMENTATION SYSTEM as shown in figure.

k. Standardization of Ca_ARF24 primer

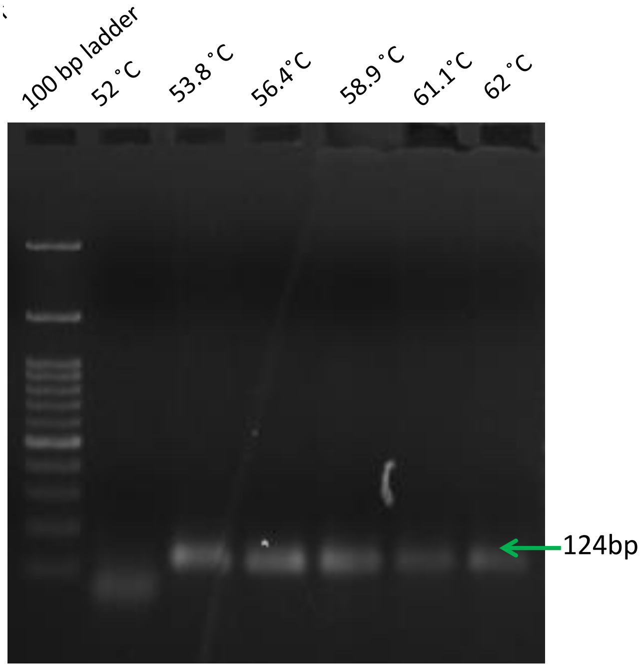


Figure 22: Standardization of annealing temperature of Ca_ARF24 Primer on *Capsicum annuum*. PCR reaction was performed with gDNA from *Capsicum annuum* for 30 cycles at various temperatures ranging from 52˚C to 62˚C. The PCR product were run on 1.2% agarose gel and visualized under ALPHA IMAGE GEL DOCUMENTATION SYSTEM as shown in figure.

l. Standardization of Ca_ARF25 primer

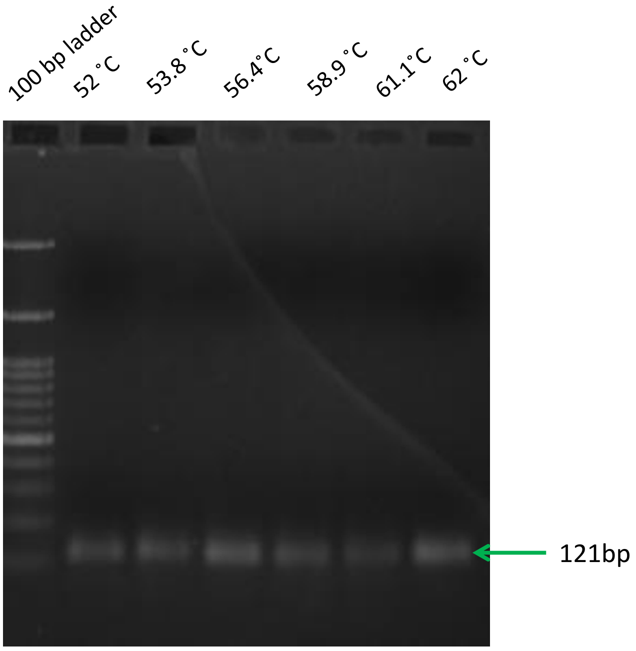


Figure 23: Standardization of annealing temperature of Ca_ARF25 Primer on *Capsicum annuum*. PCR reaction was performed with gDNA from *Capsicum annuum* for 30 cycles at various temperatures ranging from 52˚C to 62˚C. The PCR product were run on 1.2% agarose gel and visualized under ALPHA IMAGE GEL DOCUMENTATION SYSTEM as shown in figure.

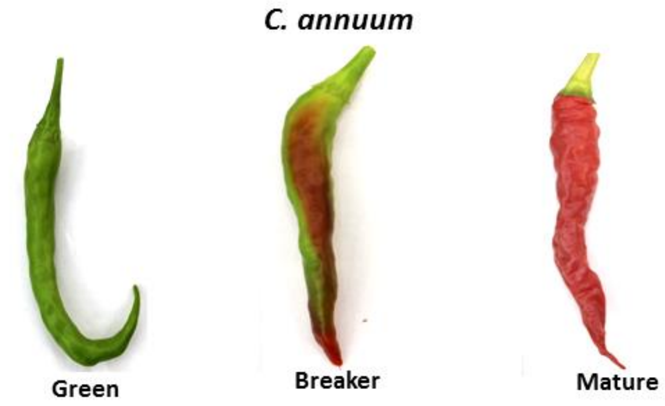
After agarose gel electrophoresis of temperature gradient PCR products, the following temperatures are standardized for different primers for 30cycle of PCR.

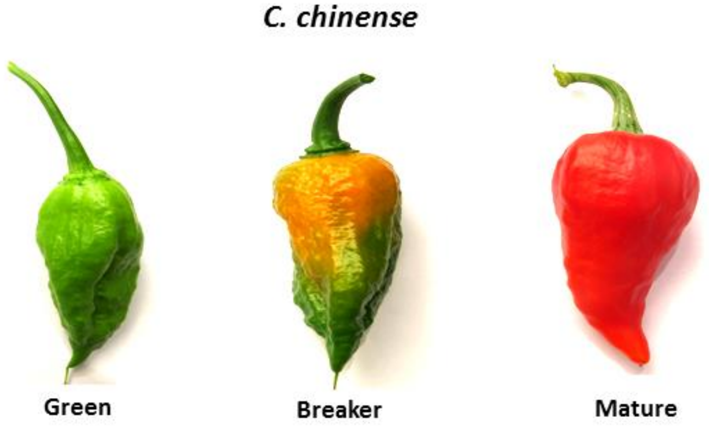
Table 9: Standardized temperature of PCR for different Ca_ARF primers

S.N.	Name of Primer	Standardized temperature (˚C)
1.	Ca_ARF3	56.4
2.	Ca_ARF4	58.9
3.	Ca_ARF6	61.1
4.	Ca_ARF9	58.9
5.	Ca_ARF12	58.9
6.	Ca_ARF14	61.1
7.	Ca_ARF15	58.9
8.	Ca_ARF18	58.9
9.	Ca_ARF19	56.4
10.	Ca_ARF23	56.4
11.	Ca_ARF24	56.4
12.	Ca_ARF25	56.4

4.6 Isolation of Ribonucleic Acid (RNA) and its quantification

Before isolating the RNA, the samples were collected and stored immediately in Liquid nitrogen as soon as possible. The collected samples were as follows:





RNA from fruit samples of different stages of *Capsicum annuum* L. ISc. Jammu and *Capsicum chinnensce* L. Peren were isolated using the TRIZOL reagent. The concentration and purity of the isolated RNA, calculated by Nano-drop were listed as following,

Table 10: Nanodrop readings of RNA

RNA sample		Concentration (ng/µl)	OD_{260}/OD_{280}	OD_{260}/OD_{230}
Capsicum species	Fruit stage			
Capsicum annuum L. ISc. Jammu	Green	3176.6	2.00	1.26
	Breaker	1389.5	1.85	0.47
	Mature	610.4	1.96	0.80
Capsicum chinnensce L. Peren	Green	1274.9	2.07	1.89
	Breaker	2561.1	2.08	2.12
	Mature	2794.6	2.09	2.04

From the readings noted in table 4, the isolated RNA has good amount from 50-100mg of tissue samples. The highest concentration of RNA was obtained from green fruit of *C.*

annuum L. Isc. Jammu which was nearly about 3200 ng and lowest concentration was also from *C. annuum* L. Isc. Jammu fruits mature stage.

The absorbance ratio of 260/280 which was in between 1.85 to 2.09 shows the RNA was of good quality. In similar way, 260/230 ratio is used as secondary measure for assessing purity of RNA. 260/230 ratio values lower than 2 indicates contaminations of phenol in samples.

4.6.1 Amplification of cDNA by PCR

The cDNA was synthesized from the RNA samples of different *Capsicum* samples. All cDNA was amplified using a constitutive gene primer EF1-α and it was found that cDNA was amplified properly. It was confirmed by 1.2% agarose gel electrophoresis of the PCR products in TBE buffer. The amplicon size of this primer is above hundred and below 200 which is shown in figure below.

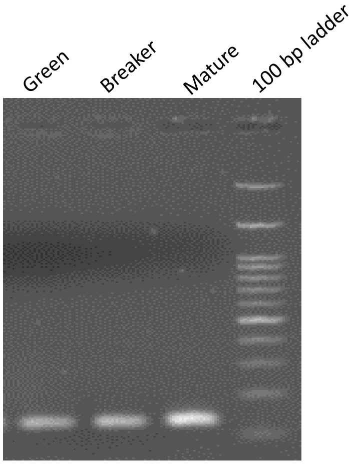


Figure 24: EF1- α primer amplified on cDNA of ISC jammu (*C. annuum*) for cDNA synthesis confirmation.

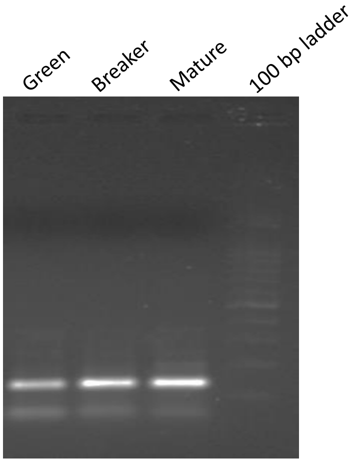


Figure 25: EF1- α primer amplified on cDNA of Peren (*Capsicum chinense)* for cDNA synthesis confirmation.

4.6.2 Expression analysis Ca_ARF genes in different *Capsicum* fruits

The expression of ARF genes are analyzed by amplifying the cDNA of different *Capsicum* fruits isolated from 3 different staged fruit tissues.

4.6.2.1 Amplification of cDNA using Ca_ARF3 primer

Semi-quantitative PCR has performed for expression analysis of Ca_ARF3 gene in different developmental stages of fruit in *C. annuum* and *C. chinense*. We observed that the expression of Ca_ARF3 gene was high at mature stage of fruit than the breaker and early stage. Similar pattern were observed in C. *chinense* but Ca_ARF3 gene expression in *C. chinense* as compare to *C. annuum*. This result shows that Ca_ARF3 gene have a role in later stages of fruit development.

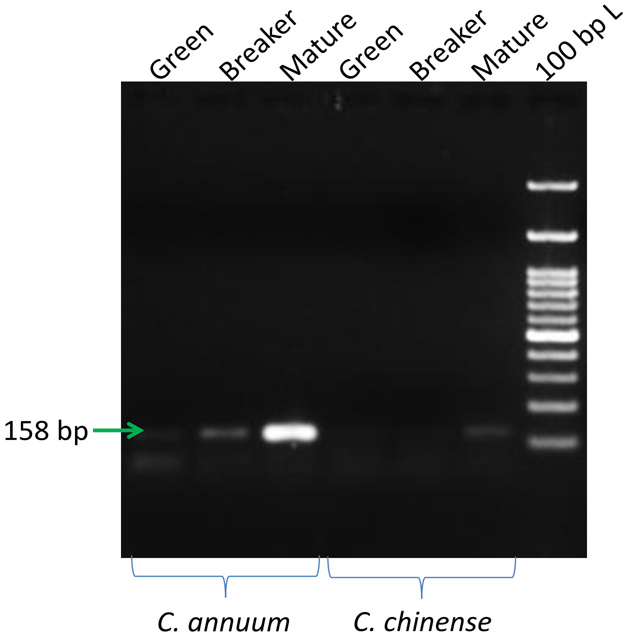


Figure 26: Semi-quantitative PCR products on 1.2% agarose gel visualized under UV captured by gel doc system

In silico analysis of Ca_ARF3 gene using *Capsicum* transcriptome shows more or less similar results, this will gives more reliability to our result.

4.6.2.2 Amplification of cDNA using Ca_ARF4 primer

Another primer Ca_ARF4 corresponding to auxin response factor 5-like expression from semi-quantitative PCR was found to be highly expressed in the mature fruit of *C. annuum.* But the expression of these genes was very low in early and breaker stage of *C. annuum* and very low expression are observed in all fruit stage of *C. chinense*.

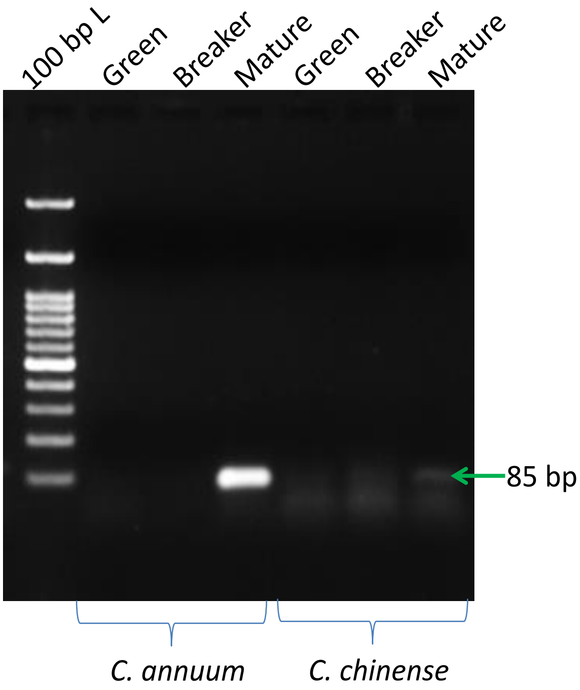


Figure 27: Semi-quantitative PCR products on 1.2% agarose gel visualized under UV captured by gel doc system

4.6.2.3 Amplification of cDNA using Ca_ARF6 primer

In our experiment Ca_ARF6 corresponds to auxin response factor 6-like gene. This gene expression is detected in both *C. annuum* and *C. chinense* but their expression seems to be very low in all fruit stages.

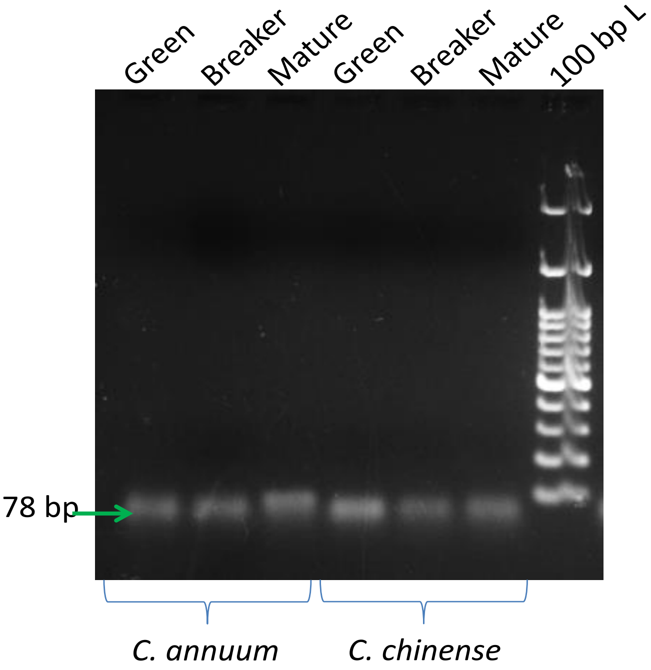


Figure 28: Semi-quantitative PCR products on 1.2% agarose gel visualized under UV captured by gel doc system

4.6.2.4 Amplification of cDNA using Ca_ARF9 primer

Ca_ARF9 primer is designed to amplify the auxin response factor 2-like gene in *Capsicum*. Semi-quantitative PCR analysis for this gene shows that it is expressed highly in mature stage in both *Capsicum* species. The expression in breaker level seems to be a bit lower than mature one in *C. annuum* and very low expression in green stage. In *C. chinense* expression is found to be very low in green and breaker stage.

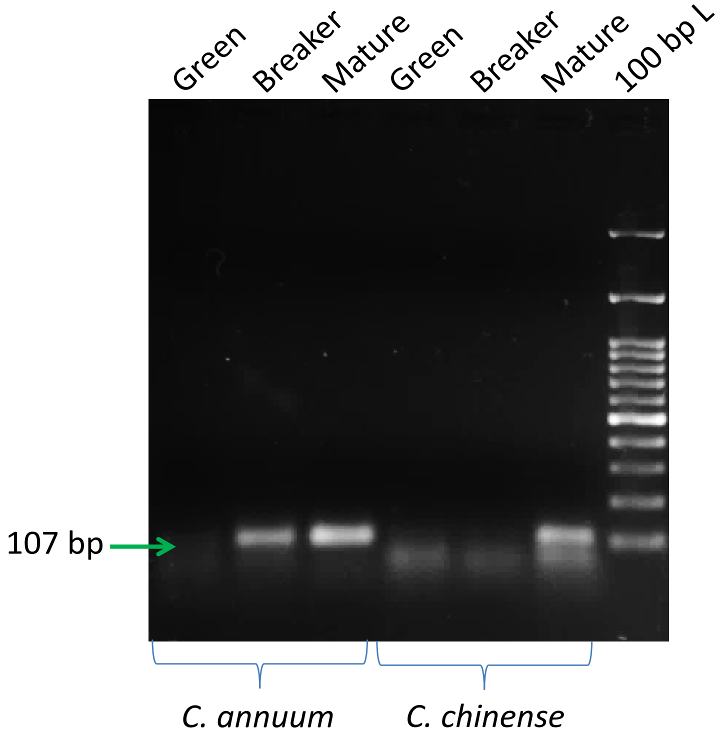


Figure 29: Semi-quantitative PCR products on 1.2% agarose gel visualized under UV captured by gel doc system

4.6.2.5 Amplification of cDNA using Ca_ARF12 primer

Primer Ca_ARF12 corresponds to gene auxin response factor 5. This gene expression pattern is found to somewhat different from the other genes for *C. annuum* as it is highly expressed in the breaker stage and mature stage. In *C. chinense* fruits it is expressed in mature stage and no expressions in other 2 earlier stages.

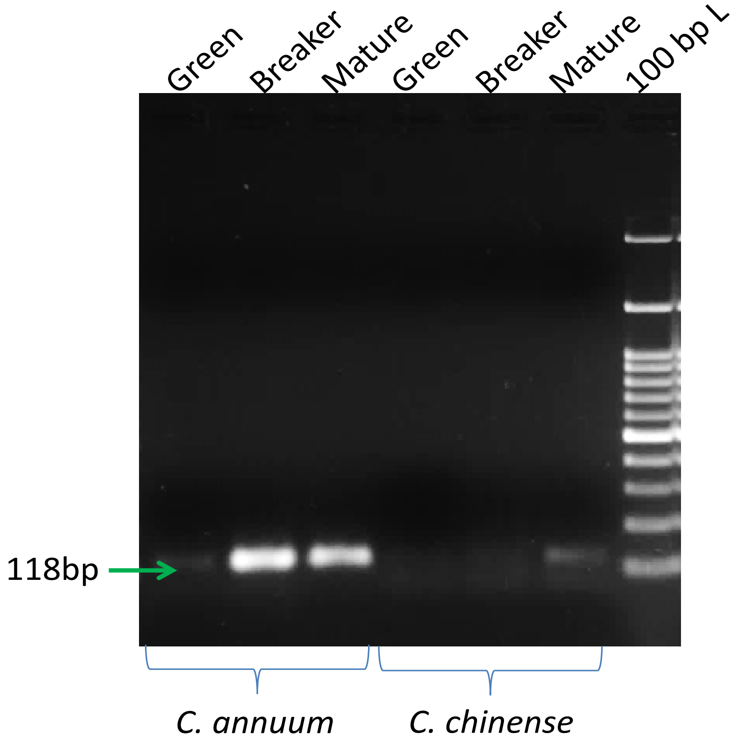


Figure 30: Semi-quantitative PCR products on 1.2% agarose gel visualized under UV captured by gel doc system

4.6.2.6 Amplification of cDNA using Ca_ARF14 primer

The expression of gene auxin response factor 8-like amplified using primer Ca_ARF14 quantified by semi-quantitative PCR shows that it is expressed moderately in all 3 stages of *C. chinense*. Their expression in *C. annuum* is found to be high in breaker fruit as compared to green and mature fruits.

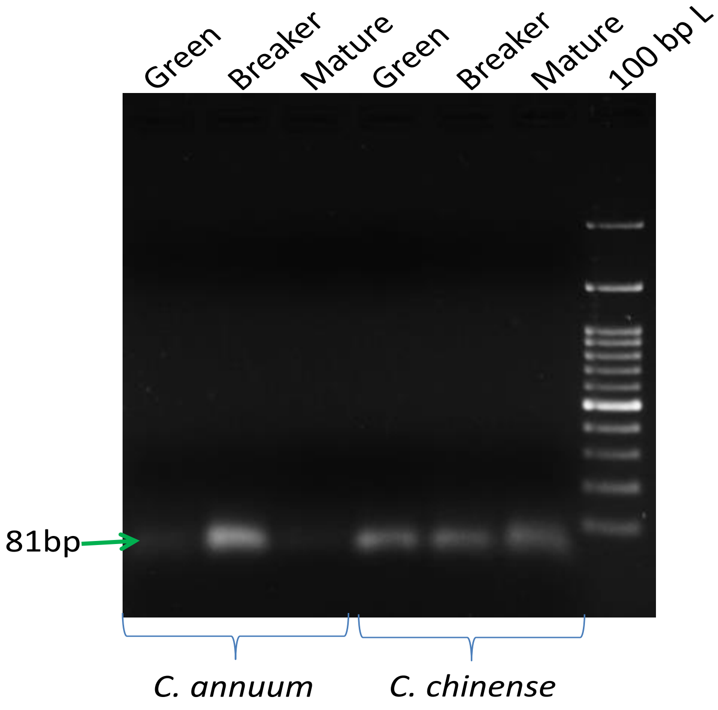


Figure 31: Semi-quantitative PCR products on 1.2% agarose gel visualized under UV captured by gel doc system

4.6.2.7 Amplification of cDNA using Ca_ARF15 primer

Semi-quantitative PCR of Ca_ARF15 primer amplified gene is noted to be expressed highly in breaker and mature stage of *C. annuum* and very low expression in green fruits. But in *C. chinense,* shows very down expression in all stages.

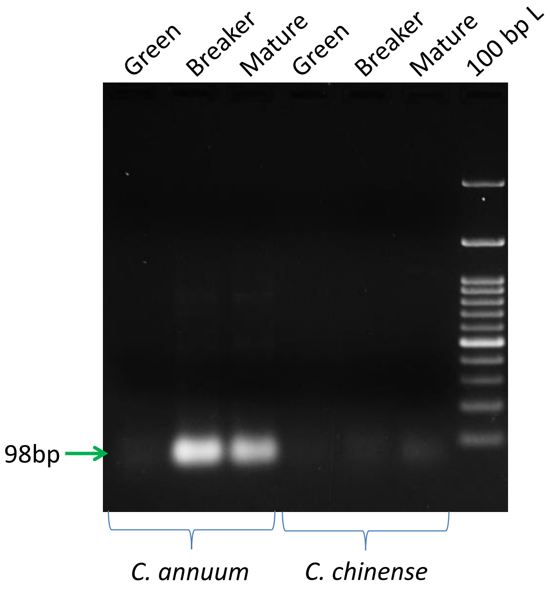


Figure 32: Semi-quantitative PCR products on 1.2% agarose gel visualized under UV captured by gel doc system

4.6.2.8 Amplification of cDNA using Ca_ARF18 primer

Ca_ARF18 primer amplification of the cDNA by semi-quantitative PCR of all 3 stages of 2 different *Capsicum species* reveals that it is expressed in all 3 stages of *C. annuum* and in mature its expression is high. But in *C. chinense* expression is observed very low amount.

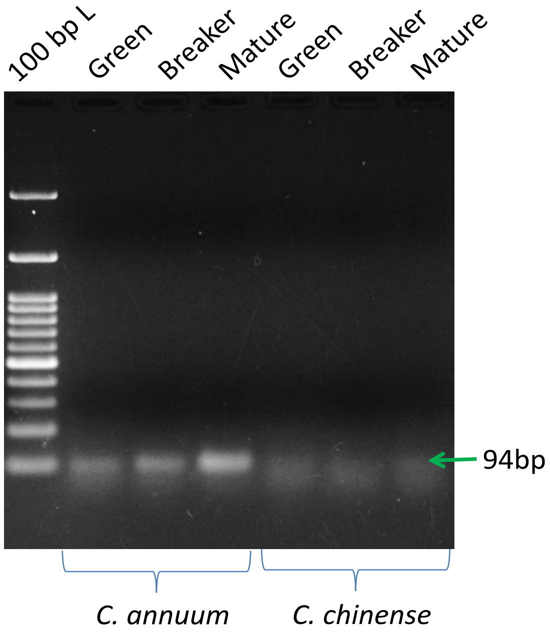


Figure 33: Semi-quantitative PCR products on 1.2% agarose gel visualized under UV captured by gel doc system

4.6.2.9 Amplification of cDNA using Ca_ARF19 primer

Ca_ARF19 amplifies 137 bp long genes in *Capsicum*. The expression of this gene is observed to be very less in breaker and mature stage of *C. annuum* and no any detectable expression was observed in all stages of *C. chinense*.

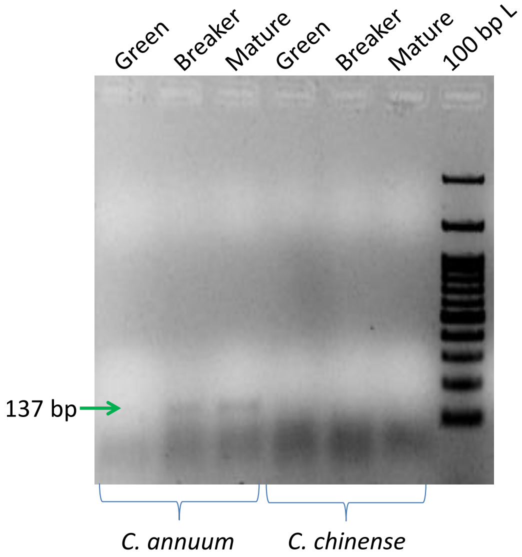


Figure 34: Semi-quantitative PCR products on 1.2% agarose gel visualized under UV captured by gel doc system

4.6.2.10 Amplification of cDNA using Ca_ARF23 primer

Similar expression pattern in both species of *Capsicum* was observed in auxin response factor 1-like gene which is amplified using the primer Ca_ARF23. In both species lowest expression is observed in mature and the expression levels rises along with the developmental stages.

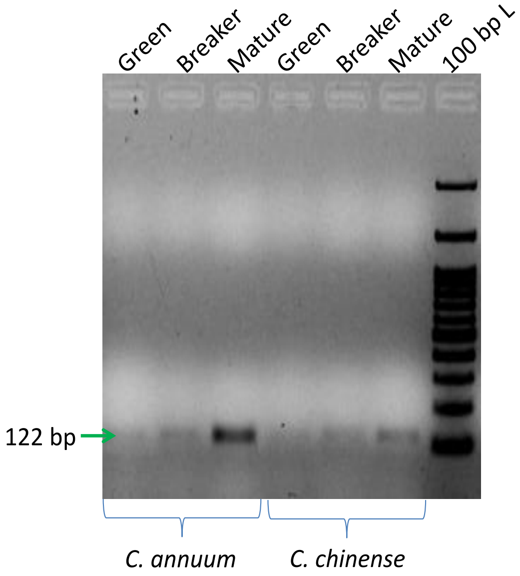


Figure 35: Semi-quantitative PCR products on 1.2% agarose gel visualized under UV captured by gel doc system

4.6.2.11 Amplification of cDNA using Ca_ARF24 primer

Auxin factor 18-like gene amplified by primer Ca_ARF24 shows its expression only in C. *annuum*. In green and mature stages it is expressed highly but very low in breaker stage; whereas in *C. chinense,* the expression is non-noticeable.

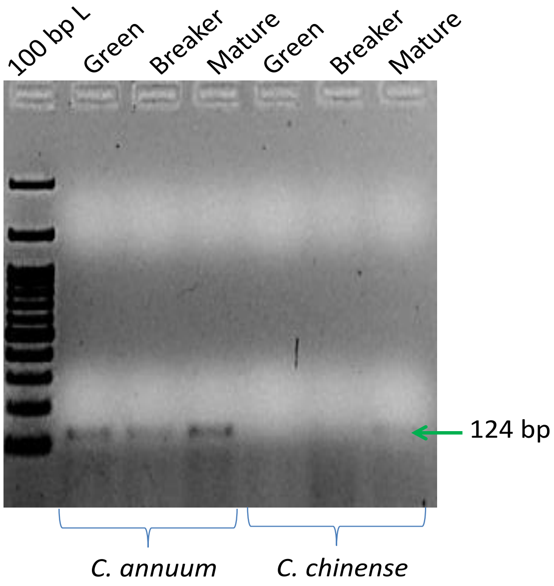


Figure 36: Semi-quantitative PCR products on 1.2% agarose gel visualized under UV captured by gel doc system

4.6.2.12 Amplification of cDNA using Ca_ARF25 primer

The amplification done by using the primer Ca_ARF25 is not observed in both *C. annuum* and *C. chinense*. As the amplicon size of this primer is 121 bp, no any amplified band is observed in the gel under UV.

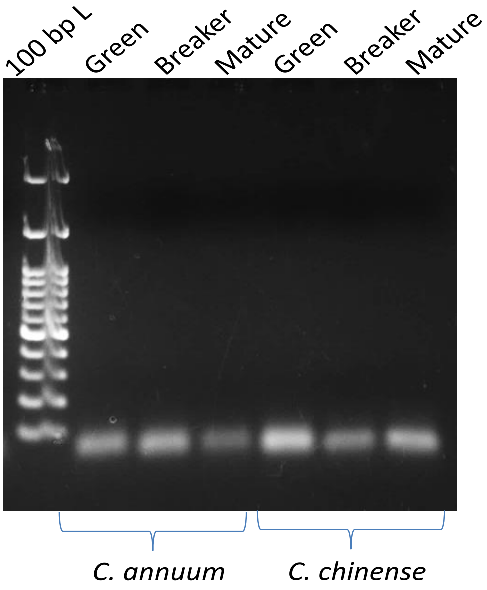


Figure 37: Semi-quantitative PCR products on 1.2% agarose gel visualized under UV captured by gel doc system

Table 11: Summary of expression analysis of Ca_ARF genes in different *Capsicum* fruits

ARF name	*C. annuum*			*C. chinense*		
	Green	Breaker	Mature	Green	Breaker	Mature
ARF 3	Low	Moderate	High	Nil	Nil	Low
ARF 4	Nil	Very low	High	Nil	Very low	Low
ARF 6	Moderate	Moderate	Moderate	Moderate	Low	Low
ARF 9	Very low	High	High	Very low	Very low	High
ARF 12	Very low	High	High	Nil	Nil	Very low
ARF 14	Very low	High	Nil	Moderate	Moderate	Moderate
ARF 15	Very low	High	High	Very low	Very low	Very low
ARF 18	Very low	Low	High	Very low	Nil	Nil
ARF 19	Nil	Very low	Very low	Nil	Nil	Nil
ARF 23	Very low	Low	High	Nil	Very low	Moderate
ARF 24	Very low	Very low	Moderate	Nil	Nil	Nil
ARF 25	Nil	Nil	Nil	Nil	Nil	Nil

This table 11 summarizes the expression level of different ARF genes. The genes which do not expressed are termed as nil. Most of ARF genes are highly expressed in the mature stage of the fruit.

4.6.3 Quantitative Real Time Polymerase Chain Reaction (qRT-PCR) for expression analysis of ARF

Real time PCR is a technique used for monitoring the progress of PCR in real time. By using this technique we can quantify even a very small amount of the PCR product and is considered as more reliable method as it provides more specificity and high sensitivity. RT-PCR results for Ca_ARF3 indicates the expression of this gene decreases significantly from green, breaker to mature in *C. annuum*, whereas in *C. chinense* the expression level is similar. While doing comparison of the expression of the stage in between the species, the expression is also found to significant. In *C. chinense,* the expression pattern is also similar with *C. annuum* in fruit stages. The results of Ca_ARF3 show that it is expressed during the early fruit development. This result exactly matched with the *in silico* analysis of laboratory as shown in heat map in figure 11.

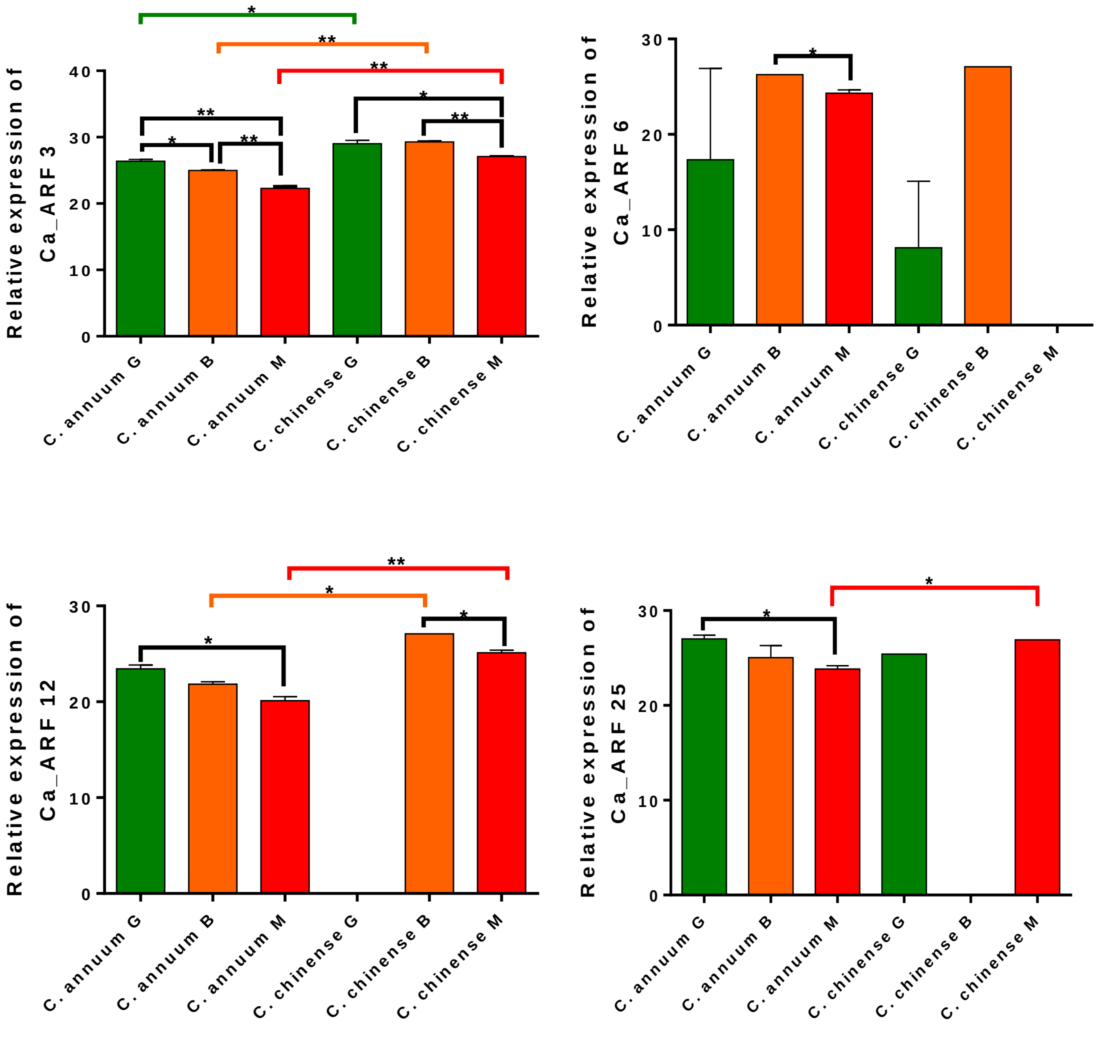


Figure 38: qRT-PCR analyses for relative expression of ARF genes in *C. annuum* and *C. chinense*. The expression analysis was performed with fruits of green, breaker and mature stage; **P<0.05, **P<0.01, ***P<0.001.*

The expression of Ca_ARF6 expression level at green fruit stage is found be lower as compared to breaker and mature stage in *C. annuum*. Again in *C. chinense* the lower level of expression is observed as compared to breaker stage but there is no expression in mature stage. Ca_ARF 12 also has expressed in all stages of *C.annuum* and decreases expressing the genes along with increase in its developmental stage. Then in *C. chinense* significant results is obtained for breaker and mature stage in comparison with the *C. annuum*. In green no expression is obtained in case of *C. chinense.* Graph analysis of Ca_ARF25 also shows decreasing pattern of expression in *C. annuum* from green, breaker to mature continuously with significance p value. In *C. chinense* the relative expreesion is observed only in green and mature leaving breaker. But the expression seems to be non-significant.

CHAPTER 5
SUMMARY

Genome-wide analysis of *Capsicum* ARF genes resulted in identification of 28 ARF genes distributed in 12 chromosomes. Genome-wide analysis of ARF genes would assist in determining the mechanisms of chili growth and development.

We identified 28 ARF genes in *C. annuum* which is higher than the number determined by previous studies (Yu, Zhan, Feng, Huang, and Sun, 2017; Zhang, Cao, Dong, and Shang, 2017) which were found to be 24 and 19 genes respective as mentioned citation. The phylogenetic tree also helped to elucidate the phylogenetic relationships between relationships between the ARF members of rice, tomato, potato, Arabidopsis and pepper. From the phylogenetic analysis, the ARF genes of *Capsicum* are only clustered with the ARF genes of tomato and potato. This may be due to similarity in function and structure of pepper ARF with tomato and potato which are members from the same family. The chromosomal map identifies 4 pseudogenes which are unplaced on scaffold.

Simple sequence repeats (SSR) are also known as microsatellite, widely distributed in genome. Due to their high reproducibility and co dominant inheritance, and extensive genome coverage makes them a power full tool for plant genetic studies (H., and K., 2002). In this study, using *Capsicum* transcription data we have identified 768, 167, 29, 13 and 6 dinucleotides, trinucleotides, tetranucleotides, pentanucleotides and hexanucleotides SSRs respectively in ARF genes of *Capsicum* which were distributed all along the 12 chromosomes. This is a preliminary analysis of SSR marker identification in *Capsicum* for ARF genes. We believe that validation of these SSR markers will help genetic study of ARF genes in *Capsicum*.

The expression of the ARF genes has been studied under normal condition without any additional treatment to plant. But the comparison is done for 2 distinct species of pepper i.e. *C. annuum* and *C. chinense.* Their fruit morphology is different and hence chosen to study the expression pattern of the ARF genes. Comparison of the *in silico* data with the laboratory generated data helps to validate the expression pattern of the ARF genes.

CHAPTER 6
CONCLUSION

In our study we analyzed ARF genes by using different bioinformatics tools. We conclude that expression of ARF genes from in-silico analysis matches exactly with Ca_ARF3 gene but the expression of other genes didn’t matched properly. This might be because of experimental error like pipetting or the experimental conditions.

Although more experimental validation are needed, the preliminary results of our work suggests that, the ARF genes are involved in fruit elongation rather than fruit size as they are expressed in *C. annuum*. Hence it is concluded that, ARF roles are involved in determining fruit length.

Apart from this, the differences may be due to difference in sample i.e. cultivar chosen for generating transcriptome *in silico* data and cultivar chosen for experiment.

REFERENCES

Anand, P., and Bley, K. (2011). Topical capsaicin for pain management: therapeutic potential and mechanisms of action of the new high-concentration capsaicin 8% patch. *British Journal of Anaesthesia*, *107*(4), 490–502. https://doi.org/10.1093/bja/aer260

Aza-González, C., Núñez-Palenius, H. G., and Ochoa-Alejo, N. (2011). Molecular biology of capsaicinoid biosynthesis in chili pepper (Capsicum spp.). *Plant Cell Reports*, *30*(5), 695–706. https://doi.org/10.1007/s00299-010-0968-8

Bagchi, G. D., and Srivastava, G. N. (2003). SPICES AND FLAVORING (FLAVOURING) CROPS | Fruits and Seeds. In *Encyclopedia of Food Sciences and Nutrition* (pp. 5465–5477). Elsevier. https://doi.org/10.1016/B0-12-227055-X/01123-8

Boer, D. R., Freire-Rios, A., van den Berg, W. A. M., Saaki, T., Manfield, I. W., Kepinski, S., … Coll, M. (2014). Structural Basis for DNA Binding Specificity by the Auxin-Dependent ARF Transcription Factors. *Cell*, *156*(3), 577–589. https://doi.org/10.1016/J.CELL.2013.12.027

Bosland, P. W., Coon, D., and Cooke, P. H. (2015). Novel formation of ectopic (nonplacental) capsaicinoid secreting vesicles on fruit walls explains the morphological mechanism for super-hot chile peppers. *Journal of the American Society for Horticultural Science*, *140*(3), 253–256.

Carey, N. S., and Krogan, N. T. (2017). The role of AUXIN RESPONSE FACTORs in the development and differential growth of inflorescence stems, *2324*(March). https://doi.org/10.1080/15592324.2017.1307492

Carrizo García, C., Barfuss, M. H. J., Sehr, E. M., Barboza, G. E., Samuel, R., Moscone, E. A., and Ehrendorfer, F. (2016). Phylogenetic relationships, diversification and expansion of chili peppers (Capsicum , Solanaceae). *Annals of Botany*, *118*(1), 35–51. https://doi.org/10.1093/aob/mcw079

Chandler, J. W. (2016a). Auxin response factors, 1014–1028. https://doi.org/10.1111/pce.12662

Chandler, J. W. (2016b). Auxin response factors. *Plant, Cell and Environment*, *39*(5), 1014–1028. https://doi.org/10.1111/pce.12662

Cheng, Y., JalalAhammed, G., Yu, J., Yao, Z., Ruan, M., Ye, Q., … Wan, H. (2016). Putative WRKYs associated with regulation of fruit ripening revealed by detailed expression analysis of the WRKY gene family in pepper. *Scientific Reports*, *6*(1), 39000. https://doi.org/10.1038/srep39000

Cichewicz, R. H., and Thorpe, P. A. (1996). The antimicrobial properties of chile peppers (Capsicum species) and their uses in Mayan medicine. *Journal of Ethnopharmacology*, *52*, 61–70.

CRICK, F. (1970). Central Dogma of Molecular Biology. *Nature*, *227*(5258), 561–563. https://doi.org/10.1038/227561a0

Diao, W.-P., Snyder, J. C., Wang, S.-B., Liu, J.-B., Pan, B.-G., Guo, G.-J., and Wei, G. (2016).

Genome-Wide Identification and Expression Analysis of WRKY Gene Family in Capsicum annuum L. *Frontiers in Plant Science*, *7*, 211. https://doi.org/10.3389/fpls.2016.00211

Dragon, F., Gallagher, J. E. G., Compagnone-Post, P. A., Mitchell, B. M., Porwancher, K. A., Wehner, K. A., … Baserga, S. J. (2002). A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis. *Nature*, *417*(6892), 967–970. https://doi.org/10.1038/nature00769

FAO data - dataset-data-filter - Commodity Balances, Crops, Primary Equivalent (FAOSTAT). (2016). Retrieved March 29, 2018, from http://ref.data.fao.org/web/guest/dataset-data-filter?entryId=ac2f9731-8464-4589-89fd-7e8242009929andtab=dataandtype=DimensionmemberanduuidResource=e24a444a-94c5-4055-af20-618ede84b242

Franco-Zorrilla, J. M., López-Vidriero, I., Carrasco, J. L., Godoy, M., Vera, P., and Solano, R. (2014). DNA-binding specificities of plant transcription factors and their potential to define target genes. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(6), 2367–2372. https://doi.org/10.1073/pnas.1316278111

Gonzalez, D. H. (2016). Introduction to Transcription Factor Structure and Function. In *Plant Transcription Factors* (pp. 3–11). Elsevier. https://doi.org/10.1016/B978-0-12-800854-6.00001-4

Guilfoyle, T. J., and Hagen, G. (2007a). Auxin response factors. *Current Opinion in Plant Biology*, *10*, 453–460. https://doi.org/10.1016/j.pbi.2007.08.014

Guilfoyle, T. J., and Hagen, G. (2007b). Auxin response factors. *Current Opinion in Plant Biology*, *10*(5), 453–460. https://doi.org/10.1016/j.pbi.2007.08.014

H., E., H., S., and K., K. (2002). SSR analysis of genetic diversity among maize inbred lines adapted to cold regions of Japan. *TAG Theoretical and Applied Genetics*, *104*(8), 1270–1277. https://doi.org/10.1007/s00122-001-0857-1

Hao, Y., Hu, G., Breitel, D., Liu, M., Mila, I., Frasse, P., … Zouine, M. (2015). Auxin Response Factor SlARF2 Is an Essential Component of the Regulatory Mechanism Controlling Fruit Ripening in Tomato. *PLOS Genetics*, *11*(12), e1005649. https://doi.org/10.1371/journal.pgen.1005649

Hegde, V., Partap, P. S., and Yadav, R. C. (2017). Plant Regeneration from Hypocotyl Explants in Capsicum (Capsicum annuum L.). *International Journal of Current Microbiology and Applied Sciences*, *6*(7), 545–557. https://doi.org/10.20546/ijcmas.2017.607.066

John A. Cooper. (n.d.). Transcription factor | biology | Britannica.com. Retrieved April 26, 2018, from https://www.britannica.com/science/transcription-factor

Joshi, R., Wani, S. H., Singh, B., Bohra, A., Dar, Z. A., Lone, A. A., … Singla-Pareek, S. L. (2016). Transcription Factors and Plants Response to Drought Stress: Current Understanding and Future Directions. *Frontiers in Plant Science*, *7*, 1029. https://doi.org/10.3389/fpls.2016.01029

Karlova, R., Chapman, N., David, K., Angenent, G. C., Seymour, G. B., and de Maagd, R. A. (2014). Transcriptional control of fleshy fruit development and ripening. *Journal of Experimental Botany*, *65*(16), 4527–4541. https://doi.org/10.1093/jxb/eru316

Kim, S., Park, M., Yeom, S.-I., Kim, Y.-M., Lee, J. M., Lee, H.-A., … Choi, D. (2014). Genome sequence of the hot pepper provides insights into the evolution of pungency in Capsicum species. *Nature Genetics*, *46*(3), 270–278. https://doi.org/10.1038/ng.2877

Krogan, N. T., Berleth, T., Marcos, D., and Weiner, A. I. (2016). The auxin response factor MONOPTEROS controls meristem function and organogenesis in both the shoot and root through the direct regulation of PIN genes. *New Phytologist*, *212*, 42–50. https://doi.org/10.1111/nph.14107

Latchman, D. S. (1993). Transcription factors: an overview Function of transcription factors. *Int. J. Exp. Path*, *74*, 417–422.

Letunic, I., and Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Research*, *44*(W1), W242-5. https://doi.org/10.1093/nar/gkw290

Li, S.-B., Xie, Z.-Z., Hu, C.-G., and Zhang, J.-Z. (2016a). A Review of Auxin Response Factors (ARFs) in Plants. *Frontiers in Plant Science*, *7*, 47. https://doi.org/10.3389/fpls.2016.00047

Li, S.-B., Xie, Z.-Z., Hu, C.-G., and Zhang, J.-Z. (2016b). A Review of Auxin Response Factors (ARFs) in Plants. *Frontiers in Plant Science*, *7*, 47. https://doi.org/10.3389/fpls.2016.00047

Liu, L., White, M. J., and Macrae, T. H. (1999). Functional domains , evolution and regulation. *Eur. j. Biochem*, *257*, 247–257. https://doi.org/10.1046/j.1432-1327.1999.00349.x

Manikharda, Takahashi, M., Arakaki, M., Yonamine, K., Hashimoto, F., Takara, K., and Wada, K. (2018). Influence of Fruit Ripening on Color, Organic Acid Contents, Capsaicinoids, Aroma Compounds, and Antioxidant Capacity of Shimatogarashi (*Capsicum frutescens*). *Journal of Oleo Science*, *67*(1), 113–123. https://doi.org/10.5650/jos.ess17156

Montiel, G. (2004). Transcription Factor Networks. Pathways to the Knowledge of Root Development. *PLANT PHYSIOLOGY*, *136*(3), 3478–3485. https://doi.org/10.1104/pp.104.051029

Park, D.-S., Yoon, G.-H., Lee, H.-S., and Choi, S.-C. (2016). Capsaicin inhibits the Wnt/β-catenin signaling pathway by down-regulating PP2A. *Biochemical and Biophysical Research Communications*, *478*(1), 455–461. https://doi.org/10.1016/j.bbrc.2016.06.075

Pattison, R. J., Csukasi, F., and Catalá, C. (2014). Mechanisms regulating auxin action during fruit development. *Physiologia Plantarum*, *151*(1), 62–72. https://doi.org/10.1111/ppl.12142

Perla, V., Nimmakayala, P., Nadimi, M., Alaparthi, S., Hankins, G. R., Ebert, A. W., and

Reddy, U. K. (2016). Vitamin C and reducing sugars in the world collection of Capsicum baccatum L . genotypes. *Food Chemistry*, *202*, 189–198. https://doi.org/10.1016/j.foodchem.2016.01.135

Perry, L., Dickau, R., Zarrillo, S., Holst, I., Pearsall, D. M., Piperno, D. R., … Zeidler, J. A. (2007). Starch Fossils and the Domestication and Dispersal of Chili Peppers (Capsicum spp. L.) in the Americas. *Science*, *315*(5814), 986–988. https://doi.org/10.1126/science.1136914

PlantTFDB - Plant Transcription Factor Database @ CBI, PKU. (2017). Retrieved May 8, 2018, from http://planttfdb.cbi.pku.edu.cn/index.php

Rademacher, E. H., Möller, B., Lokerse, A. S., Llavata-Peris, C. I., van den Berg, W., and Weijers, D. (2011). A cellular expression map of the Arabidopsis AUXIN RESPONSE FACTOR gene family. *The Plant Journal*, *68*(4), 597–606. https://doi.org/10.1111/j.1365-313X.2011.04710.x

Ramchiary, N., Kehie, M., Brahma, V., Kumaria, S., and Tandon, P. (2014). Application of genetics and genomics towards Capsicum translational research. *Plant Biotechnology Reports*, *8*(2), 101–123. https://doi.org/10.1007/s11816-013-0306-z

Saha, S., Karaca, M., Jenkins, J. N., Zipf, A. E., Reddy, O. U. K., and Kantety, R. V. (2003). Simple sequence repeats as useful resources to study transcribed genes of cotton. *Euphytica*, *130*(3), 355–364. https://doi.org/10.1023/A:1023077209170

Seo, E., Choi, D., and Choi. (2015). Functional studies of transcription factors involved in plant defenses in the genomics era. *Briefings in Functional Genomics*, *14*(4), 260–267. https://doi.org/10.1093/bfgp/elv011

Shan, W., Kuang, J., Chen, L., Xie, H., Peng, H., Xiao, Y., … Lu, W. (2012). Molecular characterization of banana NAC transcription factors and their interactions with ethylene signalling component EIL during fruit ripening. *Journal of Experimental Botany*, *63*(14), 5171–5187. https://doi.org/10.1093/jxb/ers178

Shiu, S.-H., Shih, M.-C., and Li, W.-H. (2005). Transcription factor families have much higher expansion rates in plants than in animals. *Plant Physiology*, *139*(1), 18–26. https://doi.org/10.1104/pp.105.065110

Simon, S., and Petrášek, J. (2011). Why plants need more than one type of auxin. *Plant Science*, *180*(3), 454–460. https://doi.org/10.1016/J.PLANTSCI.2010.12.007

Srinivasan, K. (2016). Biological Activities of Red Pepper (Capsicum annuum) and Its Pungent Principle Capsaicin: A Review. *Critical Reviews in Food Science and Nutrition*, *56*(9), 1488–1500. https://doi.org/10.1080/10408398.2013.772090

The Editors of Encyclopaedia Britannica. (2018). Paprika | spice | Britannica.com. Retrieved June 2, 2018, from https://www.britannica.com/topic/paprika#ref142483

Walker, S. J., Worst, T. J., and Vrana, K. E. (2003). Semiquantitative Real-Time PCR for Analysis of mRNA Levels. In *Drugs of Abuse* (pp. 211–228). New Jersey: Humana Press. https://doi.org/10.1385/1-59259-358-5:211

Wani, T. A., and Lattoo, S. K. (2017). Auxin response factor (GaARF) cloning and expression in relation to reproductive maturation in Grewia asiatica L. *Plant Gene*, *12*, 123–130. https://doi.org/10.1016/J.PLGENE.2017.10.001

WOODWARD, A. W., and BARTEL, B. (2005). Auxin: Regulation, Action, and Interaction. *Annals of Botany*, *95*(5), 707–735. https://doi.org/10.1093/aob/mci083

Yu, C., Zhan, Y., Feng, X., Huang, Z.-A., and Sun, C. (2017). Identification and Expression Profiling of the Auxin Response Factors in Capsicum annuum L. under Abiotic Stress and Hormone Treatments. *International Journal of Molecular Sciences*, *18*(12). https://doi.org/10.3390/ijms18122719

Yusuf, D., Butland, S. L., Swanson, M. I., Bolotin, E., Ticoll, A., Cheung, W. A., … Wasserman, W. W. (2012). The Transcription Factor Encyclopedia. *Genome Biology*, *13*(3), R24. https://doi.org/10.1186/gb-2012-13-3-r24

ZHANG, H., CAO, N., DONG, C., and SHANG, Q. (2017). Genome-wide Identification and Expression of ARF Gene Family during Adventitious Root Development in Hot Pepper (Capsicum annuum). *Horticultural Plant Journal*, *3*(4), 151–164. https://doi.org/10.1016/J.HPJ.2017.07.001

Zhang, Z.-X., Zhao, S.-N., Liu, G.-F., Huang, Z.-M., Cao, Z.-M., Cheng, S.-H., and Lin, S.-S. (2016). Discovery of putative capsaicin biosynthetic genes by RNA-Seq and digital gene expression analysis of pepper. https://doi.org/10.1038/srep34121

APPENDICES

10X Tris Borate EDTA (TBE) Buffer – 1000 mL

Tris Base	108 g
Boric acid	55 g
0.5M EDTA (pH 8)	40 mL

Final volume made up to 1000 mL by adding Milli-Q water.

1X TBE Buffer -1000 mL

10X TBE	100 mL
Milli-Q water	900 mL

Cetyl Trimethyl Ammonium Bromide (CTAB) buffer – 1000 mL

1M Tris Cl (pH 8)	100 mL
0.5M EDTA	40 mL
5M NaCl	280 mL
10% CTAB	200 mL
Milli-Q water	380 mL

Di-ethyl pyro carbonate (DEPC) water-1000 mL

DEPC	1 mL
MQ water	1000 mL

- **10X 3-(N-Morpholino)propanesulfonic acid (MOPS) Buffer-500 mL**

0.4 M MOPS	41.85 g
0.1M Sodium acetate	4.10 g
0.01M EDTA (pH 8)	10 mL (from 0.5M Stock) EDTA

pH 7 maintained by adding NaOH, final volume made to 500 mL by adding DEPC treated water and autoclaved and stored at 4˚C.

- **1X MOPS buffer- 1000 mL**

10X MOPS	100 mL
37% formaldehyde	20 mL
Autoclaved DEPC water	880 mL

- **6X DNA loading dye-100 mL**

Bromophenol blue	250 mg
Xylene cyanol FF	250 mg
Ficoll 400	15 g

Final volume maintained 100 mL by adding MQ water.

Genome-wide identification of Auxin Response Factor (ARF) genes and expression analysis in *Capsicum* fruits

Bidyakala Sunuwar[1]*, Nirala Ramchiary*[2] *and Rajani Malla*[1]
[1]*Central Department of Biotechnology, Tribhuvan University, Kathmandu, Nepal and*
[2]*School of Life Science, Jawaharlal Nehru University, New Delhi, India.*

ABSTRACT

One of the important plant transcription factors is Auxin Response Factor (ARF), which has vital role in regulation of growth and development of the plants, and stress tolerance. ARFs key regulator is the phyto-hormone Auxin. Very less investigation has been done about ARF family of pepper (*Capsicum annuum*) and their roles in biological processes. From the available updated genome database, 28 ARF genes were identified in pepper genome and 24 of them are distributed on 12 different chromosome. A phylogenetic tree was constructed depending on ARFs sequences derived from Pepper, Tomato, Arabidopsis, Potato and Rice and Simple Sequence Repeats (SSRs) were identified present in Capsicum ARF. In this study, expression of selected 12 ARF genes in fruits (developmental stage: early, breaker and mature) of two distinct *Capsicum* species (*Capsicum annuum* and *Capsicum chinense*) were analyzed in transcriptome level. Confirmation of expression analysis was done by quantitative real time PCR (qRT-PCR). The preliminary results of our experiment elucidates, ARF plays significant role in fruit elongation of pepper. This result will provide base for identifying proper function and molecular breeding studies of ARF genes in pepper in forthcoming days.

Keywords: Transcription factor, Auxin Response Factor (ARF), *Capsicum*, expression analysis, transcriptome, quantitative real time PCR (qRT-PCR)

INTRODUCTION

Genus *Capsicum*, commonly called as chili pepper, bell pepper, paprika or just *Capsicum* is a vegetable. *Capsicum* is indigenous genus of South and Central America. This crop is second most important members of the family Solanaceae after tomato, consisting more than 38 species. Among them only six species are cultivated namely *Capsicum annuum, Capsicum chinense, Capsicum frutescence, Capsicum pubescens, Capsicum baccatum and Capsicum assamicum* in different parts of world (Ramchiary et al. 2014). Chilies are used as vegetables, spices, flavorants, colorants and condiments all over the world in fresh or dried forms. Wide variations of intra as well interspecies are seen for fruit morphology, plant types and other agronomic traits in *Capsicum* species. The fruits of *Capsicums* are consumed in any three stages namely green, breaker and mature freshly and in dried red form also. One of the unique and important properties of *Capsicum* species is its pungency due to Capsaicin apart from other Solanaceae members. Pungency of *Capsicum* is limited to its fruit. Hence, detailed study of fruit development of *Capsicum* is necessary for improving its economic value. Chilies fruit are consumed all over the

Roles of different plants TFs are identified in various biological processes, from development of the plants to defense mechanism against different pathogens (Seo, Choi, and Choi 2015). In plants, organ development process is perpetual throughout its life cycle. Being sessile organisms, plants have to confront various environmental conditions in order to survive as compared to animals. Thus, their TFs are precisely regulated through different signaling cascades which control development and defense mechanism at molecular levels. According to different studies, 6-10% of the whole plant genome codes for TFs of plants(Franco-Zorrilla et al. 2014). Auxin is an important phyto-hormone which plays vital roles in plant growth and engineers many developmental processes. Understanding regulation of auxin and its interactions mechanisms can be fruitful for deciphering numerous aspects of plant growth and development. Auxin helps in determining root and shoot structure, organ patterning, vascular structure development, growth in tissue culture. Auxin in combination other hormones modulates various processes. In plant tissue culture, in vitro propagation of root and shoot development is induced by auxin in combination with cytokinin. The