# Comparative Genomics and Gene Expression Analysis for Yield and Quality Traits in Garlic


**By**

## Malik Attique Ur Rehman

**Department of Plant Genomics and Biotechnology**

**PARC Institute of Advanced Studies in Agriculture**

**National Agricultural Research Centre, Islamabad**

**Quaid-I-Azam University, Islamabad, Pakistan**

**2020**

# Comparative Genomics and Gene Expression Analysis for Yield and Quality Traits in Garlic

*A Thesis*
*Submitted to Quaid-i-Azam University, Islamabad in the Partial fulfillment of the requirements for the degree of*

**Master of Philosophy**
**In**
**Plant Genomics and Biotechnology**


**By**


## Malik Attique Ur Rehman

**Department of Plant Genomics and Biotechnology**

**PARC Institute of Advanced Studies in Agriculture**

**National Agricultural Research Centre, Islamabad**

**Quaid-I-Azam University, Islamabad, Pakistan**


**2020**

# CERTIFICATE

This thesis submitted by **Malik Attique Ur Rehman** from Plant Genomics and Biotechnology, PARC Institute of Advanced studies in Agriculture, Islamabad, is accepted in its current form. This thesis fulfills all the requirement for facilitating him with Degree of Master of Philosophy in in **Plant Genomics and Biotechnology**.

**Internal Examiner:** Dr. Muhammad Ramzan Khan
Principal Scientific Officer,
NIGAB, NARC, Islamabad

**External Examiner:** Dr. Tayyaba Yasmin
Associate Professor,
Department of Bio Science,
Comsats Institute of Information Technology (CIIT),
Islamabad

**Chairperson:** Dr. Shaukat Ali
Principal scientific Officer,
NIGAB, NARC, Islamabad

# AUTHOR'S DECLARATION

I **Malik Attique Ur Rehman** hereby declare that the data presented in this thesis "**Comparative Genomics and Gene Expression Analysis for Yield and Quality Traits in Garlic**" is generated myself from original research work during the scheduled period of study under the supervision of **Dr. Muhammad Ramzan Khan**. The results and material used in this thesis never presented anywhere else earlier.

Date: _____

**Malik Attique Ur Rehman**
Reg # 02361713003

# DEDICATION

I sincerely dedicate this lifetime achievement to my loving and very caring parents whose prays and support made my path smooth and comfortable to my goals. They are my mentor and I am proud of them.

# TABLE OF CONTENTS

# LIST OF FIGURES

## LIST OF TABLES

## LIST OF ABBREVIATIONS

| Word | Abbreviation |
|---|---|
| **ANS** | Anthocyanidin synthase gene |
| **BCI** | Bio-Resource Conservation Institute |
| **BLAST** | Basic Local Alignment Search Tool |
| **bps** | Base Pairs |
| **cDNA** | Complementary deoxyribonucleic acid |
| **cm** | Centimeter |
| **DNA** | Deoxyribonucleic Acid |
| **dNTPs** | Deoxyribonucleic Acid |
| **FT** | Flowering Locus T gene |
| **LFS** | Lachrymatory factor synthase gene |
| **MEGA** | Molecular Evolutionary Genetic Analysis |
| **mg** | Milligram |
| **ml** | Milliliter |
| **NARC** | National Agricultural Research Center |
| **NCBI** | National Centre for Biotechnology Information |
| **NIGAB** | National Institute for Genomics and Advanced Biotechnology |
| **PCR** | Polymerase Chain Reaction |
| **pH** | Power of Hydrogen Ion |
| **pm** | Pico mole |
| **qRT-PCR** | Quantitative reverse transcriptase PCR |
| **RNA** | Ribonucleic Acid |
| **rpm** | Revolution Per Minute |
| **SIR** | Sulphite Reductase gene |
| **ul** | Microliter |
| **uM** | Miro molar |

# ACKNOWLEDGEMENTS

I am grateful and like to thank **ALMIGHTY ALLAH**, The Most Merciful, the Beneficent, who gave me the courage and support to complete this study and thesis write up. Cordial gratitude to **PROPHET MUHAMMAD (PBUH)**. Who is forever a torch of guidance and knowledge for whole humanity. The whole facilities, abilities, opportunities, strength and powers are blessed by ALLAH almighty for the level of study and research to such a vibrant environment

I would like to express my sincere gratitude and pay special regard to my honorable supervisor **Dr. Muhammad Ramzan Khan**, Principal Scientific Officer, NIGAB, NARC, Islamabad for his patience, motivation, continuous support and efforts for completion of my thesis research work..

I would like to thank to my colleagues of session 2017-2019, for joys, happiness, sadness, fun and all academic activities during the whole academic session. I would like to thank all my teachers and Lab fellows specially **Mam Safeena** , **Mam Nazia** ,**Sir Aqeel** , **Mam Wajyah** , **Sahir Hameed** all were very kind and cooperative in the research matters. I would like to pay special thanks to my colleague **Syed Zaheer uddin** , **Saima Noor** , **Shelah Shoukat** and **Aliya Irum** they all encourages and support me allot in all the matter either related to education, research and social attraction .

At the end, I am extremely grateful to my **Parents, Brother** and **Sister** for providing me the moral and emotional support. Without my family support and love, I would not have finished this research work.

**Malik Attique Ur Rehman**

# ABSTRACT

Garlic (*Allium sativum*) is the second most widely cultivated Allium species grown in the world. The production of Pakistani garlic varieties is hampered by a number of factors. Among them scarcity of genetic data and functional characterization of the genes involved in garlic quality are the major constraints for improvement in breeding programmes. Analyzing the diversity of key gene families involved in plant architecture and adaptation among five different local and exotic garlic cultivars was attempted in this study. For this purpose, four cultivars namely "Local Red, NARC G1, China White and an Advance line" were selected. The Allium genome search revealed 12 key genes involved in yield and quality parameters like flavour, sulphur metabolism, bulbing and colour of the clove. Out of these 4 genes, Ft-Like (FT), Anthocyanidin Synthase (ANS), Lachrymatory Factor Synthase (LFS) and Sulphite Reductase (SIR) were selected for phylogenetic and gene expression studies. Homologous sequences of these genes were retrieved from the NCBI and Ensembl plants databases. Phylogenetic trees of FT, ANS, LFS and SIR form different plant species exhibited clustering into their corresponding clades. Realtime-RT-PCR analysis confirmed the expression of these yield related genes in both the bulb and leaves of five different cultivars of garlic (*Allium sativum*). The expression of these yield and quality related genes was stronger in China white as compared to other varieties. FT gene exhibited strong expression in Local red bulb. FSR gene was least expressed in the Local Red but moderately expressed in China White. SIR was strongly expressed in both bulb and leaf tissue of all the varieties under study. Morphological analysis was also carried out to compare the yield potential in different cultivars. From the mean value analysis and LSD values, "China white" variety performed better with regard to Plant Height, Stem Diameter, Flag Leaf Length, Flag Leaf width, Smaller Leaf length, Smaller Leaf width, No of cloves and 100 clove weight.

# 1 INTRODUCTION

## 1.1 Background of *Alliums sativum*

*Allium sativum* L. (Garlic) belongs botanically to the Lillaceae family and Allium genus (Amarakoon and Jayasekara, 2017). It is one of the most used bulb crop after *Allium cepa* (Onion). Garlic is widely used in medicinal and for cooking purposes. In Pakistan Allium consumers are increasing day by day but its production is not meeting the increasing demand (Durak et al. 2004; Srivastava et al. 2005; Katzer et al. 2005; Kabir et al. 2011)

According to the "theplantlist.org" Allium genus is huge including 972 recognized species that are dispersed and grown worldwide including the dry and irrigated zones on earth (Fritsch and Friesen, 2002). Due to evolution of species Allium are now grown in diverse areas, They are preferably grown in sunny and dry areas in moderately humid and dry environment because of which they have great difference in their structure, leaves, bulb and flowers (Hanelt, 1990). Due to cultivation is wide range of environment there is selection of specific key genes that control the structure and help plant to cope with diverse environmental factors and become difference from their wild relatives (Doebley et al. 2006; Baldwin et al. 2013). After analyzing the key genes linked in the formation of underground organ and plant structure, wild species can provide great information about key gens controlling the adaptation and selection during different biotic and abiotic stresses (Ross-Ibarra et al. 2007). *Allium sativum* (Garlic) grow up to height of 1.2 meter and its underground storage organ is called bulb. Garlic can be grown in variety of climate most favourable is mild climate, soft neck and hard neck garlic are two most common types of garlic that is grown all over the world (Iciek et al. 2009).

There are two major types of garlic grown all over the world, one is soft neck and other is hard neck shown in figure (Burba 2013a, b). Garlic is consumed in various way depending on the region of use, Most of the people use it in its green and fresh form (e.g. chopped, sliced or minced) and while in most other countries it is also used in dry form (e.g. powdered). In Asian and African continent garlic leaves are also used

in the dietary products and it is also the major component of the nutraceutical industry (Cavagnaro and Galmarini 2007).

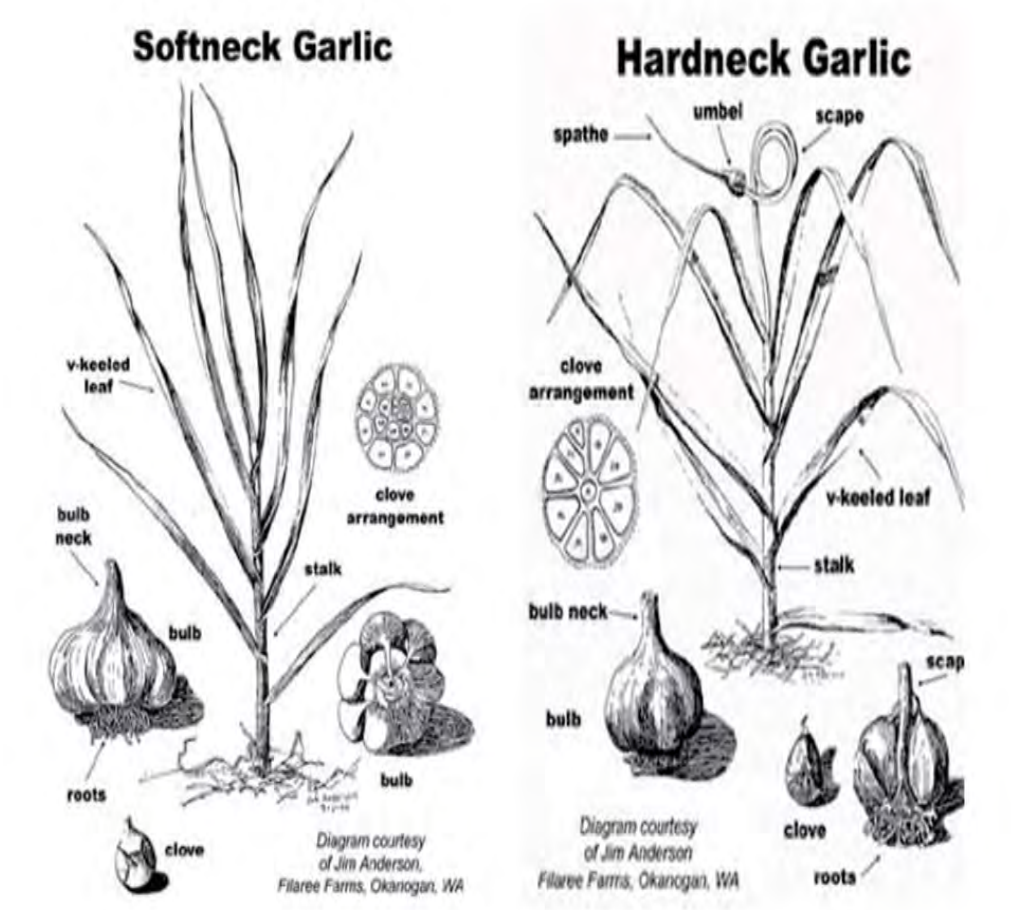


**Figure 1.1 Difference between the Hardneck garlic and Softneck garlic** difference among different parts of garlic are shown including clove, complete bulb, cloves arrangement and plant structure (Adapted from Jim Aderson, Flare Farm, and Okanogan, WA 2018)

Recent research have shown that there are about thirty five thousands protein codding regions present is alliums as compared to plants having smaller genome plants like Arabidopsis having twenty seven thousands and *Oryza sativa* (Rice) containing thirty seven thousands (Sohn et al. 2016; Wendel et al. 2016). Alliums having the larger genome is not because of the single gene family, Key traits of alliums involve sulphur assimilation, production of flavonoids, fertility, flowering and bulbing (Sohn et al. 2016; Wendel et al. 2016). Huge amount of duplication of DNA occurred in the Alliums as a result single gene family have undergone enormous expansion DNA (Jakse et al. 2008; King et al. 1998).

## 1.2 Cultivation of Allium

According to the FAO 2017 data garlic is the 4th most cultivated leguminous vegetable after tomatoes, potatoes and Cassava. More than 20 eatable Allium species are grown and consumed by humans all over the world including, Garlic, Onion, Leek, Japanese bunching onion and Shallot (van der Meer, 1997). Allium for their beautiful growing plant and unique flower, they are also used for decorative purposes in parks and home lawns. *Allium aflatunense* is used as an ornamental plant. Production of Allium seed is also cost-effective activity as gross production of allium vegetable was $ 61,348 million (FAO, 2017). According to the statistics 70% dry *Allium cepa,* 25 % dry *Allium sativum*, 4% green *Allium cepa* and 1 % leeks are used worldwide (FAO, 2017).There is increase in production of Alliums especially *Allium sativum* and *Allium cepa* production increases significantly since 2000 because per capita consumption increasing day by day due to awareness of their potential in medical field (FAO, 2017).

According to the FAO statistics 2017 garlic is sown worldwide in one and half million hectares as a result twenty five million tons of garlic is produced every year (FAO, 2017). Since 1970, with the passage of time and increasing amount of consumers the production of garlic is raised about ten times and production area is increased four times (FAO, 2017). Every year it add up massive amount of money to different countries economy, in 2014 the gross production value was about 15,129 million US dollars. Garlic is used in various forms like dried or fresh, dry processed garlic is used mostly in the dietary, cooking and health purposes. The favorable conditions for production is temperate regions, Most of the garlic growing countries trade its product for generating revenue as they can be stored for longer period of time and its quality did not degrade (FAO, 2017).

China leads the production of garlic producing about eighty percent of the total world production, India is at 2nd position in production countries of garlic. China producing more than 10 times the production of Indian garlic. Other leading producing counties include Brazil , Argentina , Uzbekistan , USA , Spain , Myanmar , Russia , Egypt , South Korea . According to the FAO statistics the list of leading importers of the garlic include Pakistan, Germany, Vietnam, Brazil, Russia, Italy, Vietnam, France and USA (FAO, 2017). Leading producing countries include China they are producing

both the fresh and dry garlic and exporting to the international market, the average dry yield of garlic is sixteen tons per hectare. A few countries also producing thirty tons per hectare of dry garlic also. Garlic is extensive crop it needs proper agronomic practices, water and fertilizer to give good yield at the end. According to the literature the production of garlic 45 years ago was increasing very slowly while 15 years ago there is peak in the production. China is the main contributor in drastic increase in production of garlic it is reported that Republic of Korea is among the larger consumer of garlic as ten kg of garlic is consumed by each person each year. Average garlic consumption is increasing day by day from last decade as consumers know the importance of it in our daily life it increases from 1.6 kg to 3kg of dry garlic per person per year (Cavagnaro and Galmarini 2007; Gonzalez et al. 2009).

## 1.3 Economic and medicinal benefits

Organo sulphur compound is the most commonly found in garlic, it is reported that this compound have anti-cancerous activity (Borkowska et al. 2013). Due to garlic medicinal benefits it is well known as therapeutic medicinal plant. In the whole it is used in various dietary and medicinal ailments, Garlic was first time published as a medicinal plant in the book of Zoroastrian holy written in the late sixth century BC (Dannesteter, 2003; Zain al-abden et al. 2017)

In Asia most of the people use *Allium sativum* (garlic) in their daily food diet and it is common ingredient of every dish preparation (Kabir et al. 2011; Assefa et al. 2015). Pakistan garlic yield is not satisfactory as compared to the world's average production yield that is 9.67 tons per hectare, Production yield in Pakistan is around 8.99 tons per hectare because of which we are spending lot of foreign exchange money to import garlic. Pakistani garlic cultivars have very low yield potential to meet our dietary demands. Because of shortage of land garlic is only grown over an area of 8 thousands hectares. (Economic Survey of Pakistan, 2018)

Chemical composition shows that most of the *Allium sativum* bulb consists of water about 65% other elements is garlic bulb are organo sulphur about 2.3%, protein in the form of allinase 2%, amino acid 1.2 % and fiber content is 1.5 % (Divya et al. 2017). According to the recent research garlic is reported having beneficial impacts on the human health, Most common health benefits include increase in detoxification of

foreign compounds, antioxidant , antimicrobial risk, reduction of cardiovascular diseases, inhibiting platelet aggregation; and risk reduction of cancer . Sliced garlic and ginger chopped stored in 95 percent of ethanol can prevent drug pathogen, clinical study indicate that anti-microbial activity (Karuppiah and Rajaram, 2012) it is reported through experimentation that *Allium sativum* (Garlic) products can be used to treat heart related ailments (Aviello et al. 2009; Colín-González et al. 2012; Chan et al. 2013; Santhosha et al. 2013). To decrease the risk of systolic blood pressure aged garlic is used that can be made by slicing the garlic and then storing it in the diluted ethanol of 20 % (Iciek et al. 2009; Ried et al. 2013; Borkowska et al. 2013).

Clinical studied have shown that most of the patient higher blood pressure can be coped with the use of garlic. It can also be used as a antihypertensive agent (Omar, 2013, Stabler et al. 2012 Garlic is the remedy that can be used for lowering high cholesterol of the patients without any sort of health risk as compare to the conventional medicine (Wang et al. 2012; Ried et al. 2013; Allison et al. 2012)

## 1.4 Yield related genes evolution in Allium species

Due to the evolution all the plants genomes undergoes wide range of variation in their genome size and genes families among different species due to rearrangements including duplication of genome and deletion of amino acids. Plants can now easily be adopted in the diverse environmental conditions (Varshney et al. 2014). Allium genome consist of all the gene families that are present in the monocots such as Oriza sativa (rice) with smaller genome size. All the Allium species have large genome size because they have repetitive DNA sequences in their genome. After the access to the latest tools of sequencing, plant genome sequences have provided tremendous information to identify the key genes playing role in the yield (Varshney et al. 2014)

All Allium species have huge genome size of 16 Gbps as compared to other plants and very complex composition of structure that cause hindrance in sequencing of their genome. (Khosa et al. 2016; McCallum 2007).While comparing Onion (*Allium cepa* ) genome size and *Oryza sativa* (Rice) genome size it is concluded that *Allium cepa* genome size is fourty times larger than oriza sativa genome , Rice genome is well annotated and provide ample amount of information to compare with other plants of different families (Jakse et al. 2008; McCallum 2007). After examining the whole

genome of different species of plant it is discovered that huge genome size is due to the intensification of the DNA repetitive sequences and replication of the whole genome, (Wendel et al. 2016).

Another example is that forefather of the rice species had 5 chromosome and with the passage of time genome duplication occurred and first resulted in ten chromosome and later then existing twelve chromosomes in the *Oryza sativa* (Rice) (Wendel et al. 2016).The history of Allium species is about four thousand years old, different drawing of allium plants has been found carved on the Egyptian pyramids walls that shows Allium was used as a daily dietary component. Growing of allium on different cropping areas caused evolution of vital yield related gens that play key role in the formation of bulb and structure of the plant, those key genes help to differentiate current cultivars from their ancestors (Baldwin et al. 2013; Doebley et al. 2006). Cultivation and distribution of allium species in different areas undergone wide range of changes in their gene families and attributes, those key gene families can be examined to drive out the information related to the key gene selection and evolution from their ancestors (Ross-Ibarra et al. 2007). While comparing the large genome of the Allium with other plants it is concluded that large genome is not due to the duplication of any single gene family. According to the recent research the *Allium cepa* have 35 thousand protein coding genes as compare to the plants having smaller genome size like Arabidopsis having 27 thousand and rice having 37 thousand protein coding genes (Sohn et al. 2016; Wendel et al. 2016).

Cultivation and distribution of allium species in different areas undergone wide range of changes in their gene families and attributes, those key gene families can be examined to drive out the information related to the key gene selection and evolution from their ancestors (Ross-Ibarra *et al,* 2007).While comparing the large genome of the Allium with other plants it is concluded that large genome is not due to the duplication of any single gene family. According to the recent research the *Allium cepa* (onion) have 35 thousand protein coding genes as compare to the plants having smaller genome size like Arabidopsis having 27 thousand and rice having 37 thousand protein coding genes (Sohn et al. 2016; Wendel et al. 2016). Allium genome comprehend less amount of gene families member so, instead of large scale genes duplication the Allium genus undergone huge amount of DNA repetitive sequences expansion (Jakse et al. 2008;

,King et al. 1998).Transcription factor gene family have different number of distinct genes in them that is most probably due to the alteration in gene family caused by genome duplication and deletion Though *Allium cepa* and *Oryza sativa* have same number of transcription factors (Guo 2013).

## 1.5 Evolution of bulbing and flowering genes in Alliums Species

PEBD (Phosphatidylethanolamine-binding domain proteins) is encoded by the FT gene family that is present in every plant genome (Turck et al. 2008; Kardailsky et al. 1999).The amount of those genes are present on distinct number among different plant species , In *Oryza sativa* (Rice) thirteen FT-like genes are found while in the model plant Arabidopsis there are only six Ft like gens are found including FT like genes(FT), Mother of FT (MFT), Twin sister of FT (TSF), Brother of FT (BFT), Arabidopsis thaliana relatives of centroradialis (ATC) and Terminal flower like 1 (TFL1) (Zheng et al. 2016; Turck et al. 2008) shown in Table 1.1. After conducting the phylogenetic analysis it is concluded that PEBP-gene family subdivided into three families Mother of FT (MFT), FT like genes (FT) and TFL1 (Terminal flower like 1)(Karlgren et al. 2011).

TFL1 and FT-like gene of PEBP-gene family have conserved sequence in almost every plant species but they perform opposite function to each other. TFL1 supress the formation of flowering while the FT-like gene act as initiator for flowering (Wickland and Hanzawa 2015; Turck et al. 2008). FT-like genes are expressed in the leaves of the plant under favourable environmental condition and then the produced protein is transported to the upper portion of the leaves to initiate blossoming and storage organ formation (Turck et al. 2008; Kardailsky et al. 1999).

Eudicots have almost 6 times less amount of FT like genes as compare to the monocots. Monocots have higher number of FT like genes as compare to dicots it might be due to the partial duplication of their genome sequence (Zheng et al. 2016). Research reported that 6 FT like genes copies were found in bulb onion (*Allium cepa)* genome. After conducting the phylogenetic association it was discovered that all 6 of them fit in to the FT like clade (Lee et al. 2013). FT-like genes are reportedly involved in the activation of flowering signal but few members of this family are also involve in other functions like bulbing and tuber formation in several plant species (Pinand Nilsson

2012). Formation of bulb might be carried out by FT like gene in Allium species and it also enhance the growth of Poplar, Tomatoes heterosis, tuber formation in potatoes and opening of stomatal cells in plants (Kinoshita et al. 2011; Lee et al. 2013; Krieger et al. 2010; Navarro et al. 2011; Hsu et al. 2011)

**Table 1.1 FT like genes present in different pants**

| Name of plant | No of copies | Function | References |
|---|---|---|---|
| ***Allium*** | 6 | Bulb formation , Inflorescence | Lee et al. (2013) |
| ***Oryza sativa*** | 13 | Inflorescence | Zheng et al. (2016) |
| ***Zea mays*** | 15 | Inflorescence | Zheng et al. (2016) |
| ***Sorghum bicolor*** | 11 | Inflorescence | Zheng et al. (2016) |
| ***Solanum tuberosum*** | 3 | Tuber formation , Inflorescence | Navaro et al. (2011) |

Allium bulb have 3 FT like genes namely ACFT1, ACFT4 and ACFT2 that are also involve in the inflorescence period of model plant Arabidopsis (Lee et al. 2013).In Allium Species flowering is promoted by the AcFT1 gene while on the other hand Flowering is supressed by the AcFT4 gene. AcFT2 play a very minute role in the flowering period of both Allium species and Arabidopsis. FT like gene expression analysis indicate that during the favourable conditions at juvenile stage the expression of the AcFT4 gene is upregulated while suppressing s the effect of AcFT1 gene same phenomena occur in both short and long day plant species . Alliums under supressed condition the expression changes antagonistically as AcFT4 expression is supressed and AcFT1 gene expression is upregulated for the development of underground part (Bulb) (Lee et al. 2013).

Phylogenetic, functional and expression study of FT like genes shows that ACFT4 supress the formation of bulb so act an inhibitor of bulbing while AcFt1 regulate the formation of the bulb in Allium so act a promoter of bulbing (Lee et al. 2013). Further experiment indicate that if there is over expression of the AcFT4 gene in the Allium it will not form underground storage organ (bulb) because of it same

structure like leeks if produced on the other hand if AcFT1 is over expressed in the Allium species there will be formation of bulb like structure (Lee et al. 2013).

There are number of genes that control the photoperiodic regulation of flowering and photoperiodic formation of underground part (Bulb) in the Allium (Brewster 2008). Different developmental process of bulbing and flowering are enabled by the response of circadian clock genes. Those genes are regulated in response to the change in the season (Sanchez and Kay, 2016).There are number of key genes that play important role in the formation of flowering and bulbing, In the Model plant Arabidopsis there is up regulation of clock gens and formation of FT protein that moves to upper part (shoot apical meristem) of the plant and promote formation of bulb (Johansson and Staiger 2015; Sawa et al. 2007).

Most of plants have only 1 copy of the FKF1 and GI genes but those plants who have undergone multiple duplication of genome such as *Allium cepa* (onion) there is only one copy of those genes (Table 1.2) (Taylor et al. 2010).In *Allium cepa* (Onion) only 2 gene duplicates of the FKF1 and GI are present, that concluded that the genome duplication occurred only in certain part of the Allium genome as a result only few genes have multiple copies (Taylor et al. 2010). Similarly in the plants like potatoes, initiation of tuber formation as a response to day-length is promoted by the StFKF1 and StG1 that activate StSP6A (orthologue of the FT like genes) (Kloosterman et al. 2013).

**Table 1.2 Different bulbing genes homologues indicated in several plants**

| Name of plant | GIGANTEA | FKF1 | References |
|---|---|---|---|
| **Arabidopsis** | 1 | 1 | Higgins et al. (2010) |
| **Rice** | 1 | 1 | Murakami et al. (2007)<br>Higgins et al. (2010) |
| **Barley** | 1 | 1 | Higgins et al. (2010) |
| **Maize** | 2 | 2 | Dong et al. (2012) |
| **Soybean** | 3 | 2 | Li et al. (2013) |
| **Medicago** | 1 | - | Kim et al. (2013) |
| **Onion** | 2 | 1 | Taylor et al. (2010) |

## 1.6 Evolution of flavonoid biosynthesis genes in Allium Species

RNA sequencing facilitated in recognising the key genes responsible for the formation of flavour in Alliums (Schwinn et al. 2016; Khosa et al. 2016).There is plenty of variation in the formation of flavonoids compounds in Alliums, Mainly flavonoid biosynthetic pathway help in the formation of three compounds such as production of bright yellowish chalcones, yellow flavonoids and red anthocyanins (Khar et al. 2008; Masuzaki et al. 2006). There are more than one key gene reported that is involved in the formation of the flavonoids Chalcone isomerase (CHI), Chalcone synthase (CHS), Dihydroflavonol 4 reductase (DRF) and ANS gens (anthocyanidin synthase) .if there is slightly mutation in these genes there will be formation of a different shade (Khar et al. 2008; Masuzaki et al. 2006)

It is reported that there are multiple copies of the DRF gene and ANS gene found in different species while in *Allium cepa* (bulb Onion) contain only 1 copy of the gene (Table 1.3). ANS and DRF genes have different catalytic activity and expression in various organs, Due to gene duplication there are multiple number of copies having diverse function DRF (Dihydroflavonol 4-reductase) enzyme (Huang et al. 2012). Anthocyanidin synthase (ANS) gene is activated by the different transcriptional factors including, WDR (WD-repeat proteins), bHLH and R2R3-MYB. Genes involved in the biosynthesis of flavonoid are directly activated by WDR, bHLH and R2R3-MYB transcriptional factors (Xu et al. 2015)

Lately, there are four new transcription factors identified R2R3-MYB that play a vital part in the production of flavonoid biosynthesis in *Allium cepa* (Bulb Onion) There is overexpression of CHS, DRF, MYB1 and ANS gene in the red *Allium cepa* (Bulb Onion). Experiments indicate that if the MYB1 gene is knock-downed there will be increase in anthocyanin production while if there is high expression of MYB1 gene it will lead to decrease in production of anthocyanin (Schwinn et al. 2016). *Allium cepa* (Bulb onion) MYB1 gene is very similar to the dicots and anthocyanin-MYB mutant of snapdragon. Overexpression of MYB1 gene in transgenic *Allium cepa* resulted in the formation of strong red colour in both leaf and callus while there was no strong pigmentation in the control plants of *Allium cepa* (Schwinn et al. 2016).

**Table 1.3 Different anthocyanin production pathway genes homologues indicated in several plants**

| Name of plants | DRF gene | ANS gene | References |
|---|---|---|---|
| **Arabidopsis** | 1 | 1 | https://www.arabidopsis.org/ |
| **Tomato** | 1 | 1 | Bongue Bartelman et al. (1994) |
| **Potato** | 1 | 1 | De Jong et al. (2003) |
| **Rice** | 1 | 2 | Shih et al. (2008) |
| **Lotus japonicus** | 6 | - | Shimada et al. (2005) |
| **Onion** | 1 | 1 | Kim et al. (2004a, b) |
| **Populus trichocarpa** | 2 | 2 | Huang et al. (2012) |

## 1.7 Evolution of sulphur assimilation genes in Allium species

The Allium is the bigger genera having six thousand reported species. Most studied and widely used species of this genera are *Allium cepa* (Onion), *Allium sativum* (garlic), *Allium ascalonicum* (shallots), *Allium schoenoprasum* (Chives) and *Allium ampeloprasum* (leeks) (Poojary et al. 2017). Alliums are rich in Biological active compounds (BAC) (Radulović et al. 2015; Zeng et al. 2017). They have huge quantity of flavonoids (Nishimura et al. 2016). Chemical composition indicate that Allium are made up of both oil soluble secondary phytochemical metabolites for example OSC secondary phytochemical metabolites and water soluble components like sugars, amino acids, vitamins and different enzymes (Koca and Tasci, 2016).

Most of the raw Allium cloves compose of 59% water and other elements present are 0.5% lipids, 2 % eatable fibres, 6% proteins and 33% carbohydrates. There are Vitamin B1, B2, B6, C and E , pectin, adenosine, postglandians, lectins, essential amino acids, biotin, nicotininc, phospholipids, fatty acids and glycolipids (Koca and Tasci, 2016).The Allium genera have 2.3 % of medicinal OCS , while Allicin is the present more only in Alliums (Koca and Tasci, 2016). Most of the allium species are popularly reported to have protective chemicals for human health (Bianchini and Vainio, 2001; Nicastro,Ross and Milner, 2015).Alliums are reported to have the anti-hyperhomocysteinic, hypolipidemic, anti-thrombotic , hypocholesterolemic and antihypertensive properties that can decrease heart related diseases risk (Saljoughian et al. 2017).

Most of the Allium species have hypoglycaemic, hypotensive, hepatoprotective, neuroprotective, antiamensic, anti-asthmatic, anti-mutagenic, anti-oxidant, anti-spamdic, anti-diabetic, anti-protozoal, anti-viral, anti-microbial (Wang et al. 2013; Bisen et al. 2016; Prakash et al. 2007; Santas et al. 2008; Nile et al. 2017).Polyphenols and organ-sulphur are bioactive compounds of Alliums and also present in other plant species,they control most of the health related activities (Putnik et al. 2011; Rosello-Soto et al. 2018; Zhu et al. 2018; Putnik Barba et al. 2017; Rosello-Soto et al. 2018; Fidelis et al. 2018) . Organo-sulphur and polyphenols are very sensitive compounds because of which during processing most of their concentration abolish (Giacometti et al. 2018; Putnik et al. 2018; BursaćKovačević et al. 2018; Barba et al. 2017). Both the compounds are helpful in the process of digestion (Arreola et al. 2015)

Human body is not that efficient to absorb all of those bioactive compound so, Bio-availability is a huge factor influencing the concentration (Granato et al. 2016)The pungent taste of the *Allium sativum* (Garlic) is mainly because of the bio active compounds like organo sulphur that is rich in health benefits (Rahman, 2007; Tepe et al. 2005). Alliin is the bio-active compounds mostly present in the unharmed bulb of Alliums subsequently to organosulfur that is converted to Allicin while chopping (Amarakoon and Jayasekara, 2017).

It is reported that photosynthetic plant species, fungi and most of the bacterias have the genes that carry out the sulphate assimilation pathway as a result of that wide range of sulphur containing compound such as alliins and glucosinolates and sulphur containing amino acids are produced (Takahashi et al. 2011). Concentration of the organo sulphur compounds is different among different allium species, most of the research is carried out on the *Allium cepa* (bulb onion). It is reported that there are multiple key genes in onions and Garlic that are involved in the synthesis of organo sulphur compound (Sun et al. 2016; Kamenetsky et al. 2015; and McManus et al. 2012; Brewster 2008)

SiR (sulphite reductase) is the key factor that control the production of sulphur compound, it help in the reduction of sulphite as a result sulphur compound is formed in the alliums (McManus et al. 2012; Takahashi et al. 2011). There is only one copy of sulphite reductase present in the *Allium cepa* genome while *Oryza sativa* have 2 copies

of SiR (Kopriva 2006; McManus et al. 2012).It is identified after conducting the garlic transcriptome analysis that there are huge number of isoforms that tiger different key enzymes to help out in the production of sulphur compound (Kamenetsky et al. 2015).

Most of the key genes of sulphur production are expressed in other part of Alliums excluding bulb, cysteine sulphur-oxides are produced in the roots and leaves of the plants then they are transferred to the cloves (Kamenetsky et al. 2015).

## 1.8 Aims and Objective

The main objectives of the research are as under:

- Phylogenetic relationship of the selected yield related genes among *Allium sativum* (Garlic) and *Allium cepa* (Onion)
- Expression analysis of selected yield and quality related genes in tissues including i.e. bulb, stem and leaves of five different genotypes of garlic

# 2 MATERIALS AND METHODS

## 2.1 Selection of *Allium sativum* cultivars

Five different soft neck and hardneck *Allium sativum* cultivars NARC-G1, Local white, Local Red, China white and Advance line were selected for the comparison of yield related genes (Figure 2.1). These selected cultivars are commonly grown all over the Pakistan. Seeds were collected from BCI (Bio-Resource Conservation Institute) of National Agricultural Research Centre (NARC), Islamabad.


**Figure 2.1 Different varieties of garlic selected for research study** (**A.**NARC-G1, **B.** Local white, **C.** Local Red, **D.** China white)

## 2.2 Morphological analysis

The fallowing morphological data was recorded during the field trials including height of the plant, stem width, flag leaf length and width, smallest leaf length and width, number of cloves and weight of the cloves. All the data was recorded randomly from different plants from each replication to remove chances of error. The data of stem diameter was observed in millimetres using digital Vernier calliper, Number of cloves (garlic seeds) were computed physically by hand and digital balance was used to

determine the weight of the cloves. All field data was analysed in various statistical tools namely Statistics 8.1 and Minitab to compute ANOVA (analyses of variance) that shows the difference among different parameters. If that difference among parameters was significant then they were carried out for further analysis to determine the LSD test (Least significant differences).

## 2.3 Identification of yield related genes

Published literature was used to identify 12 different genes from which only four key genes were selected involved in the yield and quality parameters including bulbing, flavonoid synthesis, Aroma, colour formation and biosynthesis of Sulphur shown in Table 2.1. CDS Sequences of the 4 genes Ft-Like Gene, Anthocyanidin Synthase (ANS), Lachrymatory Factor Synthase (LFS) And Sulphite Reductase (SIR) were retrieved from two different well known databases NCBI and Ensembl Plants. Basic Local Alignment Search Tool (BLAST) was used to search homology among different yield related genes in *Allium cepa* and *Allium sativum* species.

**Table 2.1 Selected gene for research work**

| Sr No. | ***Name of gene*** | ***Expression*** | ***Function*** | ***References*** |
|---|---|---|---|---|
| 1 | FT-like gene | Bulb and Flower | Blubbing and Flowering | Lee et al. (2013) |
| 2 | Anthocyanidin synthase (ANS) | Bulb | Colour formation | Kim et al. (2004) |
| 3 | Lachrymatory factor synthase (LFS) | Bulb | Flavonoids production | Imai et al. (2002) |
| 4 | Sulphite Reductase (SIR) | Bulb | Organo-sulphur formation | Kopriva 2006; McManus et al. 2012). |

## 2.4 Phylogenetic analysis of yield related genes

*Allium cepa* and *Allium sativum* yield related genes sequences were aligned with each other using Clustal W alignment (Thompson et al. 1997). Neighbour Joining (NJ) algorithm was used for the construction of phylogenetic tree with the help of MEGA 7 software (Tamura et al. 2013).

The implication of nodes was calculated using a bootstrap study of 1,000 replicates. For the surety of different domains that shows the topology of Neighbour joining tree, pairwise gape deletion mode was used in MEGA 7.0. The phylogenetic tree was constructed after aligning all the gene sequences in Clustal W in MEGA 7.0 (Thompson et al. 1994).

## 2.5 Primers designing for gene expression

Primers 3 software was used for designing the primers of *Allium sativum* (Rozen and Skaletsky, 1999) to examine the expression of yield related genes FT-like, ANS, LFS and SIR in three different tissues i.e. bulb, stem and leaves of *Allium sativum* cultivars. Those primers were then checked with the help of two different softwares Multiple Primer Analyzers (Thermofisher Scientific) and primer stat (Stothard 2000). The specificity of the designed primers was tested by UCS PCR at UCSC-In Silico PCR genome browser (https://genome.ucsc.edu/) (Rhead et al. 2009). The designed primers for yield related gene expression analysis are given in (Table 2.1)

**Table 2.2 List of primers used for gene expression**

| S.No. | Name | Sequence |
|---|---|---|
| 1 | **ANS-Forward** | 5'-ATGCCCATAGAAAGCGTGATAATT-3' |
| 2 | **ANS-Reverse** | 5'-ATGCAGTCGCGTGAGTCGAGGTCG-3' |
| 3 | **LFS-Forward** | 5'-ATGGAGCTAAATCCTGGTGCACCT-3' |
| 4 | **LFS-Reverse** | 5'-ATGCTTGCTCTGGCTTTGTATTTG-3' |
| 5 | **FT-Forward** | 5'-ATGATGGATTCGGATCCGTTAAGG-3' |
| 6 | **FT-Reverse** | 5'-GGAGCATCTGGGTCCACCATTACA-3' |
| 7 | **SIR-Forward** | 5'-ATGGAAGCGATGGCGGCGACTGCG-3' |
| 8 | **SIR- Reverse** | 5'-CTCAGGAACAGCTACAGCCGTAAT-3' |

## 2.6 Plant and sample preparation

Collected Seeds of 5 different varieties of *Allium sativum* from plant Bio Resources Conservation Institute (BCI) of National Agricultural Research Centre (NARC), Islamabad were grown at NARC. Morphological studies of four varieties were completed including plant height, No of cloves, bulb weight, stem thickness, leaf length and leaf width. Morphological data is given in the Table No 3.1. Further study regarding gene expression, comparative and phylogenetic study of gene expression and bioinformatics study was also completed in NIGAB


**Figure 2.2 Different growth stages of garlic in field**
**1:** Sowing in line **2:** emergence of garlic **3:** three leaf stage **4:** bulbing stage **5:** garlic plant at bulbing stage

Samples were collected after 2 months of sowing at bulb formation stage from the field of NARC, Islamabad. Those collected samples were then transferred to the National Institute of Genomics and Advanced Biotechnology labs for further study purpose.

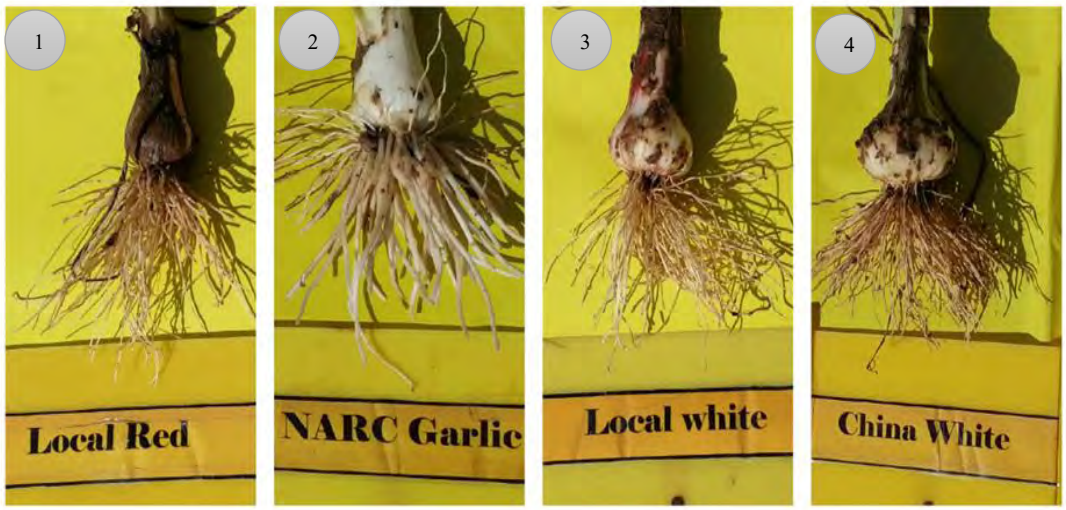


**Figure 2.3 Different varieties at bulbing stage**
1st local red, 2nd NARC G1, 3rd Local white, 4th Local white

## 2.7 Extraction of RNA

Total RNA from the 3 different tissues such as bulb, leaves and stem of *Allium sativum* cultivars namely NARC-G1, Local white, Local Red and China white was extracted by using GeneJET Plant RNA Purification Mini Kit (Thermo Scientific). 100mg of fresh tissues in liquid nitrogen were grinded mechanically in mortar and pestle. Grinded samples were then transferred in Eppendorf-tubes having 500 µL of RNA Lysis buffer. Samples were then incubated for three minutes at 56°C. Incubated samples were then centrifuged for 5 min at ≥14,000 rpm. Afterward 450 to 550 µL of supernatant (upper transparent part) was transferred into a new micro centrifuge tube. 250 µL of ethanol having 96% concentration was added .Mixed solution was then transferred to a collection tube. Column was then centrifuged for at 11000 rpm for one minute. Afterward solution obtained in the flow through was discarded. 700 µL of WB-1 (Wash buffer 1) was transferred to the purification column then centrifuged at 11000 rpm for 1 minute. Purification column was then transferred to new 2 mL collection tube while discarding both collection tube and flow through. 500 µL of WB-2 (Wash Buffer 2) was transferred to the purification column followed by centrifuge at 11000rpm for 1 min. Solution containing in Flow through was discarded. Again 500 µL of WB-2 (Wash Buffer 2) was added followed by for 1 minute. Purification column was not discarded and put into to a new 1.5 mL collection tube. For the processes of elution of RNA from the purification tube 50 µL of nuclease free water was poured and then centrifuged at

11000 rpm for 1 minute for 1 min. RNA obtained at the bottom of the tube was immediately transferred to -20°C for further use.

## 2.8 Gel Electrophoresis

It was used to analyse extracted RNAs from different tissues of *Allium sativum* samples. 1% agarose gel was prepared in a flask using 55ml distilled water, 6ml of 10X Tris-acetic TAE buffer and 0.5 grams of agarose (Appendix I). The mixture was then heated in a microwave oven for a few seconds until dissolved completely when solution become transparent then cool down for few mins. 1.5 µl 3 ml ethidium-bromide was transferred to the prepared solution for RNA bands visualization. Solution is then transferred into gel tray having combs already in the place. After several minutes when the gel got solidified combs were removed carefully. 5µl RNA form each extracted sample of different tissues were loaded into the wells along with 3µl of 6-X blue-dye (Thermofisher). First well of each row was loaded with 1kb of ladder as a standard. The loaded gel was then run in gel caster containing 1x TAE buffer for forty minutes at 80 volts. After completion of time gel was then analysed under UV Trans illuminator to confirm the presence of bands.

## 2.9 Quantification of RNA

Before proceeding synthesis of cDNA the quality, quantity and integrity of RNA was determined by using the Thermofisher TM Nano DropTM Spectrophotometer. The quantification wavelength of the spectrophotometer was adjusted at 260/280nm for all purified samples. The samples with low quality or quantity were again extracted for uniform expression of the genes.

## 2.10 Synthesis of cDNA

cDNA synthesis was carried out by means of Thermo Scientific Revert Aid reverse transcriptase-III, First Strand cDNA Synthesis Kit (Appendix II). All the reagents were dissolved, vortex for the short period of time to make them gentle in use. 2µl RNA, 1µl (10mm) reverse primer and 25µl water were mixed in an Eppendorf tube and vortex for few seconds to mix reagents homogenously Fallowed by 5 minutes incubation at 65ºC. Incubated samples were then cool down in ice for 1 minute.

Moreover, the ingredient 5-X reaction-buffer, 4µl of RNAse inhibitor 1µl and 10mM dNTPs 4µl and 1µl reverse transcriptase enzyme were added and fused then incubated for sixty minutes at 42ºC. Hence, the total reaction was carried out for 5 minutes at 70 ºC.

**Table 2.3 Quantificion values of total RNA obtained from Nano drop Spectrophotometer**

| Sr. No. | SAMPLE | VALUE (260/280nm) | RNA (µL) | WATER (µL) |
|---|---|---|---|---|
| **1** | **AB** | 258.16 | 7.75 | 5.75 |
| **2** | **AS** | 514.88 | 3.9 | 9.6 |
| **3** | **AL** | 1800 | 1.11 | 12.4 |
| **4** | **BB** | 452.83 | 4.42 | 9.08 |
| **5** | **BS** | 473.39 | 4.22 | 9.28 |
| **6** | **BL** | 1519 | 1.32 | 12.18 |
| **7** | **CB** | 1132 | 1.76 | 11.75 |
| **8** | **CS** | 275.67 | 7.25 | 6.25 |
| **9** | **CL** | 858.41 | 2.33 | 11.17 |
| **10** | **DB** | 818.98 | 2.44 | 11.06 |
| **11** | **DS** | 451.79 | 4.43 | 9.07 |
| **12** | **DL** | 1076 | 1.86 | 12.14 |
| **13** | **EB** | 338.51 | 5.90 | 7.60 |
| **14** | **ES** | 530 | 3.77 | 9.73 |
| **15** | **EL** | 1016 | 1.97 | 11.52 |

## 2.11 qRT-PCR expression analysis

The expression of yield related genes i.e. Ft-Like gene, Anthocyanidin Synthase (ANS), Lachrymatory Factor Synthase (LFS) and Sulphite Reductase (SIR) in various tissues were checked using comparative ΔCT method in real-time PCR (Applied Biosystems) with Step-One-Plus software. Gene specific primers was used. The kit used for RT-PCR was Taq Man SYBER GREEN (Appendix III) (Fermentas). Elongation factor (EF) was used as endogenous control and the target genes were Ft-Like Gene, Anthocyanidin Synthase (ANS), Lachrymatory Factor Synthase (LFS) and Sulphite Reductase (SIR). The total reaction volume was kept 10µl in which 5ul SYBER Green was used. Tubulin was used as an endogenous positive control to determine whether or not the reverse transcription and/or PCR reaction conditions are optimal. The profile of RT-PCR was set as denaturation for ten minutes at 94ºC, followed by 2nd stage 40 cycles of 40 seconds at 95ºC, forty cycles of 32 seconds at 58ºC, forty cycles of 32 seconds at 72ºC, the 3rd stage of 40 cycles of 15 seconds at 95ºC and forty cycle of 32 seconds at 56ºC. Finally, melt curve study was complete at 52ºC to 95 ºC. Amplification plot, melt curve plot, ΔCT values and standard curve data were recorded. These all data were more analysed with Step-One Plus software system

# 3 RESULTS

## 3.1 Identification and sequence analysis of yield and quality related genes in Alliums

A set of 12 individual yield and quality related genes of Allium species were identified from previous published literature and online data bases. There are plenty of key genes that are involved in the yield related parameters that are expressing in different parts of the plant like underground storage organ (Bulb), leaves and stem (Table 3.1) 4 different yield related genes under the parameters including bulbing, flavonoid synthesis, Aroma, colour formation and biosynthesis of Sulphur were selected for research i.e. genes FT-Like Gene, Anthocyanidin Synthase (ANS), Lachrymatory Factor Synthase (LFS) and Sulphite Reductase (SIR). Sequences of these 4 selected genes were retrieved from the databases NCBI (https://www.ncbi.nlm.nih.gov/) and Ensembl Plants (http://plants.ensembl.org/). Basic Local Alignment Search Tool was used to search homology among different yield related genes in *Allium cepa* and *Allium sativum* species (htpp: // www. ncbi. nlm. nih.gov/Blast/).

## 3.2 Morphological Analysis

Plants were grown in the field of National agriculture research centre, Islamabad. Morphological data was taken when complete bulbing stage was achieved by plants. Plants were selected randomly form the field, Plant Height (cm), Stem Diameter(cm), Flag Leaf Length (cm), Flag Leaf width(cm), Smaller Leaf length(cm), Smaller Leaf width(cm), No of cloves,100 clove weight (gm) data were recoded.

The data analysis regarding the mean values of stem diameter shown in Table No of different varieties of Garlic. Results indicate that a significant difference was observed among different varieties. The mean values show that maximum diameter of the stem was observed in variety China White having maximum stem diameter of (2.6cm) fallowed by variety local white, and local red while the minimum stem diameter was observed in NARC G1.

**Table 3.1 In silico study of 12 yield and quality related genes identified in Alliums**

| Sr. # | Name of gene | Plant part | Expression | Function | Presence | Reference |
|---|---|---|---|---|---|---|
| **1.** | **ACFT1** | Bulb and Flower | During juvenile stage of bulb | Promoter of Bulbing and Flowering | 6 copies | Lee et al. (2013) |
| **2.** | **ACFT4** | Bulb and Flower | During juvenile stage of bulb | Inhibitor of Bulbing and Flowering | 6 copies | Lee et al. (2013) |
| **3.** | **STSP6A** | Bulb | Formation of underground storage organ | Bulb formation initiation | --- | Kloosterman et al. (2013 |
| **4.** | **Anthocyanidin synthase (ANS)** | Bulb | During bulbing Stage | Colour formation | 1 copies in *Allium cepa* | Kim et al. (2004) |
| **5.** | **Di-hyroflavonol 4-Reductase (DRF)** | Bulb | Bulbing and Flowering | Colour formation | 1 copies in *Allium cepa* | Kim et al. (2004) |
| **6.** | **Chalcone synthase** | Bulb | Bulbing | Colour formation | --- | Masuzaki et al. (2006) |
| **7.** | **Chalcone isomerase** | Bulb | Bulbing | Colour formation | --- | Masuzaki et al. (2006) |
| **8.** | MYB Factors **MYB1** | In all plants parts | Formation of underground storage organ | Regulate colour formation | 4 | Schwinn et al. (2016) |
| **9.** | **Sulphite Reductase (SiR)** | Cysteine sulphoxides Synthesized in Leaves And Roots and then translocated to the storage organ | All stages of growth | reduction of sulphite for the synthesis of sulphur compounds in bulb onion and other plants (Takahashi et al. 2011 | 1 copies in *Allium cepa* | Kopriva (2006) McManus et al. (2012) |
| **10.** | **GIGANTEA** | Bulb and Flower | Initiates Bulbing and Flowering | Initiates Transcription of FT which initiate Bulbing and Flowering | 2 Copies | Taylor et al. (2010) |
| **11.** | **Lachrymatory Factor Synthase (LFS)** | Bulb | Flavonoid synthesis | Initiate enzyme activity for the synthesis of flavour | 2 copies | Imai et al. (2002) Masammura et al. 2012 |

The data relating the mean values for height of the plant of different varieties of Garlic are given in Table 2. Average height of the different varieties shows that China white have the maximum height of 91.2 cm followed by the Local Red 86.4 cm, Local White 81.4 cm, China White 91.2 cm, Advance line 74.6 cm and NARC G1 53.2cm. Diameter of the stem was recorded with the help of digital Vernier calliper its shows that china white have maximum stem diameter of (2.6 cm) followed by other varieties Advance line 1 (2.02 cm) Local red (1.74 cm) , local white (1.72 cm), NARC G1 (1.62 cm) . Flag leave is the biggest leave of any plant and have maximum exposure to the sun light. Flag leaf length and width data analysis shows that china white have the maximum length and width as compare to Local white, Local red and NARC G1 respectively.

The data about the mean values for number of cloves bulb-1 of different varieties of Garlic are given in Table 1. There is a clear significant variation in number of cloves bulb-1 of different varieties. The mean values show that a greater number of cloves bulb-1 (30) was noted in variety Local white followed by variety Buner Local (21) and variety China White (15) Advance line (11) respectively, while the smaller number of cloves bulb-1 (10) was recorded by variety NARC-G1

**Table 3.2 Mean values of different yield related morphological parameters**

| Varieties | Plant Height (cm) | Stem Diameter (cm) | Flag Leaf Length (cm) | Flag Leaf width (cm) | Smaller Leaf length (cm) | Smaller Leaf width (cm) | No of cloves | 100 clove weight (gm) |
|---|---|---|---|---|---|---|---|---|
| China white | 91.2 | 2.6 | 67.2 | 2.62 | 54.8 | 2.32 | 15 | 700 |
| Local red | 86.4 | 1.74 | 54.6 | 1.9 | 41.2 | 1.44 | 21 | 103 |
| Local white | 81.4 | 1.72 | 55.2 | 2.04 | 40.4 | 1.88 | 30 | 244 |
| Advance line | 74.6 | 2.02 | 51 | 2.02 | 40.2 | 1.82 | 11 | 680 |
| NARC G1 | 53.2 | 1.62 | 46.6 | 2.42 | 31.8 | 2.16 | 10 | 775 |

The data relating the mean values for 100 clove weight in grams of different varieties of Garlic are given in Table 2. There is a highly significant variation in single clove weight about different varieties of Garlic. The mean values explain us that the maximum clove weight (775g) was noted in variety NARC-1 followed by variety China white (700 g), Advance line (680 g), Local white (244g) and Local red (103g) respectively. Apart from NARC1 all the other two varieties are statistically similar to each other.

**Table 3.3 LSD values for different morphological parameters**

| Varieties | Plant Height | Stem Diameter | Flag Leaf Length | Flag Leaf width | Smaller Leaf length | Smaller Leaf width | No of cloves | 100 clove weight |
|---|---|---|---|---|---|---|---|---|
| China White | 90.06 a | 2.62 a | 67 a | 55.8 a | 55.8 a | 2.32 a | 15.6 c | 703 b |
| Local Red | 86.6 ab | 1.753 c | 54.8 b | 40.93 b | 40.93 b | 1.42 c | 18.06 b | 105.33 e |
| Local White | 81.46 bc | 1.73 c | 55.13 b | 40.33 b | 40.33 b | 1.86 b | 28.93 a | 144.66 d |
| Advance line | 76.33 c | 2.06 b | 52.6 b | 41.06 b | 41.06 b | 1.8 b | 11.53 d | 665.33 c |
| NARC_G1 | 52.73 d | 1.62 c | 47 c | 31.73 c | 31.73 c | 2.16 a | 9.86 e | 781.67 a |

**Note:** values followed by the same letter are not significantly different.

### 3.3 Phylogenetic analysis of yield related genes in *Allium sativum*.

The phylogenetic tree was generated for 40 yield related genes of *Allium cepa and Allium sativum* using neighbour joining method in the MEGA 7 programme. The evolutionary relationship of all yield related genes FT-Like gene, Anthocyanidin Synthase (ANS), Lachrymatory Factor Synthase (LFS) and Sulphite Reductase (SIR) sequences were retrieved from the NCBI data base. The un-rooted linear tree was generated with the help of MEGA 6.06 programme using multiple sequence alignment with Clustal w. To show the evolutionary relationship among yield related genes c and of both garlic and onion distance based method was used with the p-distance strategy. The clade or branch length of a tree shows the evolutionary rate. During generating of phylogenetic tree complete deletion strategy was used to remove distinct and unwanted gaps. To ascertain the accurate gene in the class of its specificity bootstrap value of 1000 replicate was used. FT-Like gene, Anthocyanidin Synthase (ANS), Lachrymatory Factor Synthase (LFS) and Sulphite Reductase (SIR) gene with definite functional and developmental relevance are arranged in a distinct clade, characterize their functional and sequential conservation.

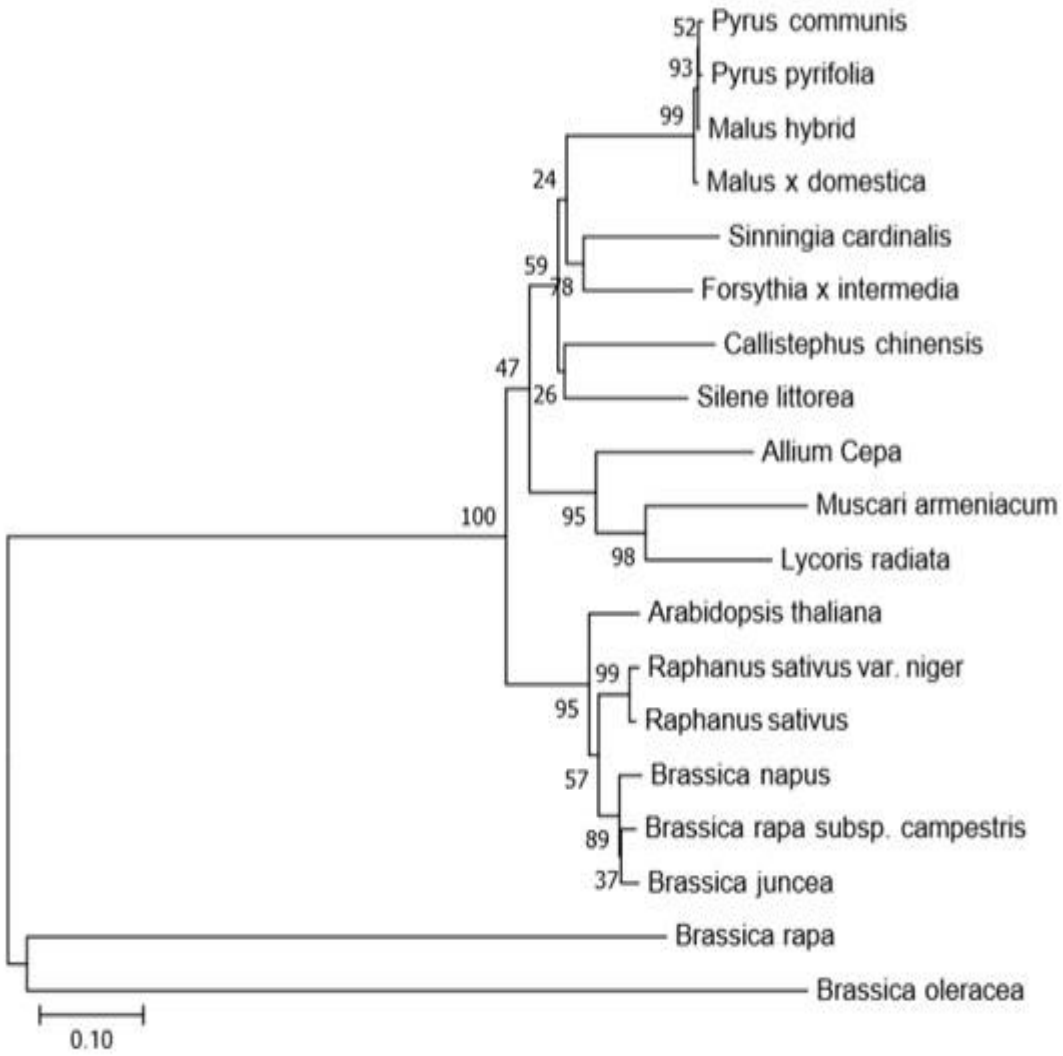


**Figure 3.1 Evolutionary relationships of ANS gene** The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 3.38421338 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. The analysis involved 19 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 562 positions in the final dataset. Evolutionary analyses were conducted in MEGA7.The scale symbolizes the nucleotide substitution rate.


**Figure 3.2 Evolutionary relationships of FSR gene** The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.98116689 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. The analysis involved 6 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 1342 positions in the final dataset. Evolutionary analyses were conducted in MEGA7. The scale symbolizes the nucleotide substitution rate.

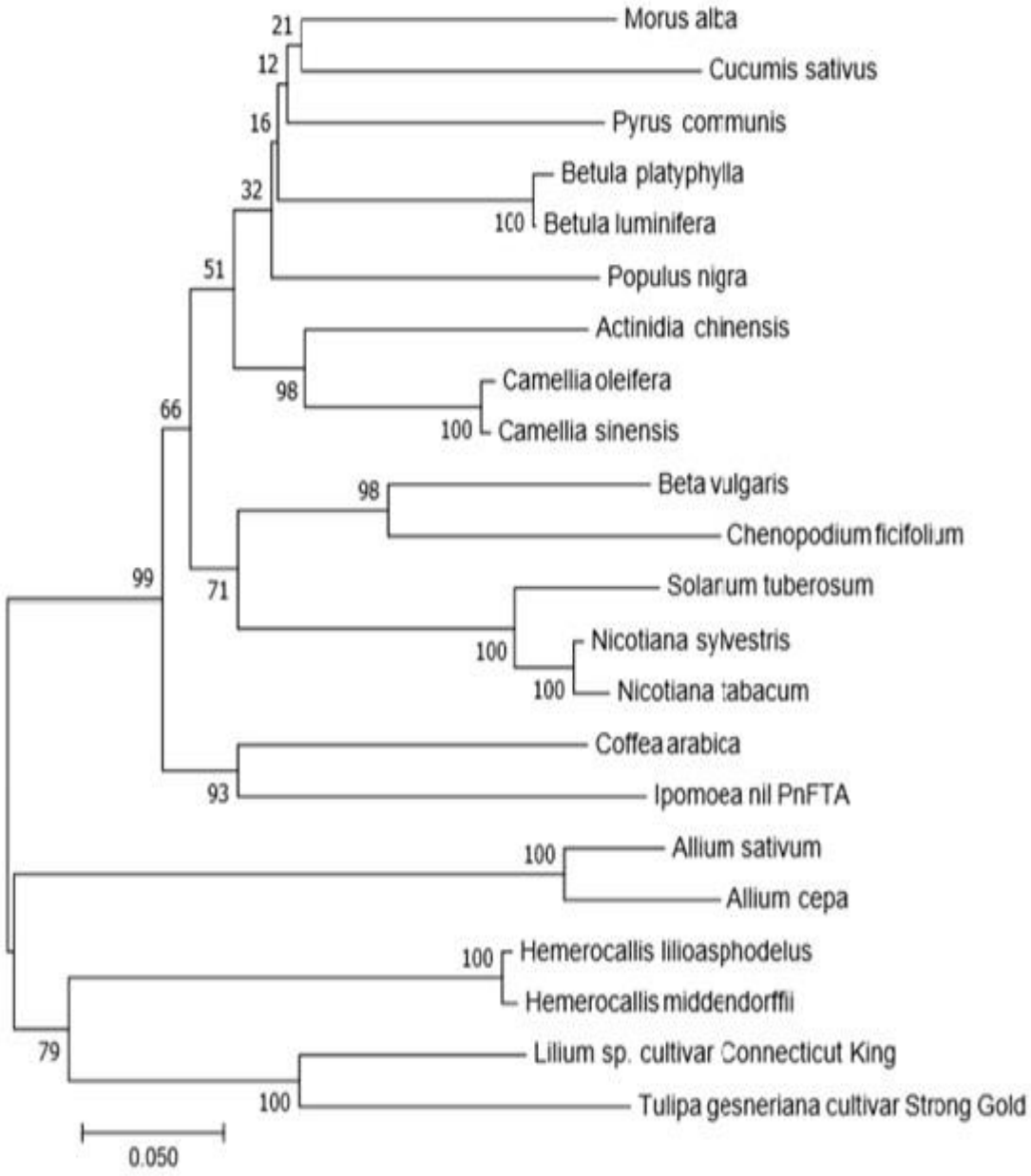


**Figure 3.3 Evolutionary relationships of FT gene** The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 3.11067148 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. The analysis involved 9 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 177 positions in the final dataset. Evolutionary analyses were conducted in MEGA7. The scale symbolizes the nucleotide substitution rate.

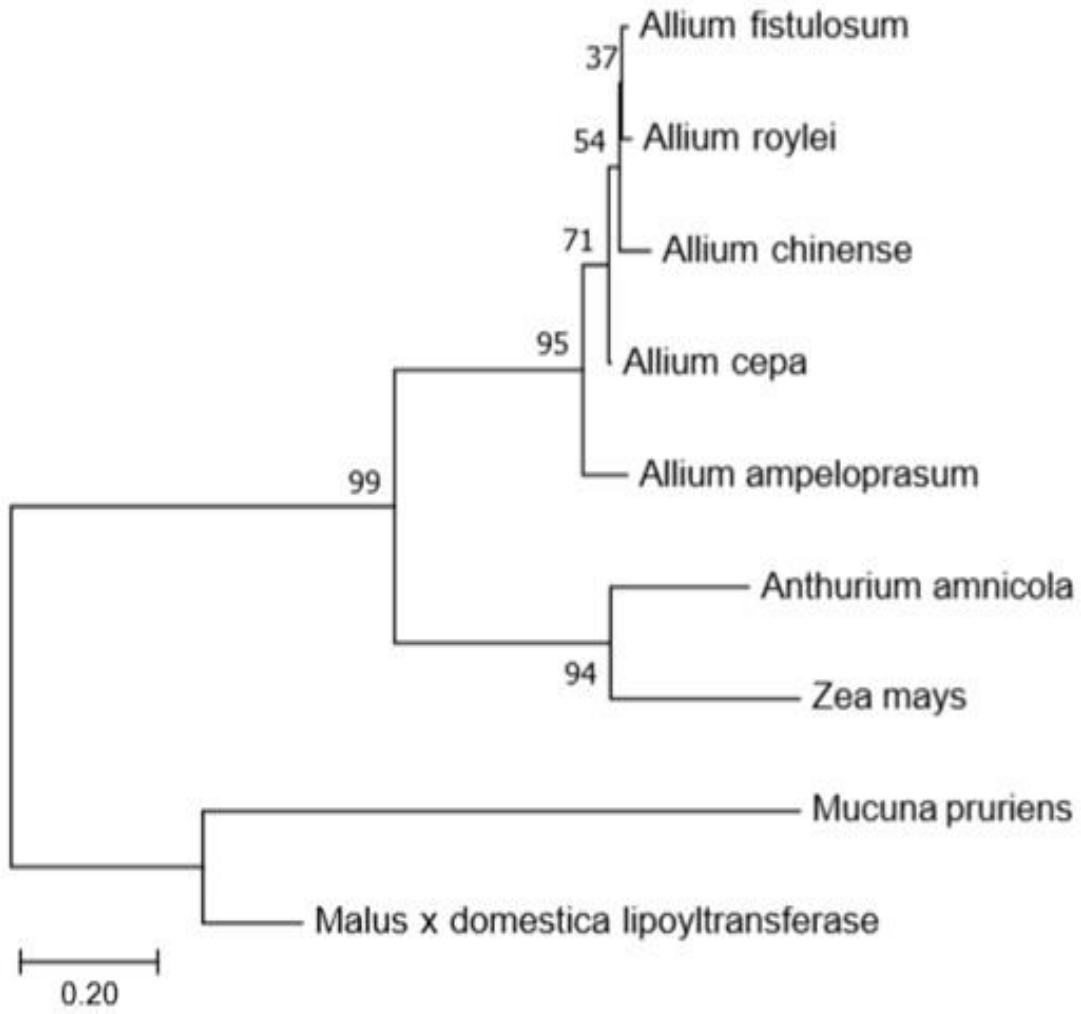


**Figure 3.4 Evolutionary relationships of LFS gene** The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 3.11067148 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. The analysis involved 9 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 177 positions in the final dataset. Evolutionary analyses were conducted in MEGA7. The scale symbolizes the nucleotide substitution rate.

## 3.4 qRT-PCR expression of yield and quality related genes in four different garlic varieties

The expression level of yield related genes FT-Like gene, Anthocyanidin Synthase (ANS), Lachrymatory Factor Synthase (LFS) and Sulphite Reductase (SIR) in four different varieties Local red, local white, NARC G1 and China white were examined with the help of Real-Time PCR analysis. The plant tissues selected for the expression of yield related genes were bulb, stem and leaves at complete bulbing stage. Threshold cycle (ΔCT) value shows the difference in the expression of yield related genes in different varieties. Results indicate that there was more expression of all the genes in tissues of variety China white, Minimum expression was observed in the local red and local white, while the expression of genes in NARC G1 was moderate. Tubulin was used as an endogenous positive control.

All the genes were expressed in all varieties under studies, The study carried out suggested that there was noticeable difference among the expression of genes in different varieties, FT like gene play an important role in the formation of bulbing , Anthocyanidin synthase (ANS) was more expressed in the Local red because of its red pigment, while Lachrymatory factor synthase (LFS) was expressed almost same in every variety.

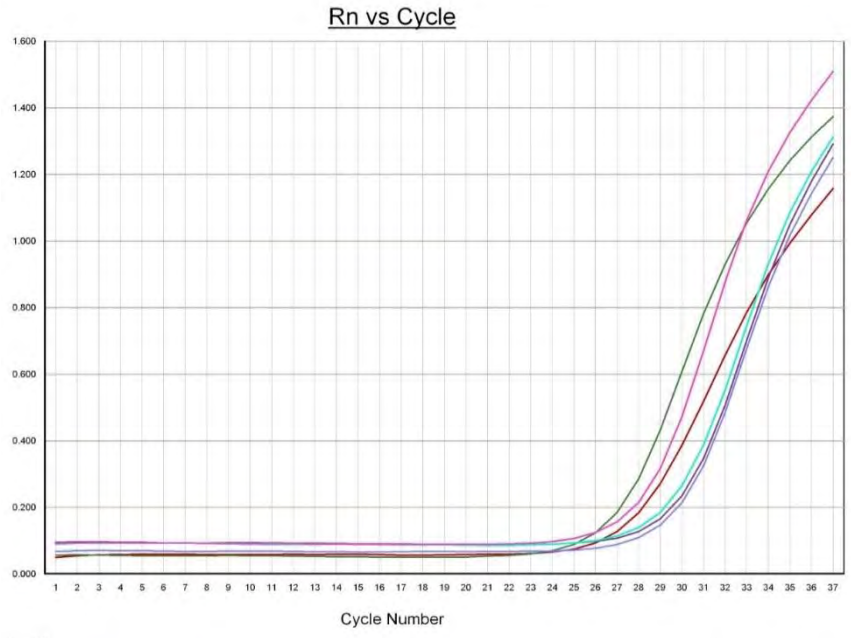

**Figure 3.5 Yield related genes expression by RT-PCR.** This figure shows the Amplification plot of different yield related genes in Local Red, Local white, China white and NARC G1 in its various tissues bulb stem and leaves. There is significant difference in the peaks at different escalating copy numbers at different time intervals

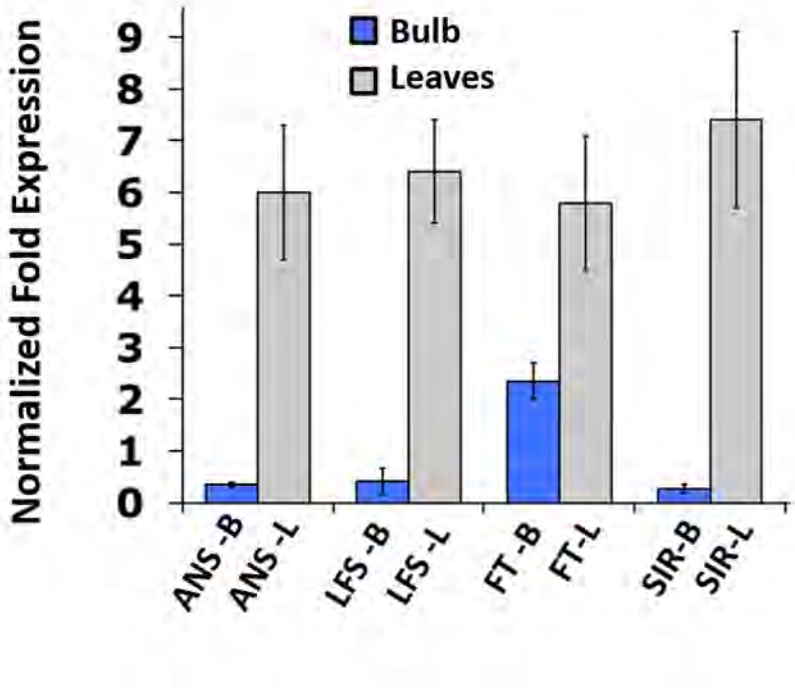


**Figure 3.6 the expression graph of different yield related genes in China white.** The bars on the graph represent the difference in genes expression in 2 different tissues. Tubulin was used as an endogenous positive control.

FT Bulb and SIR leaf there is higher level of yield related genes expression in china white and local white as compare to the other two varieties local red and NARC G1, FT-like genes were more expressed in all the varieties as compared to other gene. In A2 there is lower level of expression of yield related genes among different garlic varieties. ANS is least expressed in all the varieties, the results suggest that all the yield related genes have different expression among different garlic varieties. All the genes were least expressed in the local variety that is Local red.

### 3.5 Gel Electrophoresis of yield related genes in different varieties of garlic

The product obtained after completion of the real-time PCR were run on to the Agarose gel. The bands of yield related genes were observed with the help of ethidium bromide stained in 1% agarose gel (Appendix I). To determine the size and presence of

the gene 50bp ladder was used. The tray containing 1x TAE buffer was set on 90 volts for 40 minutes. The presence of band were confirmed after 40 minutes under Ultra Violet Trans-illuminator (UV trans-illuminator). Tubulin was used as an endogenous positive control gene.

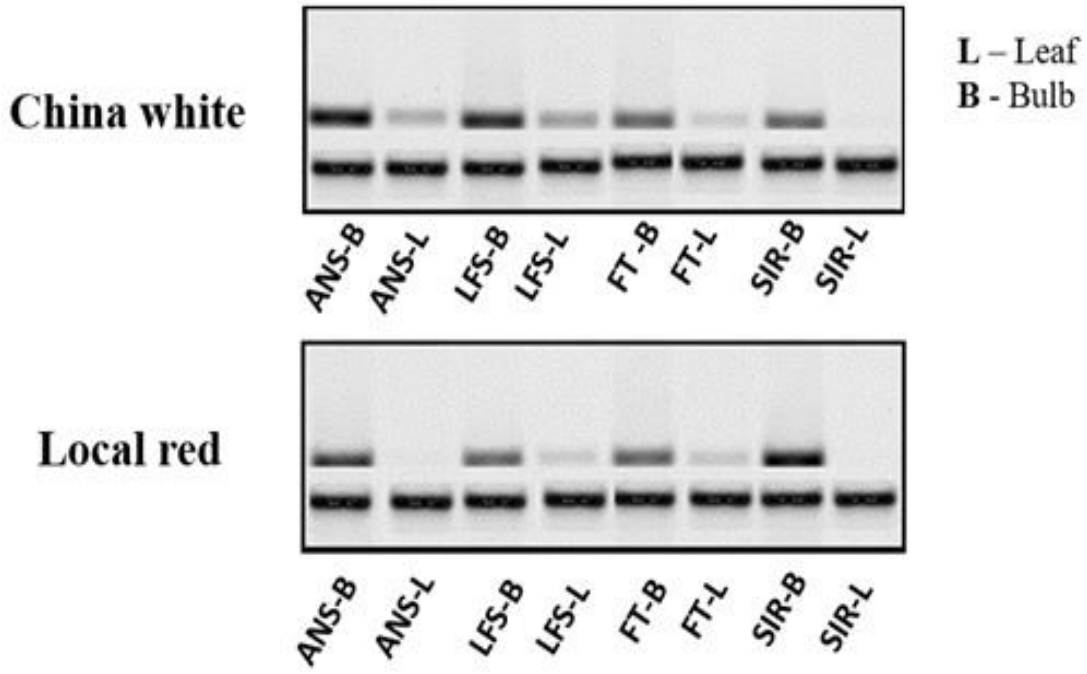


**Figure 3.7 Electropherogram for expression of yield related genes by RT-PCR.**

Expression of different yield related genes of both local and exotic varieties on the agarose gel. The bands emerged on the gel indicate the expression of each gene. There were also few unspecific bands found but all the target genes are denoted with a red line. In all the varieties there was differential genes expression found. These results indicate that the expression of FT-like gene was more in Local white and China white varieties ANS gene was rarely expressed in most of the garlic varieties. The expression of FSR was found more in both Local white and china white varieties. Similarly, the expression of LFS was lower in Local Red and NARC G1 varieties.

# 4 DISCUSSION

Garlic is the second most widely cultivated Allium species mainly grown in temperate region. Its current worldwide production is 24.9 million tons per annum, cultivated in 1.5 million hectares. Garlic gross production value for 2014 was US$15,129 million (FAO 2017). The sequencing of the bulb garlic genome and other Allium species has been hampered by their very large genome size (16 Gbps) and complex structure (Khosa et al. 2016). The bulb garlic genome is about 40 times larger than the rice genome (0.4 Gbps) (Jakse et al. 2008). Analysing the diversity of key gene families involved in plant architecture and adaptation within and among domesticated Allium and their wild allies may provide insights into the key genes under selection during domestication and dispersal (Ross-Ibarra et al. 2007).

Expression analysis of yield related genes with real-time PCR for FT-Like gene (FT), Anthocyanidin Synthase (ANS), Lachrymatory Factor Synthase (LFS) and Sulphite Reductase (SIR) (Lee et al. 2013) was further revealed by different gene expression found in two different tissues of the garlic cultivars. The expression of these genes present in both leaves and bulb of every variety. The expression of FT-Like gene was more in the leaves of each variety as compared to bulb tissues, results indicate that bulbing is activated by the Ft-like gene in the leaves while the formation of protein is carried out in the bulb portion. Due to which least expression was observed in the bulb tissues. Turck and Kardailsky reported similar results that FT-like genes are expressed in the leaves of the plant under favourable environmental condition and then the produced protein is transported to the lower portion of the bulb to initiate blossoming and storage organ formation (Turck et al. 2008; Kardailsky et al. 1999). Current study shows that there was more expression of ANS gene in the Local Red bulb tissues due to its dark red cloves colour, while least expression was observed in the NARC-G1 due to its yellowish clove membrane. It is also reported that Anthocyanidin synthase (ANS) gene is activated by the different transcriptional factors including, WDR (WD-repeat proteins), bHLH and R2R3-MYB. Genes involved in the biosynthesis of flavonoid are directly activated by WDR, bHLH and R2R3-MYB transcriptional factors (Xu et al. 2015). There are more than one key gene reported that is involved in the formation of the flavonoids including Chalcone isomerase (CHI), Chalcone synthase (CHS),

Dihydroflavonol 4 reductase (DRF) and ANS gens (anthocyanidin synthase) .if there is slightly mutation in these genes there will be formation of a different shade (Khar et al. 2008; Masuzaki et al. 2007).

There was presence of SiR gene in both the tissues of the each garlic genotype. The Expression was more observed in the Local Red due to its small clove size and pungent smell, least expression was observed in the NARC-G1 due to its large size. It is reported that there are multiple key genes in onions and Garlic that are involved in the synthesis of organo sulphur compound (Sun et al. 2016; Kamenetsky et al. 2015; McManus et al. 2012; Brewster 2008). SiR (sulphite reductase) is the key factor that control the production of sulphur compound, it help in the reduction of sulphite as a result sulphur compound is formed in the alliums (McManus et al. 2012; Takahashi et al. 2011).

Presently describe, yield related genes of the *Allium sativum* FT-Like gene, Anthocyanidin Synthase (ANS), Lachrymatory Factor Synthase (LFS) and Sulphite Reductase (SIR) gene when compared at the genomic level show similarity. Their phylogenetic analysis expresses an important correlation at the sequence level. The phylogenetic study indicates that FSR gene as compared to other genes have fever dynamicity which is balanced in the connection of genomics but bear a few duplications. All of the genes have undergone massive duplication during the evolution of the plant. In the view of previous literature, it is approved that these genes are playing a key role in the yield related parameters like bulb formation, taste, colour and plant developmental processes.

Bioinformatics and comparative study were carried out to reveal the genomic maintenance for the evolutionary and functional correlation between Yields related gens of the Garlic (*Allium sativum*) FT-Like gene (FT), Anthocyanidin Synthase (ANS), Lachrymatory Factor Synthase (LFS) and Sulphite Reductase (SIR) gene having functional and genetic conservation among them. The results showed significant difference in the studied attributes of the garlic varieties. A lot of evolutionary dynamic expression was found in those yield related genes among exotic and local species NARC G1 Local red, Local white, Advance line and china white. There was pattern of conservation among these yield related genes sequences were also found. The basic

relationship between these Allium species were clearly indicated by comparative study. A lot of syntenic patterns and conservation was found which shows that there is a lot of relevance in their sequence and function. Results indicated that there is duplication event among these species.

**Future perspective**

In future, further analysis of yield related genes is required. This study will be helpful to find all Yield related genes of the garlic (*Allium sativum*) in five different garlic varieties bulb and leaves tissues. Editing of these yield related genes will be possible with CRISPR-Case-9 to improve the quality and quantity of the garlic species. Bioinformatics, comparative study of these genes will be useful in future research work. The different expression patterns of yield and quality related genes are also helpful to study the nature of garlic plants and their pathways related to transcription and regulation. Further analysis of yield related genes may be possible which are involved in the regulation of different pathways related to aroma, taste, bulbing and flowering. It will be useful to use different computational tools to find more yield and quality related genes in Allium species. The expression of different genes will be involved in improving quality and yield can be further analyzed for future study purposes. This study will bring up more information about future research projects related improvement of local garlic varieties for better yield and good quality garlic. The comparison of Garlic (*Allium sativum*) with other plants genome will also provide helpful information in studying different pathways involved in improvement of current varieties.

# REFERENCES

Allison, G.L., Lowe, G.M., Rahman, K., 2012. Aged garlic extract inhibits platelet activation by increasing intracellular cAMP and reducing the interaction of GPIIb/IIIa receptor with fibrinogen. Life Sci. 91, 1275–1280.

Amarakoon, S., Jayasekara, D. 2017. A review on garlic (*Allium sativum* L.) as a functional food. Journal of Pharmacognosy and Phytochemistry, 6(6), 1777-1780. ISSN 2349-823 Bayili, R. G., Abdoul-Lati

Anonymous. Economic Survey. Government of Pakistan, Finance Division, Economic Adviser's Wing, Islamabad, Pakistan. 2018.

Arreola, R., Quintero-Fabián, S., López-Roa, R. I., Flores-Gutiérrez, E. O., Reyes-Grajeda, J. P., Carrera-Quintanar, L., & Ortuño-Sahagún, D. (2015). Immunomodulation and anti-inflammatory effects of garlic compounds. Journal of Immunology Research, 2015, 1–13. https:/ /doi.org/ 10.1155/2015 / 401630.

Assefa AG, Mesgina SH, Abrha YW. Effect of inorganic and organic fertilizers on the growth and yield of garlic crop (*Allium sativum* L.) In Northern Ethiopia. Journal of Agricultural Science. 2015; 7(4):80.

Aviello, G., Abenavoli, L., Borrelli, F., Capasso, R., Izzo, A.A., Lembo, F., Capasso, F., 2009. Garlic: empiricism or science. Natural product communications. 4, 1785–1796.

Balch PA. Prescription for Nutritional Healing, 3rd Ed. New York: Avery. 2000; pp. 97.

Baldwin S, Revanna R, Pither-Joyce M, Shaw M, Wright K, Thomson S, Moya L, Lee R, Macknight R, McCallum J (2013) Genetic analyses of bolting in bulb onion (*Allium cepa* L). Theor App Genet 127:535–547

Barba, F. J., Putnik, P., Bursać Kovačević, D., Poojary, M. M., Roohinejad, S., Lorenzo, J. M., & Koubaa, M. (2017). Impact of conventional and non-conventional processing on prickly pear (Opuntia spp.) and their derived products: From

preservation of beverages to valorization of by-products. Trends in Food Science & Technology, 67, 260–270. https://doi.org/10.1016/j.tifs.2017.07.012.

Bianchini, F., & Vainio, H. (2001). Allium vegetables and organosulfur compounds: Do they help prevent cancer? Environmental Health Perspectives, 109(9), 893–902.

Bisen, P. S., & Emerald, M. (2016). Nutritional and therapeutic potential of garlic and onion (Allium sp.). Current Nutrition & Food Science, 12(3), 190–199. https://doi.org/ 10.2174/1573401312666160608121954.

Borkowska, A., Knap, N., Antosiewicz, J., 2013. Diallyl trisulfide is more cytotoxic to prostate cancer cells PC-3 than to noncancerous epithelial cell line PNT1A: a possible role of p66Shc signaling axis. Nutr. Cancer 65, 711–717.

Brewster JL (2008) Onions and other vegetable Alliums. CABI, Oxfordshire, UK

Burba JL (2013a) Grupos Ecofisiológicos (GE), de ajos en Argentina y su equivalencia internacional. In 100 temas sobre producción de ajo. Mejoramiento genético y producción de semillas de ajo. Ediciones INTA, La Consulta, Mendoza, Argentina, vol 2, pp 8–16

Burba JL (ed) (2013b) 100 temas sobre producción de ajo. Ediciones INTA, La Consulta, Mendoza, Argentina, vol 1–5, 500 p

Bursać Kovačević, D., Maras, M., Barba, F. J., Granato, D., Roohinejad, S., Mallikarjunan, K., ... Putnik, P. (2018). Innovative technologies for the recovery of phytochemicals from Stevia rebaudiana Bertoni leaves: A review. Food Chemistry, 268, 513–521. https://doi.org/10.1016/j.foodchem.2018.06.091.

Capasso, A., 2013. Antioxidant action and therapeutic efficacy of *Allium sativum* L. Molecules 18, 690–700.

Cavagnaro PF, Galmarini CR (2007) Garlic. In: Kole, C (ed) Genome mapping and molecular breeding in plants, vegetables, vol 5. Springer, Berlin, 375, pp 349–365

Chan, J.Y., Yuen, A.C., Chan, R.Y., Chan, S.W., 2013. A review of the cardiovascular benefits and antioxidant properties of allicin. Phytother. Res. 27, 637–646.

CHEN, S., SHEN, X., CHENG, S., LI, P., DU, J., CHANG, Y., MENG, H. 2013. Evaluation of Garlic Cultivars for Polyphenolic Content and Antioxidant Properties. Plos One, 8, 11, e79730. http: //dx. doi. org/10. 1371/ journal. pone. 0079730

Colín-González, A.L., Santana, R.A., Silva-Islas, C.A., Chánez-Cárdenas, M.E., Santamaría, A., Maldonado, P.D., 2012. The antioxidant mechanisms underlying the aged garlic extract- and S-allylcysteine-induced protection. Oxid. Med. Cell Longev. 2012, 907162.

Dannesteter, J., 2003. AVESTA: VENDIDAD: Fargard 20: The origins of medicine. Translated from Sacred Books of the East, American Edition. The Christian Literature Company, New York. 1898, (online), Available at: www.avesta.org.

Das, I., Acharya, A., & Saha, T. (2012). Protective effect of garlic in skin cancer. In V. R. Preedy (Ed.). Handbook of diet, nutrition and the skin (pp. 300–317). Wageningen: Wageningen Academic Publishers.

Doebley JF, Gaut BS, Smith BD (2006) The molecular genetics of crop domestication. Cell 127:1309–1321

Doebley JF, Gaut BS, Smith BD (2006) The molecular genetics of crop domestication. Cell 127:1309–132

Durak I, Kavutcu M, Aytaç B, Avcı A, Devrim E, Özbek H, Öztürk HS. Effects of garlic extract consumption on blood lipid and oxidant/ antioxidant parameters in humans with high blood cholesterol. The Journal of Nutritional Biochemistry. 2004; 15(6):373-7.

FAO.2017.FAOSTAT.http://www.fao.org/faostat/en/#data

Fidelis, M.,Santos, J. S., Escher,G. B., VieiradoCarmo, M.,Azevedo, L.,Cristina daSilva, M., ... Granato, D. (2018). In vitro antioxidant and antihypertensive

compounds from camu-camu (Myrciaria dubia McVaugh, Myrtaceae) seed coat: A multivariate structure-activity study. Food and Chemical Toxicology, 120, 479–490. https://doi.org/10. 1016/j.fct.2018.07.043.

Fritsch RM, Friesen N (2002) Evolution, domestication and taxonomy. In: Rabinowitch HD, Currah L (eds) Allium crop science: recent advances. CABI, Wallingford, UK, pp 5–30

Fritsch RM, Friesen N (2002) Evolution, domestication and taxonomy. In: Rabinowitch HD, Currah L (eds) Allium crop science: recent advances. CABI, Wallingford, UK, pp 5–3

Giacometti, J., BursaćKovačević,D.,Putnik, P.,Gabrić,D., Bilušić,T.,Krešić, G.,...Režek Jambrak, A. (2018). Extraction of bioactive compounds and essential oils from mediterranean herbs by conventional and green innovative techniques: A review. Food Research International, 113, 245–262. https://doi.org/10.1016/j.foodres.2018. 06.036.

Gonzalez R, Soto V, Sance M, Camargo A, Galmarini CR (2009) Variability of solids, organosulfur compounds, pungency and health-enhancing traits in garlic (*Allium sativum* L.) cultivars belonging to different ecophysiological groups. J Agric Food Chem 57:10282–10288

Granato, D., Magalhães Carrapeiro, M.d., Fogliano, V., & van Ruth, S. M. (2016). Effects of geographical origin, varietal and farming system on the chemical composition and functional properties of purple grape juices: A review. Trends in Food Science & Technology, 52, 31-48

Guo YL (2013) Gene family evolution in green plants with emphasis on the origination and evolution of Arabidopsis thaliana genes. Plant J 73:941–951. https://doi.org/10.1111/tpj.12089

Hanelt P (1990) 'Taxonomy, evolution and history' in onions and allied crops. In: Rabinowitch HD, Brewster JL (eds) Botany, physiology, and genetics, vol 1. CRC Press, Boca Raton, pp 1–26

Hanelt P (1990) 'Taxonomy, evolution and history' in onions and allied crops. In: Rabinowitch HD, Brewster JL (eds) Botany, physiology, and genetics, vol 1. CRC Press, Boca Raton, pp 1–26

Hsu CY, Adams JP, Kim H, No K, Ma C, Strauss SH, Drnevich J, Vandervelde L, Ellis JD, Rice BM, Wickett N, Gunter LE, Tuskan GA, Brunner AM, Page GP, Barakat A, Carlson JE, dePamphilis CW, Luthe DS, Yuceer C (2011) FLOWERING LOCUS T duplication coordinates reproductive and vegetative growth in perennial poplar. Proc Natl Acad Sci USA 108:10756–10761

Huang Y, Gou J, Jia Z, Yang L, Sun Y, Xiao X, Song F, Luo K (2012) Molecular cloning and characterization of two genes encoding dihydroflavonol-4-reductase from populus trichocarpa. PLoS ONE 7(2): e30364. https:// doi.org/10. 1371/journal.pone.0030364

Iciek, M., Kwiecień, I., Włodek, L., 2009. Biological properties of garlic and garlic-derived organosulfur compounds. Environ. Mol. Mutagen. 50, 247–265.

Jakse J, Meyer JDF, Suzuki G, McCallum J, Cheung F, Town CD, Havey MJ (2008) Pilot sequencing of onion genomic DNA reveals fragmented transposable elements, low gene densities, and significant gene enrichment after methy filtration. Mol Genet Genomics 280:287–292

Jakse J, Meyer JDF, Suzuki G, McCallum J, Cheung F, Town CD, Havey MJ (2008) Pilot sequencing of onion genomic DNA reveals fragmented transposable elements, low gene densities, and significant gene enrichment after methy filtration. Mol Genet Genomics 280:287–292

Jakse J, Meyer JDF, Suzuki G, McCallum J, Cheung F, Town CD, Havey MJ (2008) Pilot sequencing of onion genomic DNA reveals fragmented transposable elements, low gene densities, and significant gene enrichment after methy filtration. Mol Genet Genomics 280:287–292

Johansson M, Staiger D (2015) Time to flower: interplay between photoperiod and the circadian clock. J Exp Bot 66:719–730. https://doi.org/10.1093/jxb/eru441

Kabir MA, Rahim MA, Majumder DAN. Effect of different levels of organic manures on the growth and yield of garlic. Journal of Agroforestry and Environment. 2011; 5(1):73-76.

Kamenetsky R, Faigenboim A, Mayer ES, Michael TB, Gershberg C, Kimhi S, Esquira I, Shalom SR, Eshel D, Rabinowitch HD, Sherman A (2015) Integrated transcriptome catalogue and organ-specific profiling of gene expression in fertile garlic (*Allium sativum* L.). BMC Genom 16:12

Kardailsky I, Shukla VK, Ahn JH, Dagenais N, Christensen SK, Nguyen JT, Chory J, Harrison MJ, Weigel D (1999) Activation tagging of the floral inducer FT. Science 286:1962–1965

Karlgren A, Gyllenstrand N, Källman T, Sundström JF, Moore D, Lascoux M, Lagercrantz U (2011) Evolution of the PEBP gene family in plants: functional diversification in seed plant evolution. Plant Physiol 156:1967–1977

Karuppiah, P., Rajaram, S., 2012. Antibacterial effect of *Allium sativum* cloves and Zingiber officinale rhizomes against multiple-drug resistant clinical pathogens. Asian Pac. J. Trop. Biomed. 2, 597–601.

Katzer G. Spices Pages: Garlic (*Allium sativum*, garlic). In: Purseglove, J.W. 1987. Tropical crops dicotyledons. Longman, Singapore publishers (Pvt.) Ltd. 2005.

Khar A, Jakse J, Havey MJ (2008) Segregations for onion bulb colors reveal that red is controlled by at least three loci. J Am Soc Hortic Sci 133:42–47

Khosa JS, McCallum J, Dhatt AS, Macknight R (2016b) Enhancing onion breeding using molecular tools. Plant Breed 135:9–20

Khosa JS, McCallum J, Dhatt AS, Macknight R (2016b) Enhancing onion breeding using molecular tools. Plant Breed 135:9–20

King JJ, Bradeen JM, Bark O, McCallum J, Havey MJ (1998) A low density genetic map of onion reveals a role for tandem duplication in the evolution of extremely large diploid genome. Theor Appl Genet 96:52–62

King JJ, Bradeen JM, Bark O, McCallum J, Havey MJ (1998) A low density genetic map of onion reveals a role for tandem duplication in the evolution of extremely large diploid genome. Theor Appl Genet 96:52–62

Kinoshita T, Ono N, Hayashi Y, Morimoto S, Nakamura S, Soda M, Kato Y, Ohnishi M, Nakano T, Inoue S, Shimazaki K (2011) FLOWERING LOCUST regulates stomatal opening. Curr Biol 21:1232–1238

Kloosterman B, Abelenda JA, Carretero-Gomez M, Oortwijn M, De Boer JM, Kowitwanich K, Horvath BM, Eck HJ, Smaczniak C, Prat S, Visser RGF, Bachem CWB (2013) Naturally occurring allele diversity allows potato cultivation in northern latitudes. Nature 495:246–250

Koca, I., & Tasci, B. (2016). Garlic as a functional food. Acta Horticulturae, (1143), 139–146. https://doi.org/10.17660/ActaHortic. 2016.1143.20.

Kopriva S (2006) Regulation of sulfate assimilation in Arabidopsis and beyond. Ann Bot 97:479–495

Krieger U, Lippman ZB, Zamir D (2010) The flowering gene SINGLE FLOWER TRUSS drives heterosis for yield in tomato. Nat Genet 42:459–463

Lee R, Baldwin S, Kenel F, McCallum J, Macknight R (2013) FLOWERING LOCUS T genes control onion bulb formation and flowering. Nat Commun 4:2884

Masuzaki S, Shigyo M, Yamauchi N (2006) Complete assignment of structural genes involved in flavonoid biosynthesis influencing bulb color to individual chromosomes of the shallot (*Allium cepa* L.). Genes Genet Syst 81:255–263

McCallum J (2007) Onion. In: Kole C (ed) Genome mapping and molecular breeding in plants, vegetables, vol 5. Springer, Berlin, 375 pp 331–347

McCallum J (2007) Onion. In: Kole C (ed) Genome mapping and molecular breeding in plants, vol 5. Springer, Heidelberg, Berlin, New York, pp 331–342

McManus MT, Joshi S, Searle B, Pither-Joyce M, Shaw M, Leung S, Albert N, Shigyo M, Jakse J, Havey MJ, McCallum J (2012) Genotypic variation in sulfur assimilation and metabolism of onion (*Allium cepa* L.) III. Characterization of sulfite reductase. Phytochem 83:34–42

Navarro C, Abelenda JA, Cruz-Oro E, Cuellar CA, Tamaki S, Silva J, Shimamoto K, Prat S (2011) Control of flowering and storage organ formation in potato by Flowering Locus T. Nature 478:119–122

Nicastro, H. L., Ross, S. A., & Milner, J. A. (2015). Garlic and onions: Their cancer prevention properties. Cancer Prevention Research (Philadelphia, PA), 8(3), 181–189. https://doi.org/10.1158/1940-6207.CAPR-14-0172.

Nile, S. H., Nile, A. S., Keum, Y. S., & Sharma, K. (2017). Utilization of quercetin and quercetin glycosides from onion (*Allium cepa* L.) solid waste as an antioxidant, urease and xanthine oxidase inhibitors. Food Chemistry, 235, 119–126. https://doi.org/10. 1016/j.foodchem.2017.05.043.

Omar, S.H., 2013. Garlic and cardiovascular diseases. Nat. Prod. 3661–3696.

Pin P, Nilsson O (2012) The multifaceted roles of FLOWERING LOCUS T in plant development. Plant Cell Environ 35:1742–1755

Poojary, M. M., Putnik, P., Bursać Kovačević, D., Barba, F. J., Lorenzo, J. M., Dias, D. A., & Shpigelman, A. (2017). Stability and extraction of bioactive sulfur compounds from Allium genus processed by traditional and innovative technologies. Journal of Food Composition and Analysis, 61, 28 –39

Prakash, D., Singh, B.N., &Upadhyay, G. (2007). Antioxidantand freeradical scavenging activities of phenols from onion (*Allium cepa* ). Food Chemistry, 102(4), 1389–1393.

Putnik, P., Lorenzo, J., Barba, F., Roohinejad, S., Režek Jambrak, A., Granato, D., ... Bursać Kovačević, D. (2018). Novel food processing and extraction technologies of high-added value compounds from plant materials. Foods, 7 https://www.mdpi.com/ 2304-8158/7/7/106.

Radulović, N. S., Miltojević, A. B., Stojković, M. B., & Blagojević, P. D. (2015). New volatile sulfur-containing compounds from wild garlic (Allium ursinum L., Liliaceae). Food Research International, 78,1 –10. https://doi.org/10.1016/j.foodres.2015.11. 019.

Ramirez, D. A., Locatelli, D. A., González, R. E., Cavagnaro, P. F., & Camargo, A. B. (2017). Analytical methods for bioactive sulfur compounds in Allium: An integrated review and future directions. Journal of Food Composition and Analysis, 61,4 –19. https://doi.org/10.1016/j.jfca.2016.09.012.

Ricroch A, Yockteng R, Brown S, Nadot S (2005) Evolution of genome size across some cultivated Allium species. Genome 48:511–520

Ried, K., Frank, O.R., Stocks, N.P., 2013. Aged garlic extract reduces blood pressure in hypertensives: a dose-response trial. Eur. J. Clin. Nutr. 67, 64–70.

Rosello-Soto, E., Barba, F. J., Putnik, P., Bursac Kovacevic, D., Lorenzo, J. M., & Cantavella-Ferrero, Y. (2018). Enhancing bioactive antioxidants' extraction from “Horchata de Chufa” by-products. Foods, 7(10), https://doi.org/10.3390/ foods7100161

Ross-Ibarra J, Morrell PL, Gaut BS (2007) Plant domestication, a unique opportunity to identify the genetic basis of adaptation. Proc Natl Acad Sci 104(suppl 1):8641–8648

Ross-Ibarra J, Morrell PL, Gaut BS (2007) Plant domestication, a unique opportunity to identify the genetic basis of adaptation. Proc Natl Acad Sci 104(suppl 1):8641–8648

Saljoughian, S., Roohinejad, S., Bekhit, A. E.-D. A., Greiner, R., Omidizadeh, A., Nikmaram, N., & Mousavi Khaneghah, A. (2017). The effects of food essential

oils on cardiovascular diseases: A review. Critical Reviews in Food Science and Nutrition, 1–18. https://doi.org/10.1080/10408398.2017.1279121.

Sanchez SE, Kay SA (2016) The plant circadian clock: from a simple timekeeper to a complex developmental manager. Cold Spring Harb Perspect Biol 8:12. https:// doi.org/10.1101/cshperspect. a027748

Santas, J., Carbo, R., Gordon, M., & Almajano, M. (2008). Comparison of the antioxidant activity of two Spanish onion varieties. Food Chemistry, 107(3), 1210–1216. https:// doi.org/10.1016/j.foodchem.2007.09.056.

Santhosha, S.G., Jamuna, P., Prabhavathi, S.N., 2013. Bioactive components of garlic and their physiological role in health maintenance: a review. Food Biosci. 3, 59–74.

Sawa M, Nusinow DA, Kay SA, Imaizumi T (2007) FKF1 and GIGANTEA complex formation is required for day-length measurement in Arabidopsis. Science 318:261–265

Schwinn KE, Ngo H, Kenel F, Brummell DA, Albert NW, McCallum JA, Pither-Joyce M, Crowhurst NS, Eady C, Davies KM (2016) The Onion (*Allium cepa* L.) R2R3-MYB Gene MYB1 regulates anthocyanin biosynthesis. Front Plant Sci 7:1865. https://doi.org/10.3389/fpls.2016.01865

Shalom SR, Gillett D, Zemach H, Kimhi S, Forer I, ZutahyY,TamY,Teper-Bamnolker P, Kamenetsky R, Eshel D (2015) Storage temperature controls the timing of garlic bulb formation via shoot apical meristem termination. Planta 242:951–962

Sohn SH, Ahn YK, Lee TH, Lee JE, Jeong MH, Seo CH, Chandra R, Kwon YS, Kim CW, Kim DS, Won SY, KimJS,ChoiD(2016) Construction of a draft reference transcripts of onion (*Allium cepa* ) using long-read sequencing. Plant Biotechnol Rep 10:383–39

Sohn SH, Ahn YK, Lee TH, Lee JE, Jeong MH, Seo CH, Chandra R, Kwon YS, Kim CW, Kim DS, Won SY, Kim J S, Choi D (2016) Construction of a draft reference transcripts of onion (*Allium cepa* ) using long-read sequencing. Plant Biotechnology Rep 10:383–390

Srivastava R, Agarwal A, Tiwari RS, Kumar S. Effect of micronutrients, zinc and boron on yield, quality and storability of garlic (*Allium sativum*). Indian Journal of Agricultural Science. 2005; 75(3):157-9

Stabler, S.N., Tejani, A.M., Huynh, F., Fowkes, C., 2012. Garlic for the prevention of cardiovascular morbidity and mortality in hypertensive patients. Cochrane Libr. 8.

Sun XD, Yu XH, Zhou SM, Liu SQ (2016) De novo assembly and characterization of the Welsh onion (Allium fitulosum L.) transcriptome using Illumina technology. Mol Genet Genomics 291:647–659

Takahashi H, Kopriva S, Giordano M, Saito K, Hell R (2011) Sulfur assimilation in photosynthetic organisms: molecular functions and regulations of transporters and assimilatory enzymes. Ann Rev of Plant Biol 62:157–184

Taylor A, Massiah AJ, Thimas B (2010) Conservation of Arabidopsis thaliana photoperiodic flowering genes in onion (*Allium cepa* L.). Plant Cell Physiol 51:1638– 1647

Tepe, B., SOKMEN, M., AKPULAT , HA., SOKMEN, A. 2005. In vitro antioxidant activities of the methanol extracts of five Allium species from Turkey. Food Chemistry. 92, 89-92. http://dx.doi.org/10.1016/j.foodchem.2004.07.016

Tsubura, A., Lai, Y.C., Kuwata, M., Uehara, N., Yoshizawa, K., 2011. Anticancer effects of garlic and garlic-derived compounds for breast cancer control. Anti-Cancer Agents Med. Chem. 11, 249–253.

Turck F, Fornara F, Coupland G (2008) Regulation and identity of florigen: Flowering Locus T moves center stage. Annu Rev Plant Biol 59:573–594

Van Der Meer QP (1997) Old and new crops within edible Alliums. Acta Hort 433:17–31

Varshney RK, Terauchi R, McCouch SR (2014) Harvesting the promising fruits of genomics: applying genome sequencing technologies to crop breeding. PLoS Biol 12:e1001883

Wallace IV, G.C., Haar, C.P., Vandergrift III, W.A., Giglio, P., Ray, S.K., Patel, S.J., Das, A., 2013. Multi-targeted DATS prevents tumor progression and promotes apoptosis in an animal model of glioblastoma via HDAC-inhibition. J. Neuro-oncol. 114, 43–50.

Wang, B.-S., Huang, G.-J., Lu, Y.-H., & Chang, L.-W. (2013). Anti-inflammatory effects of an aqueous extract of Welsh onion green leaves in mice. Food Chemistry, 138(2–3), 751–756. https://doi.org/10.1016/j.foodchem.2012.11.106.

Wang, H.C., Pao, J., Lin, S.Y., Sheen, L.Y., 2012. Molecular mechanisms of garlic-derived allyl sulfides in the inhibition of skin cancer progression. Ann. NY Acad. Sci. 1271, 44–52.

Wendel JF, Jackson SA, Meyers BC, Wing RA (2016) Evolution of plant genome architecture. Genome Biol 17:37. https://doi.org/10.1186/s13059-016-0908-1

Wendel JF, Jackson SA, Meyers BC, Wing RA (2016) Evolution of plant genome architecture. Genome Biol 17:37. https://doi.org/10.1186/s13059-016-0908-1

Wickland DP, Hanzawa Y (2015) The FLOWERING LOCUS T/TERMINAL FLOWER 1 gene family: Functional evolution and molecular mechanisms. Mol Plant 8:983–997

Xu G, Kong H (2007) Duplication and divergence of floral MADS-box genes in grasses: evidence for the generation and modification of novel regulators. J Integr Plant Biol 49:927–939

Zain al-abdeen, S., Abdullah, I.T., Al-Salihi, S., 2017. The synergism effect of aqueous garlic extract and ciprofloxacin against some multi-resistant bacteria. J. Microbiol. Biotechnol. Res. 3, 136–142.

Zeng, Y., Li, Y., Yang, J., Pu, X., Du, J., Yang, X., ... Yang, S. (2017). Therapeutic role of functional components in Alliums for preventive chronic disease in human being. Evidence-Based Complementary and Alternative Medicine, 2017,1 –13. https://doi.org/ 10.1155/2017/9402849.

Zheng XM, Wu FQ, Zhang X, Lin QB, Wang J, Guo XP, Lei CL, Cheng ZJ, Zou C, Wan JM (2016b) Evolution of the PEBP gene family and selective signature on FT-like clade. J Syst Evol 54:502–510

Zhu, Z., Li, S., He, J., Thirumdas, R., Montesano, D., & Barba, F. J. (2018). Enzymeassisted extraction of polyphenol from edible lotus (Nelumbo nucifera) rhizome knot: Ultra-filtration performance and HPLC-MS2 profile. Food Research International, 111, 291–298. https: // doi. org/10. 1016/j. foodres. 2018. 05.047.

# APPENDICES

**Appendix-I**

**10X TAE Buffer Composition**

| Reagent | Quantity |
|---|---|
| Tris-Base | 242g |
| Glacial Acetic Acid | 57.1g |
| 0.5 EDTA | 100ml |

**1X TAE Buffer**

| Reagents | Quantity |
|---|---|
| 50X TAE | 20ml |
| Distilled Water | 980ml |

**Appendix-II**

**Synthesis of first strand cDNA**

Total RNA extracted was used as a template for synthesis of cDNA according to the manufacturer instructions using the Reverse Transcriptase III Kit (Fermentas).

PCR mixture performed in 25µl reaction volume. Following reagents were used for the synthesis of first strand cDNA:

4µl 10X Reverse Transcriptase Buffer

3µl 10nM dNTPs mixture

1µl gene specific primer

4µl RNA

1µl Reverse Transcriptase

12µl double nuclease free water

PCR conditions were used as follows:

65ºC for 5 minutes

Chilled sample for 1 minute

55ºC for 60 minutes

72ºC for 10 minutes

Final hold 4 ºC for ∞

**Appendix-III**

**Quantitative real time –PCR**

To determine the expression of yield related genes in *Allium sativum* different tissues, Real Time PCR was used using SYBER green master mix.

For 25µl reaction PCR reagents detail are as follow:

| | |
|---|---|
| SYBER Green Master mix | 12 |
| Forward primer | 0.8 |
| Reverse primer | 0.8 |
| DNA template | 3 |
| Double distilled water | up to 25 |

For amplification of yield related genes with gene specific primer PCR profile is as follow:

Initial denaturation at 94ºC for 10 minutes

Denaturation at 95ºC for 40 seconds

Annealing at 58ºC for 32 seconds 37 cycles

Extension at 72ºC for 32 seconds

Final extension at 95ºC for 15 seconds

Melt curve analysis at 52ºC to 95ºC

Final hold 4 ºC for ∞

## Appendix-IV

### *Allium cepa* clone Pool1 FT-like protein gene, complete cds GenBank: KC677631.1

>KC677631.1:5-202,843-904, 2857-2897, 2969-3204 *Allium cepa* clone Pool1 FT-like protein gene, complete cds

ATGGCAAGAGAAAGTGACCCATTAATTGTAGGTAGAATAGTTGGTGATGTAATATCTCCATTCACAAGAA
GTGTTCCTTTAAGGGTAATATACCCAGCAAAAGAGGTAGCCAATGGGCGTGAGTTTAAACCATCACAAAT
AACTCAGCAACCTAGAGTTGAGATCGGAAGTGATGATCTCAGAACCTTCTATACGCTAGTGATGGTGGAC
CCAGACGCCCCTAGCCCAAGTAATCCACACCTTAGGGAATACTTGCATTGGATGGTGTCAGACATTCCAG
GAGGTACAGGATCGAACTTTGGGCGAGAAACCGTCTGCTATGAAAGCCCGAGACCAACAGCTGGTATCCA
CCGTTTCGTTTTTATCCTGTTCCAGCAGCTTGGGCGCCAAACTGTATATGCACCGAACTGGCGGCAAAAC
TTTAATACTCGAGAATTTGCAGAAAATTACAACCTTGGGTCTCCAGTTGCTGCAGTTTACTTCAACTGCC
AGAGAGAATCAGGATCTGGTGGGAGGAGAATATACACGGATTACTAA

### *Allium sativum* flowering locus T mRNA, complete cds GenBank: KP711810.1

>KP711810.1 *Allium sativum* flowering locus T mRNA, complete cds

ATGATGGATTCGGATCCGTTAAGGTTGGGTAGAATAGTAGGTGATGTTATAGACCCGTTTACCAGAAGGG
TGTCGCTTAGGGCTGTCTACTCATGCAGAGAAGTTGCAAATGGACGCGAGTTTAGGCCTTCTCAAGTTGC
CCTACAGCCAAGAATTGAAATTGGCGGCGGTGACCTTAGAAACTCCTATGCACTTGTAATGGTGGACCCA
GATGCTCCAAGCCCAAGCAATCCCCATCTAAGAGAATACTTGCACTGGTTGGTCACAGACATTCCAGGAA
GCACAAGTGCAAGCTTCGGTCAAGAAAGGATATGCTATGAAAGTCCAAGGCCATCCTTAGGAATACACAG
GTTTGCCTTCATATTATTTCAGCAGCTTGGTCGCGAGACCGTGTGCTCTCCAAACTATAGGCAGAACTTC
AGCTCCAGAGTTTTTGCAGAAGTATACAACTTGGGTTCTCCTGTTGCTGCTCTTTATTTCAACTGCCAGA
GAGAAGCTGGTTCAGGTGGGAGGAGAACTTATACATGA

### *Allium cepa* anthocyanidin synthase (ANS) gene, ANS-L allele, and complete cds 1071bp GenBank: EF192598.1

>EF192598.1:3090-3586, 3694-4267 *Allium cepa* anthocyanidin synthase (ANS) gene, ANS-L allele, complete cds

ATGCCCATAGAAAGCGTGATAATTGCACCACCAGCTCCAAGAGTAGAGACCCTCTCGAAATCCAACCTCC
ACTCCATTCCTTTGGAATACATCCGGCCAGAGCACGAGCGGGCCTGCCTAGGCGATGCATTAGAACAGCT
TCACAACTCCAACTCCGGCCCACAAATTCCGATAACCGACCTCGACTCACGCGACTGCATTGAGAAAGTG
ACCAAAGCAGCCAAAGAATGGGGAGTCATGCACATTGTGAACCATGGAATCTCTAGCGAGTTGATGGAGA

AAGTGCGTGCAGCTGGCAAGGCTTTTTTCAATCTCCCACTCGAGGCTAAAGAGGAATATGCTAACGATCA
ATCTAAAGGGAAGATCCAAGGGTATGGAAGTAAACTGGCAAATAATGCAAGCGGGCAACTCGAGTGGGAG
GACTACTTTTTTCATCTCATTTTTCCTGATGACAAGGTCGATCTCTCCGTTTGGCCTAAACAGCCTTCCG
ACTACATTGAAATTATGCAAGAGTTTGGAAGTCAGCTGAGAATATTAGCCAGCAAAATGCTATCCATACT
TTCATTGGGCTTACAACTACCAACCAAGGACAGGCTAGAACAAGAACTAAAAGGACCAGAAGACTTACTT
CTCCAGCTGAAAATAAACTACTACCCAAAATGCCCCCAGCCACATCTGGCATTAGGAGTGGAAGCCCACA
CGGACGTGAGTGCCTTATCGTTCATACTGCACAACAACGTTCCAGGCCTTCAAGTGTTGTACGAGGGTGA
ATGGGTAACTGCGAAACTCGTTCCAGATTCCTTGATTGTCCATGTTGGGGATTCGCTTGAAATTTTGAGT
AATGGGATTTACAAAAGTGTGTTGCATCGCGGTTTGGTTAATAAGGAGAAGGTGAGGATTTCTTGGGCTG
TGTTCTGTGAACCGCCAAAGGATGCTGTTGTGCTTAAGCCATTGGATGAGGTTGTGACAGACGATGCGCC
CGCAAGGTACACTCCTCGCACGTTTGCACAGCATCTGGAACGGAAGTTGTTTAAAAAGAAAGTTGGAGAT
TTGGATGACTCTGATGTTTGA

## *Allium cepa* anthocyanidin synthase (ANS) gene, ANS-h1 allele, complete cds 1071 bp GenBank: EF192475.2

>EF192475.2:3505-4001,4108-4681 *Allium cepa* anthocyanidin synthase (ANS) gene, ANS-h1 allele, complete cds

ATGACCATAGAAAGCGTGATAATTGCACCACCAGCTCCAAGAGTAGAGACCCTCTCGAAATCCAACCTCC
ACTCCATCCCTCTGGAATACATCCGGCCCGAGCACGAGCGGGCCTGTCTAGGCGATGCATTAGAACAGCT
TCACAACTCCAACTCCGGCCCACAAATTCCGATAACCGACCTCGATTCAAGCGACTGCATTGAGAAAGTG
ACCAAAGCAGCCAAAGAATGGGGAGTCATGCACATTGTGAACCATGGAATCTCTAGCGAGTTGATGGAGA
AAGTGCGTGCAGCTGGCAAGGCTTTTTTCAATCTCCCACTCGAGGCTAAAGAGGAATATGCTAACGATCA
ATCTAAAGGGAAGATCCAAGGGTATGGAAGTAAACTGGCAAATAATGCAAGCGGGCAACTCGAGTGGGAG
GACTACTTTTTTCATCTCATTTTTCCTGATGACAAGGTCGATCTCTCCGTTTGGCCTAAACAGCCTTCCG
ACTACATTGAAATTATGCAAGAGTTTGGAAGTCAGCTGAGAATATTAGCCAGCAAAATGCTATCCATACT
TTCATTGGGCTTACAACTACCAACCAAGGACAGGCTAGAACAAGAACTAAAAGGACCAGAAGACTTACTT
CTCCAGCTGAAAATAAACTACTACCCAAAATGCCCCCAGCCACATCTGGCATTAGGAGTGGAAGCCCACA
CGGACGTGAGTGCCTTATCGTTCATACTGCACAACAACGTTCCAGGCCTTCAAGTGTTGTACGAGGGTGA
ATGGGTAACTGCGAAACTCGTTCCAGATTCCTTGATTGTCCATGTTGGGGATTCGCTTGAAATTTTGAGT
AATGGGATTTACAAAAGTGTGTTGCATCGCGGTTTGGTTAATAAGGAGAAGGTGAGGATTTCTTGGGCTG
TGTTCTGTGAACCGCCAAAGGATGCTGTTGTGCTTAAGCCATTGGATGAGGTTGTGACAGACGATGCGCC
CGCAAGGTACACTCCTCGCACGTTTGCACAGCATCTGGAACGGAAGTTGTTTAAAAAGAAAGTTGGAGAT
TTGGATGACTCTGATGTTTGA

***Allium cepa* genotype dark red F3 line anthocyanidin synthase gene, complete cds**

**GenBank: AY585678.1**

>AY585678.1:1-497,605-1166 *Allium cepa* genotype dark red F3 line anthocyanidin synthase gene, complete cds

ATGCCCATAGAAAGCGTGATAATTGCACCACCAGCTCCAAGAGTAGAGACCCTCTCGAAATCCAACCTCC
ACTCCATTCCTTTGGAATACATCCGGCCAGAGCACGAGCGGGCCTGCCTAGGCGATGCATTAGAACAGCT
TCACAACTCCAACTCCGGCCCACAAATTCCGATAACCGACCTCGACTCACGCGACTGCATTGAGAAAGTG
ACCAAAGCAGCCAAAGAATGGGGAGTCATGCACATTGTGAACCATGGAATCTCTAGCGAGTTGATGGAGA
AAGTGCGTGCAGCTGGCAAGGCTTTTTTCAATCTCCCACTCGAGGCTAAAGAGGAATATGCTAACGATCA
ATCTAAAGGGAAGATCCAAGGGTATGGAAGTAAACTGGCAAATAATGCAAGCGGGCAACTCGAGTGGGAG
GACTACTTTTTTCATCTCATTTTTCCTGATGACAAGGTCGATCTCTCCGTTTGGCCTAAACAGCCTTCCG
ACTACATTGAAATTATGCAAGAGTTTGGAAGTCAGCTGAGAATATTAGCCAGCAAAATGCTATCCATACT
TTCATTGGGCTTACAACTACCAACCAAGGACAGGCTAGAACAAGAACTAAAAGGACCAGAAGACTTACTT
CTCCAGCTGAAAATAAACTACTACCCAAAATGCCCCCAGCCACATCTGGCATTAGGAGTGGAAGCCCACA
CGGACGTGAGTGCCTTATCGTTCATACTGCACAACAACGTTCCAGGCCTTCAAGTGTTGTACGAGGGTGA
ATGGGTAACTGCGAAACTCGTTCCAGATTCCTTGATTGTCCATGTTGGGGATTCGCTTGAAATTTTGAGT
AATGGGATTTACAAAAGTGTGTTGCATCGCGGTTTGGTTAATAAGGAGAAGGTGAGGATTTCTTGGGCTG
TGTTCTGTGAACCGCCAAAGGATGCTGTTGTGCTTAAGCCATTGGATGAGGTTGTGACAGACGATGCGCC
CGCAAGGTACACTCCTCGCACGTTTGCACAGCATCTGGAACGGAAGTTGTTTAAAAAGAAAGGTTGGAGA
TTTGGATGA

***Allium cepa* cultivar CUDH2107 ferredoxin sulfite reductase mRNA, complete cds 1,890 bp linear mRNA GenBank: JX020763.1**

>JX020763.1 *Allium cepa* cultivar CUDH2107 ferredoxin sulfite reductase mRNA, complete cds

ATGGAAGCGATGGCGGCGACTGCGATGGCAAAGGATCCGACGGGGCAGATGATGCTGAGTGGATTTCAAG
GGCTCAGATCGGTTGGGTCGGTTCCGGTTGGAAGGTCCGTTAAGGCGTTGCCGATTCCGTCTTCTTTTTC
TTCTTCCGTTATTACGGCTGTAGCTGTTCCTGAGGTCAAAAGAAGCAAGGTAGAAATCTTTAAAGAACAG
AGCAACTTTCTCAGATATCCACTAAATGAGGAGTTGTTAACTGATGCACCAAATATAAACGAAGCTTCCA
CCCAGCTAATAAAATTCCATGGTAGTTATCAGCAATACGACAGAGATGAACGTGGTAAAAGATCTTATTC
GTTTATGCTTCGAACAAAGAATCCATGCGGAAAAGTCCCAAACGAGTTGTACATAGCTATGGATACTCTT
GCTGATGAATTTGGAATTGGAACCCTTCGGTTAACCACGAGACAGACATTTCAATTGCATGGCATTTTAA
AGAAGAACCTCAAAACTGTGATGAGTACTATTATTCATAATATGGGGTCGACTCTTGGGGCTTGTGGGGA
TTTGAACCGAAATGTTCTTGCACCTGCTGCACCATACACTAAAAAGGAGTATGTTTTTGCTCAAGAGACC

GCAGATAATATTGCTGCTCTTCTTACACCTCAATCAGGTTTTTACTATGATATGTGGGTAGATGGTGAGA
AAATCATGACTTCAGAGCCACCCGAAGTTACAAAAGCAAGAAATGACAATTCTCATGGCACTAATTTTCC
TGATTCACCCGAGCCAATTTATGGCACTCAATTTTTGCCACGAAAGTTCAAAATTGCAGTCACTGTGCCA
ACTGACAACTCAGTTGATATCTTAACCAATGATATTGGTCTCGTCGTAATATCTGATAGTAATGGGGAGC
CTCAGGGTTTTAATCTCTATGTTGGTGGAGGAATGGGAAGAGCGCATAGAACTGACACCACTTTTCCTCG
TTTGGCTGAGCCATTGGGTTATGTACCAAAAGAAGATATATTATATGCTATCAAAGCAATTGTCGTTACC
CAAAGAGAAAATGGAAGAAGAGATGACCGTAAATATAGCAGAATGAAGTATCTGATTAGTGCATGGGGAA
TTGAGAAATTCCGTGATGTTGTTGAGCAATATTATGGAAAGAAATTTGAGGCATTCCGTGAGTTACCTGA
ATGGGAGTTTAGGAGTTATCTAGGTTGGCATGAACAGGGCAGCGAGAAATTGTTCTGTGGCCTACACATT
GATAATGGTCGTCTTAAGGGCCAAGCAAAGAAGACTTTAAGAGAAATTATTGAAAAGCACAACCTGAGTG
TGAGAATCACCCCAAATCAAAACATGATTTTGTGTGACATTGATCCTTCCTGGAAAGAGTCTATCAATGC
AGCTCTTTCTCAAGTTGGTATTCTGGAACCTCAATACGTAGATCCTCTGAATATCACATCTATGGCATGT
CCTGCTCTGCCTCTATGTCCACTAGCAATTGCAGAAGCTGAGAGAGGGATACCTGATATTCTCAAAAGGG
TGCGAGCAGTTATTGATAAGGTAGGTATGGGGAAAGAAGAATCAGTGGTGATTAGGATAACTGGTTGTCC
TAATGGCTGTGCCAGACCTTACATGGCTGAGCTAGGGTTTGTTGGTGATGGCCCAAATAGCTACCAGATA
TGGCTTGGAGGAACACCTAACCAGAGTACACTAGCGAAGTGTTTTATGAATAAGGTGAAGATCCAAAAAT
TTGAGAGCGTTTTAGAACCGCTGTTCAATGACTGGAAAGTAAATCGCAAGACAGAAGAATCATTTGGTGA
ATTTACAAACCGAATTGGCTTTGAAAAGCTGCTTGAGGTTGTAGAGCAGTGGGATAGTTCTAAAAATTAA

### *Allium cepa* var. aggregatum lfs mRNA for lachrymatory factor synthase, complete cds GenBank: AB094593.1

>AB094593.1:55-564 *Allium cepa* var. aggregatum lfs mRNA for lachrymatory factor synthase, complete cds

ATGGAGCTAAATCCTGGTGCACCTGCTGTAGTCGCTGATAGTGCTAACGGAGCTCGAAAATGGAGCGGCA
AAGTCCATGCTTTGCTTCCAAATACAAAGCCAGAGCAAGCATGGACACTACTAAAAGACTTTATTAACCT
TCACAAGGTCATGCCTTCGTTGTCAGTCTGTGAACTGGTAGAAGGTGAGGCCAATGTTGTTGGTTGTGTT
CGCTACGTTAAAGGTATAATGCACCCAATAGAAGAGGAATTTTGGGCCAAGGAGAAGCTGGTGGCGCTGG
ATAATAAGAACATGAGCTACAGTTATATTTTTACTGAGTGTTTTACAGGGTACGAGGATTACACGGCTAC
CATGCAAATAGTGGAGGGTCCTGAGCACAAGGGAAGTAGATTTGACTGGTCTTTTCAGTGCAAGTATATC
GAGGGTATGACTGAATCTGCATTCACCGAGATTCTGCAGCATTGGGCTACTGAGATAGGTCAGAAAATCG
AAGAGGTTTGCAGTGCTTGA

**Allium chinense lfs mRNA for lachrymatory factor synthase, complete cds GenBank: AB094592.1**

>AB094592.1:54-563 Allium chinense lfs mRNA for lachrymatory factor synthase, complete cds

ATGGAGCAAAATTCTGGTACGCTTGCTGTAGTCACTGATGGTGCTAAAGGAGCTGCAAAATGGAGAGGCA
AAGTCCATGCTTTGCTTCCAAATACAAAACCAGAGCAAGCATGGACACTACTAAAAGACTTTATTAACCT
TCACAAGATCATGCCTTCGTTGTCAGTTTGTGAACTGGTTGAAGGCGAGGCCAATGTTGTTGGTTGTGTT
CGCCACGTTAAAGGTATAATGCACCCAATGGAAGAGGAATTTTGGGCCAAGGAGAAGCTGGTTGCACTGG
ATGATAAGAACATGAGCTGTAGTTATATTTTTGTTGAGTGTTTTACAGGGTACGAGGATTACACAGCTAC
CATGCAAATAGTGGAGGGATCTGAGCACAAGGGATGTAGATTTGACTGGTCTTTTCAGTGTAAGTATATC
GAGGGTATGACTGAATCTGCATTCACCGATGTTCTGCAGCATTGGGCTACTGAGATTGGTCAGAAAATTG
AAGAGATTTGCAATGCTTGA

***Allium cepa* lfs mRNA for lachrymatory factor synthase, complete cds GenBank: AB089203.1**

>AB089203.1:53-562 *Allium cepa* lfs mRNA for lachrymatory factor synthase, complete cds

ATGGAGCTAAATCCTGGTGCACCTGCTGTAGTCGCTGATAGTGCTAACGGAGCTCGAAAATGGAGCGGCA
AAGTCCATGCTTTGCTTCCAAATACAAAGCCAGAGCAAGCATGGACACTACTAAAAGACTTTATTAACCT
TCACAAGGTCATGCCTTCGTTGTCAGTCTGTGAACTGGTAGAAGGTGAGGCCAATGTTGTTGGTTGTGTT
CGCTACGTTAAAGGTATAATGCACCCAATAGAAGAGGAATTTTGGGCCAAGGAGAAGCTGGTGGCGCTGG
ATAATAAGAACATGAGCTACAGTTATATTTTTACTGAGTGTTTTACAGGGTACGAGGATTACACGGCTAC
CATGCAAATAGTGGAGGGTCCTGAGCACAAGGGAAGTAGATTTGACTGGTCTTTTCAGTGCAAGTATATC
GAGGGTATGACTGAATCTGCATTCACCGAGATTCTGCAGCATTGGGCTACTGAGATAGGTCAGAAAATCG
AAGAGGTTTGCAGTGCTTGA

***Allium cepa* clone 4F10-155 lachrymatory-factor synthase (LFS) gene, complete cds GenBank: JN798504.1**

>JN798504.1:13357-13866 *Allium cepa* clone 4F10-155 lachrymatory-factor synthase (LFS) gene, complete cds

ATGGAGCTAAATCCTGGTGCACCTGCTGTAGTCGCTGATAGTGCTAACGGAGCTCGAAAATGGAGCGGCA
AAGTCCATGCTTTGCTTCCAAATACAAAGCCAGAGCAAGCATGGACACTACTAAAAGACTTTATTAACCT
TCACAAGGACATGCCTTCGTTGTCAGTCTGTGAACTGGTAGAAGGTGAGGCCAATGTTGTTGGTTGTGTT
CGCTACGTTAAAGGTATAATGCACCCAATAGAAGAGGAATTTTGGGCCAAGGAGAAGCTGGTGGCGCTGG
ATAATAAGAACATGAGCTACAGTTATATTTTTACTGAGTGTTTTACAGGGTACGAGGATTACACGGCTAC

CATGCAAATAGTGGAGGGTCCTGAGCACAAGGGAAGTAGATTTGACTGGTCTTTTCAGTGCAAGTATATC
GAGGGTATGACTGAATCTGCATTCACCGAGATTCTGCAGCATTGGGCTACTGAGATAGGTCAGAAAATCG
AAGAGGTTTGCAGTGCTTGA

**_Allium sativum_ mRNA, similar to onion lachrymatory factor synthase sequence**

**GenBank: AB164240.1**

TGCTAAAGAAGCACCAAAATGGACTGGCAAAGTCTACGCATTGCTTCCAAATACAAAGGCAGAGCAAGCAT
GGACACTGCTAAAAGACTTTGTTAACCTTGACAAGGTCATGCCTTCATTGTCAGTTTGTGAATTGGTAGAA
GGTAAGGCCAATGTTGTCGGTTGTGTTCGCTTCGTTAAAGGCTTGATGCACCCAATGGAACAGGAATTTTG
GGCCAAGGAGAAGCTGGTGGCACTGGATGACATGACCATGAGCTACAGCTATATTTTTACTGAGTGTTTTA
CCGGGTACGAGGATTACACGGCTACCATGCAAATAGTCGATGGATGTGCGGACAAAGGATGTAGATTTGAC
TGGTCGTTTCAGTGCAAGAACGTTGAAGGTATGACTGCATCTGCGTTCGCCGATGTTCTGCAGCATTGGGC
TACTGAGATTGGTCAGAAAATTGAAGAGATTTGCAATGCTTGATCACAAGCATCGGTTCTGGTTGTGATGC
ATTGTGTGTTTCAAGTCACCGTCTTGTGATTTAATAATATGCATGTAATCTGTAAGCCTGGATGTTGTGTT
TTCAATAAAAAAGAAATTGCCTTGTGCAATGTTCTAATCCTAAATAAAAAAAA

```
     Alignment: C:\Users\PC-09\Desktop\FT like final.txt


                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                    5         15         25         35         45         55

KC677631.1      ATGGCAAGAG AAAGTGACCC ATTAATTGTA GGTAGAATAG TTGGTGATGT AATATCTCCA
KP711810.1      ---ATGATGG ATTCGGATCC GTTAAGGTTG GGTAGAATAG TAGGTGATGT TATAGACCCG
Clustal Co            *  * *    ** **  ****   *  ********** * ********  ***   **

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                    65         75         85         95        105        115

KC677631.1      TTCACAAGAA GTGTTCCTTT AAGGGTAATA TACCCAGCAA AAGAGGTAGC CAATGGGCGT
KP711810.1      TTTACCAGAA GGGTGTCGCT TAGGGCTGTC TACTCATGCA GAGAAGTTGC AAATGGACGC
Clustal Co      ** ** **** * **  *  *  ****   *  *** **   *  *** ** **  ***** **

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   125        135        145        155        165        175

KC677631.1      GAGTTTAAAC CATCACAAAT AACTCAGCAA CCTAGAGTTG AGATCGGAAG TGATGATCTC
KP711810.1      GAGTTTAGGC CTTCTCAAGT TGCCCTACAG CCAAGAATTG AAATTGGCGG CGGTGACCTT
Clustal Co      *******  * * ** *** *   * *  **  ** *** *** * ** **  *  * *** **

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   185        195        205        215        225        235

KC677631.1      AGAACCTTCT ATACGCTAGT GATGGTGGAC CCAGACGCCC CTAGCCCAAG TAATCCACAC
KP711810.1      AGAAACTCCT ATGCACTTGT AATGGTGGAC CCAGATGCTC CAAGCCCAAG CAATCCCCAT
Clustal Co      **** ** ** ** * ** **  ********* ***** ** * * ********  ***** **

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   245        255        265        275        285        295

KC677631.1      CTTAGGGAAT ACTTGCATTG GATGGTGTCA GACATTCCAG GAGGTACAGG ATCGAACTTT
KP711810.1      CTAAGAGAAT ACTTGCACTG GTTGGTCACA GACATTCCAG GAAGCACAAG TGCAAGCTTC
Clustal Co      ** ** **** ******* ** * ****  ** **********  ** * *** *   * * ***

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   305        315        325        335        345        355

KC677631.1      GGGCGAGAAA CCGTCTGCTA TGAAAGCCCG AGACCAACAG CTGGTATCCA CCGTTTCGTT
KP711810.1      GGTCAAGAAA GGATATGCTA TGAAAGTCCA AGGCCATCCT TAGGAATACA CAGGTTTGCC
Clustal Co      ** * *****    * ***** ****** **  ** *** *     ** ** ** * * ** *

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   365        375        385        395        405        415

KC677631.1      TTTATCCTGT TCCAGCAGCT TGGGCGCCAA ACTGTATATG CACCGAACTG GCGGCAAAAC
KP711810.1      TTCATATTAT TTCAGCAGCT TGGTCGCGAG ACCGTGTGCT CTCCAAACTA TAGGCAGAAC
Clustal Co      ** **  * * * ******** *** *** *  ** ** *    * ** ****    **** ***

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   425        435        445        455        465        475

KC677631.1      TTTAATACTC GAGAATTTGC AGAAAATTAC AACCTTGGGT CTCCAGTTGC TGCAGTTTAC
KP711810.1      TTCAGCTCCA GAGTTTTTGC AGAAGTATAC AACTTGGGTT CTCCTGTTGC TGCTCTTTAT
```

```
Clustal Co   ** *   *   ***  ***** ****   *** *** * ** * **** ***** ***  ****

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|..
                485        495        505        515        525        535
KC677631.1   TTCAACTGCC AGAGAGAATC AGGATCTGGT GGGAGGAGAA TATACACGGA TTACTAA
KP711810.1   TTCAACTGCC AGAGAGAAGC TGGTTCAGGT GGGAGGAGAA CTTATACATG A------
Clustal Co   ********** ******** *  ** ** *** **********   ** **
```

```
     Alignment: C:\Users\PC-09\Desktop\LFS Final.txt


                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                    5         15         25         35         45         55

AB094593.1      ATGGAGCTAA ATCCTGGTGC ACCTGCTGTA GTCGCTGATA GTGCTAACGG AGCTCGAAAA
AB094592.1      ATGGAGCAAA ATTCTGGTAC GCTTGCTGTA GTCACTGATG GTGCTAAAGG AGCTGCAAAA
AB089203.1      ATGGAGCTAA ATCCTGGTGC ACCTGCTGTA GTCGCTGATA GTGCTAACGG AGCTCGAAAA
JN798504.1      ATGGAGCTAA ATCCTGGTGC ACCTGCTGTA GTCGCTGATA GTGCTAACGG AGCTCGAAAA
AB164240.1      ---------- ---------- ---------- ---------- -TGCTAAAGA AGCACCAAAA
Clustal Co                                                    ****** *  ***   ****

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                    65         75         85         95        105        115

AB094593.1      TGGAGCGGCA AAGTCCATGC TTTGCTTCCA AATACAAAGC CAGAGCAAGC ATGGACACTA
AB094592.1      TGGAGAGGCA AAGTCCATGC TTTGCTTCCA AATACAAAAC CAGAGCAAGC ATGGACACTA
AB089203.1      TGGAGCGGCA AAGTCCATGC TTTGCTTCCA AATACAAAGC CAGAGCAAGC ATGGACACTA
JN798504.1      TGGAGCGGCA AAGTCCATGC TTTGCTTCCA AATACAAAGC CAGAGCAAGC ATGGACACTA
AB164240.1      TGGACTGGCA AAGTCTACGC ATTGCTTCCA AATACAAAGG CAGAGCAAGC ATGGACACTG
Clustal Co      ****  **** ***** * **  ********* ********   ********** *********

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   125        135        145        155        165        175

AB094593.1      CTAAAAGACT TTATTAACCT TCACAAGGTC ATGCCTTCGT TGTCAGTCTG TGAACTGGTA
AB094592.1      CTAAAAGACT TTATTAACCT TCACAAGATC ATGCCTTCGT TGTCAGTTTG TGAACTGGTT
AB089203.1      CTAAAAGACT TTATTAACCT TCACAAGGTC ATGCCTTCGT TGTCAGTCTG TGAACTGGTA
JN798504.1      CTAAAAGACT TTATTAACCT TCACAAGGAC ATGCCTTCGT TGTCAGTCTG TGAACTGGTA
AB164240.1      CTAAAAGACT TTGTTAACCT TGACAAGGTC ATGCCTTCAT TGTCAGTTTG TGAATTGGTA
Clustal Co      ********** ** ******* * *****  * ******** * ******* ** **** ****

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   185        195        205        215        225        235

AB094593.1      GAAGGTGAGG CCAATGTTGT TGGTTGTGTT CGCTACGTTA AAGGTATAAT GCACCCAATA
AB094592.1      GAAGGCGAGG CCAATGTTGT TGGTTGTGTT CGCCACGTTA AAGGTATAAT GCACCCAATG
AB089203.1      GAAGGTGAGG CCAATGTTGT TGGTTGTGTT CGCTACGTTA AAGGTATAAT GCACCCAATA
JN798504.1      GAAGGTGAGG CCAATGTTGT TGGTTGTGTT CGCTACGTTA AAGGTATAAT GCACCCAATA
AB164240.1      GAAGGTAAGG CCAATGTTGT CGGTTGTGTT CGCTTCGTTA AAGGCTTGAT GCACCCAATG
Clustal Co      *****  *** **********  ********* ***  ***** ****  * ** *********

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   245        255        265        275        285        295

AB094593.1      GAAGAGGAAT TTTGGGCCAA GGAGAAGCTG GTGGCGCTGG ATAATAAGAA CATGAGCTAC
AB094592.1      GAAGAGGAAT TTTGGGCCAA GGAGAAGCTG GTTGCACTGG ATGATAAGAA CATGAGCTGT
AB089203.1      GAAGAGGAAT TTTGGGCCAA GGAGAAGCTG GTGGCGCTGG ATAATAAGAA CATGAGCTAC
JN798504.1      GAAGAGGAAT TTTGGGCCAA GGAGAAGCTG GTGGCGCTGG ATAATAAGAA CATGAGCTAC
AB164240.1      GAACAGGAAT TTTGGGCCAA GGAGAAGCTG GTGGCACTGG ATGACATGAC CATGAGCTAC
Clustal Co      *** ****** ********** ********** ** ** **** ** * * **  ********

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   305        315        325        335        345        355

AB094593.1      AGTTATATTT TTACTGAGTG TTTTACAGGG TACGAGGATT ACACGGCTAC CATGCAAATA
```

```
AB094592.1   AGTTATATTT TTGTTGAGTG TTTTACAGGG TACGAGGATT ACACAGCTAC CATGCAAATA
AB089203.1   AGTTATATTT TTACTGAGTG TTTTACAGGG TACGAGGATT ACACGGCTAC CATGCAAATA
JN798504.1   AGTTATATTT TTACTGAGTG TTTTACAGGG TACGAGGATT ACACGGCTAC CATGCAAATA
AB164240.1   AGCTATATTT TTACTGAGTG TTTTACCGGG TACGAGGATT ACACGGCTAC CATGCAAATA
Clustal Co   ** ******* **  ****** ****** *** ********** **** ***** **********

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                365        375        385        395        405        415

AB094593.1   GTGGAGGGTC CTGAGCACAA GGGAAGTAGA TTTGACTGGT CTTTTCAGTG CAAGTATATC
AB094592.1   GTGGAGGGAT CTGAGCACAA GGGATGTAGA TTTGACTGGT CTTTTCAGTG TAAGTATATC
AB089203.1   GTGGAGGGTC CTGAGCACAA GGGAAGTAGA TTTGACTGGT CTTTTCAGTG CAAGTATATC
JN798504.1   GTGGAGGGTC CTGAGCACAA GGGAAGTAGA TTTGACTGGT CTTTTCAGTG CAAGTATATC
AB164240.1   GTCGATGGAT GTGCGGACAA AGGATGTAGA TTTGACTGGT CGTTTCAGTG CAAGAACGTT
Clustal Co   ** ** **    ** * ****  *** ***** ********** * ********  *** *  *

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                425        435        445        455        465        475

AB094593.1   GAGGGTATGA CTGAATCTGC ATTCACCGAG ATTCTGCAGC ATTGGGCTAC TGAGATAGGT
AB094592.1   GAGGGTATGA CTGAATCTGC ATTCACCGAT GTTCTGCAGC ATTGGGCTAC TGAGATTGGT
AB089203.1   GAGGGTATGA CTGAATCTGC ATTCACCGAG ATTCTGCAGC ATTGGGCTAC TGAGATAGGT
JN798504.1   GAGGGTATGA CTGAATCTGC ATTCACCGAG ATTCTGCAGC ATTGGGCTAC TGAGATAGGT
AB164240.1   GAAGGTATGA CTGCATCTGC GTTCGCCGAT GTTCTGCAGC ATTGGGCTAC TGAGATTGGT
Clustal Co   ** ******* *** ******  *** ****   ********* ********** ****** ***

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                485        495        505        515        525        535

AB094593.1   CAGAAAATCG AAGAGGTTTG CAGTGCTTGA ---------- ---------- ----------
AB094592.1   CAGAAAATTG AAGAGATTTG CAATGCTTGA ---------- ---------- ----------
AB089203.1   CAGAAAATCG AAGAGGTTTG CAGTGCTTGA ---------- ---------- ----------
JN798504.1   CAGAAAATCG AAGAGGTTTG CAGTGCTTGA ---------- ---------- ----------
AB164240.1   CAGAAAATTG AAGAGATTTG CAATGCTTGA TCACAAGCAT CGGTTCTGGT TGTGATGCAT
Clustal Co   ******** * ***** **** ** *******

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                545        555        565        575        585        595

AB094593.1   ---------- ---------- ---------- ---------- ---------- ----------
AB094592.1   ---------- ---------- ---------- ---------- ---------- ----------
AB089203.1   ---------- ---------- ---------- ---------- ---------- ----------
JN798504.1   ---------- ---------- ---------- ---------- ---------- ----------
AB164240.1   TGTGTGTTTC AAGTCACCGT CTTGTGATTT AATAATATGC ATGTAATCTG TAAGCCTGGA
Clustal Co

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                605        615        625        635        645        655

AB094593.1   ---------- ---------- ---------- ---------- ---------- ----------
AB094592.1   ---------- ---------- ---------- ---------- ---------- ----------
AB089203.1   ---------- ---------- ---------- ---------- ---------- ----------
JN798504.1   ---------- ---------- ---------- ---------- ---------- ----------
AB164240.1   TGTTGTGTTT TCAATAAAAA AGAAATTGCC TTGTGCAATG TTCTAATCCT AAATAAAAAA
Clustal Co

             ..

AB094593.1   --
```

```
AB094592.1   --
AB089203.1   --
JN798504.1   --
AB164240.1   AA
Clustal Co
```

```
     Alignment: C:\Users\PC-09\Desktop\ANS FINAL.txt


                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                    5         15         25         35         45         55

EF192598.1      ACTCCATTCC TTTGGAATAC ATCCGGCCAG AGCACGAGCG GGCCTGCCTA GGCGATGCAT
EF192475.2      ACTCCATCCC TCTGGAATAC ATCCGGCCCG AGCACGAGCG GGCCTGTCTA GGCGATGCAT
AY585678.1      ACTCCATTCC TTTGGAATAC ATCCGGCCAG AGCACGAGCG GGCCTGCCTA GGCGATGCAT
Clustal Co      ******* ** * ******** ******** * ********** ****** *** **********

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   65         75         85         95        105        115

EF192598.1      TAGAACAGCT TCACAACTCC AACTCCGGCC CACAAATTCC GATAACCGAC CTCGACTCAC
EF192475.2      TAGAACAGCT TCACAACTCC AACTCCGGCC CACAAATTCC GATAACCGAC CTCGATTCAA
AY585678.1      TAGAACAGCT TCACAACTCC AACTCCGGCC CACAAATTCC GATAACCGAC CTCGACTCAC
Clustal Co      ********** ********** ********** ********** ********** ***** ***

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   125        135        145        155        165        175

EF192598.1      GCGACTGCAT TGAGAAAGTG ACCAAAGCAG CCAAAGAATG GGGAGTCATG CACATTGTGA
EF192475.2      GCGACTGCAT TGAGAAAGTG ACCAAAGCAG CCAAAGAATG GGGAGTCATG CACATTGTGA
AY585678.1      GCGACTGCAT TGAGAAAGTG ACCAAAGCAG CCAAAGAATG GGGAGTCATG CACATTGTGA
Clustal Co      ********** ********** ********** ********** ********** **********

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   185        195        205        215        225        235

EF192598.1      ACCATGGAAT CTCTAGCGAG TTGATGGAGA AAGTGCGTGC AGCTGGCAAG GCTTTTTTCA
EF192475.2      ACCATGGAAT CTCTAGCGAG TTGATGGAGA AAGTGCGTGC AGCTGGCAAG GCTTTTTTCA
AY585678.1      ACCATGGAAT CTCTAGCGAG TTGATGGAGA AAGTGCGTGC AGCTGGCAAG GCTTTTTTCA
Clustal Co      ********** ********** ********** ********** ********** **********

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   245        255        265        275        285        295

EF192598.1      ATCTCCCACT CGAGGCTAAA GAGGAATATG CTAACGATCA ATCTAAAGGG AAGATCCAAG
EF192475.2      ATCTCCCACT CGAGGCTAAA GAGGAATATG CTAACGATCA ATCTAAAGGG AAGATCCAAG
AY585678.1      ATCTCCCACT CGAGGCTAAA GAGGAATATG CTAACGATCA ATCTAAAGGG AAGATCCAAG
Clustal Co      ********** ********** ********** ********** ********** **********

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   305        315        325        335        345        355

EF192598.1      GGTATGGAAG TAAACTGGCA AATAATGCAA GCGGGCAACT CGAGTGGGAG GACTACTTTT
EF192475.2      GGTATGGAAG TAAACTGGCA AATAATGCAA GCGGGCAACT CGAGTGGGAG GACTACTTTT
AY585678.1      GGTATGGAAG TAAACTGGCA AATAATGCAA GCGGGCAACT CGAGTGGGAG GACTACTTTT
Clustal Co      ********** ********** ********** ********** ********** **********

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   365        375        385        395        405        415

EF192598.1      TTCATCTCAT TTTTCCTGAT GACAAGGTCG ATCTCTCCGT TTGGCCTAAA CAGCCTTCCG
EF192475.2      TTCATCTCAT TTTTCCTGAT GACAAGGTCG ATCTCTCCGT TTGGCCTAAA CAGCCTTCCG
AY585678.1      TTCATCTCAT TTTTCCTGAT GACAAGGTCG ATCTCTCCGT TTGGCCTAAA CAGCCTTCCG
```

```
Clustal Co  ********** ********** ********** ********** ********** **********

            ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
               425        435        445        455        465        475

EF192598.1  ACTACATTGA AATTATGCAA GAGTTTGGAA GTCAGCTGAG AATATTAGCC AGCAAAATGC
EF192475.2  ACTACATTGA AATTATGCAA GAGTTTGGAA GTCAGCTGAG AATATTAGCC AGCAAAATGC
AY585678.1  ACTACATTGA AATTATGCAA GAGTTTGGAA GTCAGCTGAG AATATTAGCC AGCAAAATGC
Clustal Co  ********** ********** ********** ********** ********** **********

            ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
               485        495        505        515        525        535

EF192598.1  TATCCATACT TTCATTGGGC TTACAACTAC CAACCAAGGA CAGGCTAGAA CAAGAACTAA
EF192475.2  TATCCATACT TTCATTGGGC TTACAACTAC CAACCAAGGA CAGGCTAGAA CAAGAACTAA
AY585678.1  TATCCATACT TTCATTGGGC TTACAACTAC CAACCAAGGA CAGGCTAGAA CAAGAACTAA
Clustal Co  ********** ********** ********** ********** ********** **********

            ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
               545        555        565        575        585        595

EF192598.1  AAGGACCAGA AGACTTACTT CTCCAGCTGA AAATAAACTA CTACCCAAAA TGCCCCCAGC
EF192475.2  AAGGACCAGA AGACTTACTT CTCCAGCTGA AAATAAACTA CTACCCAAAA TGCCCCCAGC
AY585678.1  AAGGACCAGA AGACTTACTT CTCCAGCTGA AAATAAACTA CTACCCAAAA TGCCCCCAGC
Clustal Co  ********** ********** ********** ********** ********** **********

            ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
               605        615        625        635        645        655

EF192598.1  CACATCTGGC ATTAGGAGTG GAAGCCCACA CGGACGTGAG TGCCTTATCG TTCATACTGC
EF192475.2  CACATCTGGC ATTAGGAGTG GAAGCCCACA CGGACGTGAG TGCCTTATCG TTCATACTGC
AY585678.1  CACATCTGGC ATTAGGAGTG GAAGCCCACA CGGACGTGAG TGCCTTATCG TTCATACTGC
Clustal Co  ********** ********** ********** ********** ********** **********

            ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
               665        675        685        695        705        715

EF192598.1  ACAACAACGT TCCAGGCCTT CAAGTGTTGT ACGAGGGTGA ATGGGTAACT GCGAAACTCG
EF192475.2  ACAACAACGT TCCAGGCCTT CAAGTGTTGT ACGAGGGTGA ATGGGTAACT GCGAAACTCG
AY585678.1  ACAACAACGT TCCAGGCCTT CAAGTGTTGT ACGAGGGTGA ATGGGTAACT GCGAAACTCG
Clustal Co  ********** ********** ********** ********** ********** **********

            ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
               725        735        745        755        765        775

EF192598.1  TTCCAGATTC CTTGATTGTC CATGTTGGGG ATTCGCTTGA AATTTTGAGT AATGGGATTT
EF192475.2  TTCCAGATTC CTTGATTGTC CATGTTGGGG ATTCGCTTGA AATTTTGAGT AATGGGATTT
AY585678.1  TTCCAGATTC CTTGATTGTC CATGTTGGGG ATTCGCTTGA AATTTTGAGT AATGGGATTT
Clustal Co  ********** ********** ********** ********** ********** **********

            ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
               785        795        805        815        825        835

EF192598.1  ACAAAAGTGT GTTGCATCGC GGTTTGGTTA ATAAGGAGAA GGTGAGGATT TCTTGGGCTG
EF192475.2  ACAAAAGTGT GTTGCATCGC GGTTTGGTTA ATAAGGAGAA GGTGAGGATT TCTTGGGCTG
AY585678.1  ACAAAAGTGT GTTGCATCGC GGTTTGGTTA ATAAGGAGAA GGTGAGGATT TCTTGGGCTG
Clustal Co  ********** ********** ********** ********** ********** **********

            ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
```

```
                845        855        865        875        885        895

EF192598.1   TGTTCTGTGA ACCGCCAAAG GATGCTGTTG TGCTTAAGCC ATTGGATGAG GTTGTGACAG
EF192475.2   TGTTCTGTGA ACCGCCAAAG GATGCTGTTG TGCTTAAGCC ATTGGATGAG GTTGTGACAG
AY585678.1   TGTTCTGTGA ACCGCCAAAG GATGCTGTTG TGCTTAAGCC ATTGGATGAG GTTGTGACAG
Clustal Co   ********** ********** ********** ********** ********** **********

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                905        915        925        935        945        955

EF192598.1   ACGATGCGCC CGCAAGGTAC ACTCCTCGCA CGTTTGCACA GCATCTGGAA CGGAAGTTGT
EF192475.2   ACGATGCGCC CGCAAGGTAC ACTCCTCGCA CGTTTGCACA GCATCTGGAA CGGAAGTTGT
AY585678.1   ACGATGCGCC CGCAAGGTAC ACTCCTCGCA CGTTTGCACA GCATCTGGAA CGGAAGTTGT
Clustal Co   ********** ********** ********** ********** ********** **********

             ....|....| ....|....| ....|....| ....|....| ..
                965        975        985        995
EF192598.1   TTAAAAAGAA AG-TTGGAGA TTTGGATGAC TCTGATGTTT GA
EF192475.2   TTAAAAAGAA AG-TTGGAGA TTTGGATGAC TCTGATGTTT GA
AY585678.1   TTAAAAAGAA AGGTTGGAGA TTTGGATGA- ---------- --
Clustal Co   ********** ** ******* *********
```

# Thesis MPhil Attique

ORIGINALITY REPORT

| 15% | 7% | 9% | 8% |
|---|---|---|---|
| SIMILARITY INDEX | INTERNET SOURCES | PUBLICATIONS | STUDENT PAPERS |

PRIMARY SOURCES

| | | |
|---|---|---|
| 1 | Submitted to Higher Education Commission Pakistan<br>Student Paper | 3% |
| 2 | "The Allium Genomes", Springer Science and Business Media LLC, 2018<br>Publication | 3% |
| 3 | juniperpublishers.com<br>Internet Source | 2% |
| 4 | Submitted to University of Leicester<br>Student Paper | 1% |
| 5 | link.springer.com<br>Internet Source | 1% |
| 6 | www.nestlenutrition-institute.org<br>Internet Source | <1% |
| 7 | krishikosh.egranth.ac.in<br>Internet Source | <1% |
| 8 | Predrag Putnik, Domagoj Gabrić, Shahin Roohinejad, Francisco J. Barba et al. "An overview of organosulfur compounds from | <1% |