Molecular Biology and Biochemical Study of Lachrymatory Factor Synthase in *Allium cepa* L. (onion)

A Dissertation Submitted to

the Graduate School of Life and Environmental Sciences,

the University of Tsukuba

in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy in Agricultural Science

Noriya MASAMURA

TABLE OF CONTENTS

Chapter I. GENERAL INTRODUCTION

According to recent estimations, approximately 750 species are included in the genus *Allium* (STEARN 1992). This genus is widely distributed over the Holarctic region from the dry subtropics to the boreal zone (FRITSCH and FRIESEN 2002). *Allium cepa* L. (2n = 2x = 16), which belongs to section *Cepa* in the genus, is one of the most important species of *Allium*. The species was subdivided into two groups: Common onion and Aggregatum (FRITSCH and FRIESEN 2002; HANELT 1990). The bulb onion of the common onion group is one of the most cultivated vegetables in the world, and its annual production was estimated to be approximately 70 million tons in 2007, which ranked second in value after the tomato on a list of cultivated vegetable crops worldwide ("FAOSTAT" 2007). The shallot (genomes AA) of the Aggregatum group is an important genetic resource for the improvement of the common onion because of their close relationship (HAVEY 2002; TASHIRO *et al.* 1982).

The consumption of onions can be traced back to the time of ancient Egypt, as Herodotus reported that laborers in the pyramid era of the Egyptian dynasties were fed onions and garlic (*A. sativum*), according to an inscription on the pyramid wall. From that era to present times, most people have been familiar with the lachrymatory smell arising when bulbs of onion are cut. This LF was identified as propanethial *S*-oxide almost 40 years ago by Brodnitz and Pascale (BRODNITZ and PASCALE 1971). Like other sulfur-containing flavor constituents of onion, such as thiosulfinates, thiols, and disulfides, LF has long been believed to form spontaneously following the reaction of (*E*)-(+)-*S*-(1-propenyl)-L-cysteine sulfoxide ((*E*)-PRENCSO) and alliinase. This scheme suggests that the formation of LF is inevitable without forfeiting the characteristic onion flavor. However, Imai *et al.* (2002) reported that, in contrast to the commonly accepted

scheme of the non-enzymatic reaction, the formation of LF required a novel enzyme, which they named LFS.

The discovery of LFS has proven that the formation of the LF is under the control of a specific single gene, which suggests the possibility of producing an onion that lacks the tearing property (tearless onion) via a genetic modification (GM) approach, mutagenesis, and traditional breeding. This tearless property would also lead to low pungency because LF induces mouth burn and pungency (RANDLE 1997). The market for low-pungency onions is expanding in the United States. The consumption of low-pungency onions now accounts for 15%–25% of total onion consumption in the United States. However, low-pungency onions that are currently on the market have poor storage characteristics. Therefore, there is an obvious need for a long-storage low-pungency onion.

Onions have been used in folk medicine as well as for food. Modern medical science has confirmed the health benefits of onions in both laboratory studies and large-scale epidemiological studies. The findings show that *Allium* vegetable consumption benefits the cardiovascular system and gut health as well as enhances cancer avoidance (GRIFFITHS *et al.* 2002).

Considering onions from the perspective of health benefits, three groups of chemical compounds fructo-oligosaccharides (fructan), flavonoids, mainly quercetin and its derivatives, and flavor precursors alk(en)yl cysteine sulfoxides (ACSOs) have received a lot of attention in recent years (GRIFFITHS *et al.* 2002). Fructans confer benefits as a prebiotic. A detailed analysis of their synthetic pathway and related genes in *Allium* was reported by McCallum *et al.* (2006) and Yaguchi *et al.* (2008). The variation in quercetin content among cultivars was investigated by Okamoto *et al.* (2006) and Lombard *et al.* (2004), and Mogren *et al.* (2006, 2007) reported that environmental factors, such as

nitrogen fertility, lifting time, and global radiation, also seem to affect quercetin content. Attempts to develop onions with increased quercetin content have been progressing, according to Okamoto *et al.* (2006). Several onion varieties with high quercetin content have already been marketed in Japan.

Another important compound group ACSOs, which are abundant in the genus *Allium* (ROSE *et al.* 2005), are precursors of the characteristic flavor reported above and of substances valuable to human health. Studies focusing on synthetic pathways and accumulation in onion for controlling ACSO content have been reported to date by McCallum *et al.* (2005) and Randle *et al.* (1995). Masamura *et al.* (2011) assigned six candidate genes (adenosine 5'-phosphosulfate reductase (APSR), *O*-acetylserine thiol-lyase (OASTL), serine acetyltransferase (SAT), glutamyl cysteine synthase (GCS), glutathione synthase (GS), and γ-glutamyl transpeptidase (GGT)) for the biosynthesis of ACSOs to the chromosomes of *A. cepa*, and revealed the accumulation level of ACSOs in *A. fistulosum* – shallot monosomic addition lines throughout the year. The ACSO concentration was correlated with the anti-platelet aggregation activity of the onion extract (GALMARINI *et al.* 2001).

The formation of organosulfur compounds that impart flavor and human health benefits begins with the enzymatic decomposition of ACSOs following tissue disruption; in onion, LFS and alliinase catalyze the first two reactions in the pathway. Because of the difficulty in controlling the complicated reactions following enzymatic decomposition, targeting the expression levels of the two enzymes and/or ACSO content would be a more effective and direct approach to regulate the content of the health-beneficial organosulfur compounds. Therefore, suppressing the LFS gene would also be effective for the production of an onion with health benefits arising from sulfur-containing compounds.

Namely, suppressing LFS turns off the conversion of 1-propenylsulfenic acid to LF (leading to tearlessness and low pungency); this, in turn, increases the yield of thiosulfinates, to which the flavor and health benefits of the onion are attributed. Eady *et al.* (2008) demonstrated large shifts in organosulfur secondary compound profiles in onions in which LFS activity was suppressed by RNAi. In addition, Aoyagi *et al.* (2011) reported that the substance specifically produced in an LFS-suppressed onion showed higher inhibitory activity of cyclooxygenase-I (COX-1) than aspirin. Therefore, LFS is an important target for molecular breeding of onions.

Breeders have focused primarily on improving pest and disease resistance of the bulb onion. In addition, genetic analysis of cytoplasmic male-sterility has been addressed as a matter of priority (McCallum 2007). Efforts in these studies have generated direct benefits for onion producers. The breeding of LFS-suppressed onions will offer advantages to the consumer such as tearlessness, low pungency, and health benefits.

Because onion is a biennial, out-crossing, and highly heterozygous species, genetic and genomic resources of onion have previously been limited. In addition to this specificity, the very large genome size of the *Allium* species (Ricroch *et al.* 2005) complicates genomic cloning and sequencing strategies. However, recently, an EST resource and PCR-based maps of onion (Kuhl *et al.* 2004; Martin *et al.* 2005) were developed. Complete sets of *A. fistulosum* – shallot monosomic addition lines ($2n = 2x+1 = 17$, FF+1A – FF+8A; (Shigyo *et al.* 1996), namely, *A. fistulosum* carrying each of the eight different chromosomes of shallot. In combination with the EST resource, the PCR-based maps and the chromosome addition lines have revealed the chromosomal location and genetic map position of genes responsible for important properties of the onion.

We studied LFS by using molecular biological, biochemical, and analytical approaches to contribute to the mutagenesis approaches and the traditional breeding of LFS-suppressed (tearless, low-pungency, health enhancing) onions. The objectives of the present study are as follows: (1) to determine the genomic organization of the LFS gene, (2) to identify catalytic amino acid residues and develop a three-dimensional model structure of LFS with the aim of elucidating the unique catalytic reaction mechanism of this enzyme, (3) to understand the reaction mechanism of LFS.

This dissertation is composed of five chapters, including this one. Chapter II deals with determination of the genomic organization of the LFS gene. Chapter III gives information on catalytic amino acid residues and a three-dimensional model structure of LFS. Chapter IV is a discussion of the reaction mechanism of LFS, and Chapter V contains a general discussion. The results of studies conducted for the purposes of this dissertation, as described above, are compiled here. The studies were conducted at the Somatech Center, House Foods Corporation, Japan and the Laboratory of Vegetable and Crop Science, Division of Agrobiology, Department of Biological and Environmental Science, Faculty of Agriculture, Yamaguchi University, Japan from 2006 to 2012 (MASAMURA *et al.* 2012 a; b; c).

Chapter II. CHROMOSOMAL ORGANIZATION AND SEQUENCE DIVERSITY OF GENES ENCODING LACHRYMATORY FACTOR SYNTHASE IN *ALLIUM CEPA* L.

INTRODUCTION

The onion nuclear genomes is notable for its great size, 17 pg or 15 Gbp per 1C (ARUMUGANATHAN and EARLE 1991; BENNETT and LEITCH 1995; RICROCH *et al.* 2005), one of the largest among cultivated plants. There are limited genomic resources available in onion caused by the huge genome size and genetic studies are further complicated because it is a biennial, out-crossing, and highly heterozygous species. However, an EST resource and PCR based map (KUHL *et al.* 2004; MARTIN *et al.* 2005) has been recently developed. Combination with these resources and chromosome addition lines (SHIGYO *et al.* 1996) has revealed chromosomal location and genetic map position of genes responsible for important properties such as carbohydrate accumulation (MCCALLUM *et al.* 2006; MASUZAKI *et al.* 2006a; YAGUCHI 2008), flavonoid biosynthesis (MASUZAKI *et al.* 2006b), and ACSOs biosynthesis (MASAMURA *et al.* 2011). The physical distribution of AFLP markers along *Allium* chromosomes has been studied via the integration of recombination and physical maps in a trihybrid population, *A. cepa* x (*A. roylei* x *A. fistulosum*) (KHRUSTALEVA *et al.* 2005). Direct physical mapping of genes on onion chromosomes is limited due to the genome abundance with repetitive elements (STACK and COMINGS 1979; PEARCE *et al.* 1996). Fluorescence *in situ* hybridization (FISH) was successfully applied for the detection of specific loci using large genomic clones as probes mostly in plant species with small gene-rich genomes, such as *Arabidopsis thaliana* (KOORNNEEF *et al.* 2003) or rice (JIANG *et al.* 1995). However, in some cases

using repetitive DNA that blocks probe hybridization of repetitive sequences allows for the detection of genes inserted in a Bacterial Artificial Chromosome (BAC) (LAMB *et al.* 2007; SZINAY *et al.* 2008).

The most distinctive attribute of onion is the tearing property conferred by lachrymatory factor (LF; propanethial *S*-oxide). LF is formed from 1-propenylsulphenic acid by lachrymatory factor synthase (LFS) (IMAI *et al.* 2002). 1-propenyl sulphenic acid is a putative reaction product derived from 1-propenyl cysteine sulfoxides ((*E*)-PRENCSO) by alliinase. Suppressing LFS gene turns off the conversion of 1-propenylsulfenic acid to LF, which in turn increases the yield of thiosulfinates thought to be responsible for onion flavor, and health beneficial properties of onion. Eady *et al.* (2008) previously demonstrated large shifts in organosulfur secondary compound profiles in onions in which LFS activity was suppressed by RNAi. These studies suggest that LFS is an important target for molecular breeding in onion. LFS cDNAs have been cloned from other five lachrymatory *Allium* species (*A. ampeloprasum*, *A. cepa* Aggregatum group, *A. chinese*, *A. fistulosum*, and *A. porrum*) and all recombinant proteins from them showed LFS enzymatic activity. In addition, strong homologies were observed in these LFS cDNA and no homologous sequences were yielded in GenBank search (IMAI *et al.* 2005; MASAMURA *et al.* 2012c). These results suggested that the LFS gene is strongly conserved among lachrymatory *Allium* species and is only distantly related to proteins in other higher plant taxa. Because LF is such a bioactive and distinctive compound, it is likely that strong selective forces have acted on LFS genes during evolution and domestication of *Allium*.

In this study, as a model of functional gene in huge genome, we determined the genome organization of LFS genes by sequence analysis, genetic mapping and physical

methods in order to understand evolution of LFS in *Allium* and contribute to targeting molecular breeding and mutagenesis approaches for manipulating onion quality.

MATERIALS AND METHODS

Genetic analyses of LFS gene by using monosomic addition line and mapping population

The plant materials were a complete set of *A. fistulosum*-shallot monosomic addition lines [$2n = 2x + 1 = 17$, FF+1A (plant number 130), FF+2A (141), FF+3A (5), FF+4A (10), FF+5A (26), FF+6A (308), FF+7A (324), FF+8A (240)] and parental control plants, Japanese bunching onion (*A. fistulosum* cv. Kujyo-hoso, $2n = 2x = 16$, FF) and shallot (*A. cepa* Aggregatum group 'Chiang Mai', $2n = 2x = 16$, AA) (SHIGYO *et al.* 1996). They were grown in an experimental field at Yamaguchi University (34°N, 131°E).

Genomic DNA was extracted from the frozen base of leaf sheath tissues by DNeasy Plant Mini Kit (QIAGEN, Hilden, Germany). PCR were performed an 25 μl reaction mixture containing 2.5 μl of template, 1:10 or 1:100 diluted cDNA or genomic DNA (20 ng/μl), 0.125 μl of *Taq* polymerase [Ampli*Taq* GOLD (5 U/μl), Applied Biosystems, Foster City, CA, USA], 2.5 μl 10x PCR buffer, 1.5 μl of $MgCl_2$ (25 mM), 0.5 μl of forward primer (25 μM) and 0.5 μl of reverse primer (25 μM) of primer set (cepaLFS) and 2.5 μl of dNTP mixture (2 mM). Nucleotide sequences of the primers were shown in Table 1. PCR was carried out in GeneAmp 2400 or GeneAmp 9600 (Applied Biosystems, Foster City, CA, USA) with the following amplification program, an initial heating to activate the *Taq* polymerase for 94°C for 10 min, followed by 35 cycles of 94°C for 1 min, 65°C or 68°C for 1 min and 72°C for 1 min, and then a final elongation at 72°C for 10 min. The PCR products were detected by electrophoresis in 2% agarose gels.

DNA templates and genetic map data from the interspecific *Allium* cross *A. cepa* x *A. roylei* were used as described by van Heusden *et al.* (2000a; b). Primer sets used in this study are shown in Table 1. Design, PCR and analysis methods for SSCP and SSR markers were described previously (McCallum *et al.* 2006, 2007, 2008). Linkage analysis was performed using Joinmap Ver. 4 (van Ooijen).

Screening of BAC clones containing LFS gene and its molecular characterization

The partial BAC library of onion (Suzuki *et al.* 2001) was used for the PCR screening with onion LFS specific primers amplifying a 459 bp fragment comprising most of the LFS ORF. Nucleotide sequence of the primer set (LFSorf) used were shown in Table 1. PCR was performed with Ampli-*Taq* GOLD (Applied Biosystems, Foster City, CA, USA) for 40 cycles of denaturation for 1 min at 94°C annealing for 1 min at 58°C and extension for 1 min at 72°C followed by a final extension for 7 min at 72°C.

BAC DNA purified by QIAGEN plasmid Midi kit (Qiagen, Hilden, Germany) was digested with *Eco*RI or *Hin*dIII and separated on 0.7% agarose gel. After electrophoresis, DNA was transferred to HybondN$^+$ membrane (GE Healthcare UK Ltd., Buckinghamshire, UK). The digoxigenin (DIG)-labeled LFS specific DNA probe (452 bps) was prepared by PCR reaction using PCR DIG Labeling Mix (Roche Diagnostics, Manheim, Germany). Hybridization was carried out in 5x SSC, 50% formamide at 38.5°C for overnight. The membrane was washed in 0.5x SSC, 0.1% SDS at 65°C for 30 min. After blotting and washing, we detected the positive signal with fluorescence according to the instruction of the DIG fluorescent detection Kit (Roche Diagnostics, Manheim, Germany) with ECF (GE Healthcare UK Ltd., Buckinghamshire, UK) as the substrate.

Purified BAC clone 2E8/10 was nebulised and shotgun cloned into vector pSmartHC and 752 clones were sequenced from both ends (mean sequence length *ca.* 700 bp) by GATC Biotech Ltd Germany (Konstanz, Germany http://www.gatc-biotech.com). Trace files were assembled using the Staden package ver. 1.6 (STADEN 1996) and Sequencher® (Gene Codes Corporation, Ann Arbor, MI, USA). The shotgun sequencing of BAC clone 4F10/155 was performed at Hokkaido System Science Co. Ltd (Sapporo, Japan) with the GS FLX Titanium system (454 Life Sciences, Roche Branford, CT, USA). The GS FLX Titanium generated ca.8.5 million bases (about 30 thousand reads) giving 85-fold coverage of the insert DNA of the clone. The reads were assembled into contigs with GS De Novo Assembler software (454 Life Sciences, Roche Branford, CT, USA). The contigs (~20 kb) containing LFS from both BAC clones were annotated using BLAST (ALTSCHUL *et al.* 1997) searches against Genbank plant protein and nucleotide databases. BLASTN comparison between contigs was visualized using genoPlotR (GUY *et al.* 2010). Tandem repeats were identified using EMBOSS quicktandem (RICE *et al.* 2000). Sequences were submitted to Genbank (Accessions JN798503 and JN798504).

BAC FISH analysis

Young root meristems of bulb onion (*Allium cepa* cv. Khalcedon) were collected in water saturated α-bromnaphthalene (1:1000). The root meristems were kept in α-bromnaphthalene overnight at 4°C. After this treatment the roots were fixed in methanol/acetic acid fixative (3:1) for at least 1h at room temperature. The roots were rinsed three times in distilled water and once in 10 mM sodium citrate buffer (pH 4.5) before transferring to an enzyme mix containing 0.1% pectolyase Y23 (Sigma-Aldrich, St. Louis, Missouri, USA), 0.1% cellulase RS (Yakult Pharmaceutical, Tokyo, Japan) and

0.1% cytohelicase (Bio Sepra, Idstein, Germany) in citrate buffer for 35–50 min at 37°C. The roots were carefully transferred to water and then left on ice until further use. Somatic metaphase chromosome spreads were made according to (PIJNACKER and FERWERDA 1984) with slight modification. In brief, dissected meristemic tissue were placed in very clean grease-free slides and cell suspension were made with fine needles, then 60% acetic acid were dropped onto the cells for further maceration on a hot block at 42°C with stirring for 2 min. 100 µl ethanol/acetic acid fixative (3:1) were added around droplet of acetic acid. Finally slides were rinsed briefly in 96% ethanol and were air dried.

Onion Cot-100 DNA was prepared as described by Zwick *et al.* (1997) with some modification. Total genomic DNA was isolated using CTAB method (ROGERS and BENDICH 1988) and was sonicated to a fragment size of about 1 kb. Sheared DNA was denatured in 0.3 M NaCl at 95°C for 10 min, and then let it re-anneal at 65°C for 31 h 40 min. The exact renaturation time was calculated taking in account the initial DNA concentration and percent of G-C fraction in onion genome. The remaining ssDNA was digested with S1 endonuclease (Fermentas, Burlington, ON, Canada, http://www.fermentas.com) with final concentration 1 U /µg for 90 min at 37°C. The reaction was stopped and DNA was extracted by adding 300 µl chloroform/iso-amylalcohol (24:1).

BAC DNA was isolated using GeneJETTM Plasmid Miniprep Kit (Fermentas, Burlington, ON, Canada, http://www.fermentas.com). We labeled the BAC DNA with digoxigenin-11-dUTP by Nick-translation standard protocol (Roche Diagnostics, Manheim, Germany). Slide pre-treatment was performed according to common FISH procedure. Slides were dried overnight at 37°C, treated with RNAse (0.1 mg/ml) in 2x

SSC for 1 h and then with pepsin (5 μg/ml) in 0.01 M HCl, after that treatment slides were incubate in 4% paraformaldehyde for 10 min. Slides were washed in 2x SSC three times for 5 min. between each steps of treatment. Finally, slides were dehydrated in 70%, 90% and 100% ethanol for 3 min each and air dry. The hybridization mixture consisted of 50% (v/v) deionized formamide, 10% (w/v) dextran sulphate, 2x SSC, 0.25% sodium dodecyl sulphate, 12.5 ng/μl probe DNA and 0.3 μg/μl Cot-100. The mixture was boiled for 5 min and subsequently placed on ice for 5 min. 40 ml of the mixture was administered to the chromosome preparations, covered with a coverslip (22 x 22 mm), and denaturated for 5 min at 80°C. An 81% stringency washing was applied. The slides were washed in 2x SSC for 5 min at 42°C, in 50% (v/v) formamide in 2x SSC twice for 7 min at 42°C, in 2x SSC for 3 min at 42°C. Stable sites of the probe hybridization were detected with anti-Digoxigenin FITC made in sheep (Roche Diagnostics, Manheim, Germany) and amplified with anti-sheep FITC, made in rabbit and anti-rabbit FITC, made in goat (Vector Laboratories, Burlingame, CA, USA). Chromosomes were counterstained with DAPI in Vectashield antifade (Vector Laboratories, Burlingame, CA, USA). Slides were examined under a Zeiss Axio Imager microscope (Carl Zeiss MicroImaging, Jena, Germany equipped with epifluorescence illumination, and filter sets for DAPI, TexasRed and FITC. Selected images were captured using a digital camera Axio Cam MRm. Image processing and thresholding were performed using AxioVision—a modular image-processing and analysis system for modern microscopy. Final image optimization was performed using Photoshop™ (Adobe Inc., San Jose, CA, USA).

Re-sequencing of LFS Amplicons

DNA was isolated as described previously (McCALLUM *et al.* 2006) from *Allium roylei* and the following onion genotypes selected on the basis of previous genetic diversity (McCALLUM *et al.* 2008) and pungency phenotypes: 'W202A', 'Texas Grano 438', 'BYG15-23', 'Alisa Craig 43', 'Colossal Grano', 'Early Longkeeper P12', 'W429A', 'Houston Grano', 'Tearless F_1', 'Faridpuri'. PCR reactions were performed in a 15 µl volume and contained 0.5 µM of primers LFS5L and R (Table 1) , 200 µM of dNTP's, 1.5 mM of Mg^{2+}, 10–30 ng of template DNA, 0.375 U of ThermoPrime*Taq* polymerase (Thermo Scientific, Massachusetts, USA) and 1x the manufacturers ReddyMix ™ PCR buffer. Reactions were performed in a Gene Amp® PCR System 9700 thermocycler (Applied Biosystems, Foster City, CA, USA). Cycling conditions were 2 min at 95°C followed by a touchdown which included four cycles of 30 sec at 95°C, 30 sec of 62°C to 59°C where this annealing temperature decreased by 1°C each cycle and 30 sec at 72°C. This touchdown was followed by 36 cycles of 30 sec at 95°C, 30 sec at 58°C and 30 sec at 72°C and a final extension of 10 min at 72°C. PCR products were visualized by electrophoresis on a 1% LE agarose plus 1% NuSieve (FMC Bioproducts, Rockland, ME, USA) agarose gel stained with ethidium bromide.

PCR products were then ligated into the pCR® 4 TOPO® Vector using a TOPO® TA Cloning® Kit (Invitrogen, Carlsbad, CA, USA) and transformed into DH5α by heat shock. Twelve colonies for each of the 12 onion lines were selected after growth on LB containing Amplicilin 100 µg/ml. These were streaked on Amplicilin 100 µg/ml and tested by PCR (see above). Eight colonies containing LFs amplicons from each of the twelve lines were used to inoculate a 1 ml LB (Ampicillin 100 µg/ml) overnight culture. Plasmids were then isolated using a Perfectprep® Plasmid Isolation Kit (Eppendorf, Hamburg, Germany). LFS plasmid inserts were then Sanger-sequenced with M13

forward primer using Big Dye® Terminators v3.1 (Applied Biosystems, Foster City, CA, USA). Sequencing reactions were run on an ABI PRISM® 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA). Traces were trimmed and aligned using Geneious software (DRUMMOND *et al.* 2010) and submitted into Genbank (Accession numbers HQ738844–HQ738918). Alignments were cropped to the region corresponding to bases 368–632 of Genbank Acc AB089203 and maximum likelihood trees were constructed using PHYML (GUINDON and GASCUEL 2003) with 100 bootstraps. Nucleotide diversity calculations were performed using DnaSP (LIBRADO and ROZAS 2009).

Identification of LFS Transcripts by 454 Sequencing

Raw 454 flowgram files were used from Genbank BioProject accession PRJNA60277 for shoot transcriptome sequencing of doubled haploid onion line 'CUDH2150' (Cornell University; Genbank BioSample 138247) and 'Nasik Red' (USDA-ARS PI 271311; Genbank BioSample 138248). Reads were aligned to the LFS reference sequence (AB089203) using Roche gsMapper software using minimum overlap identity of 98% in the assembly step and minimum overlap length of 300 for 'DH2150' (Titanium reads) and 150 for 'Nasik Red' (GS-FLX reads).

Amplification of LFS haplotypes in the FF+5A

PCR were performed an 30 µl reaction mixture containing 3.0 µl of genomic DNA (20 ng/µl) purified from two different plants [FF+5A(26) and FF+5A(71)] of FF+5A or shallot parent, 0.15 µl of *Taq* polymerase [Ex*Taq*(5 U/µl), TakaraBio, Ohtsu, Japan], 3.0 µl 10x Ex buffer, 0.15 µl of forward primer (25 µM) and 0.15 µl of reverse primer (25

μM) of primer set (LFShaplo) and 2.4 μl of dNTP mixture (2.5 mM). Nucleotide sequences of the primers were shown in Table 1. PCR was carried out in GeneAmp 2400 or GeneAmp 9600 (Applied Biosystems, Foster City, CA, USA) with the following amplification program, an initial heating to activate the *Taq* polymerase for 94°C for 3 min, followed by 35 cycles of 94°C for 1 min, 64.9°C for 1 min and 72°C for 1 min. The 5 μl of PCR reactions were subjected to restriction endonuclease digestion with *Nsi*I and *Sph*I enzymes and the fragments were analyzed by electrophoresis in 2.0% agarose gels.

Table 1 Previously unpublished primer sets used in this study.

Primer Set	Marker type	Genbank Accession Number	Forward Primer Reverse Primer
cepaLFS	—	AB089203	ACAAAGCCAGAGCAAGCATGGACA CTGCAAACCTCTTCGATTTTCTGACCTATC
LFS5	Heteroduplex	AB089203	GCACTAGAACTTGCAAAAAGCA TGAGATAGGTCAGAAAATCGAAGA
ACP052	SSCP	CF445004, CF445805, CF445805	TTCCCTCCTCACTCCCTACA CGACCACAAACACAAGCAAC
ACM295	SSR	CF445600	AGATCCGTCCCATGAAACT GATCCGCTTCTGAAATCTCG
ACM021	SSR	CF448154	AAAACCCTCAACATCTCACTCC TCTCTTCTTCCTCGTCCTGC
ACM076	SSR	CF449018	ATTAGAAACATCCATCGCCG CGCGATCATCATTTTCCATA
ACP003	SSCP	BE205590	AAGCTCTTAAAGCTGCTGATGG ATGCACGATAGCACAAGACATC
LFSorf	—	AB089203	ACAAAGCCAGAGCAAGCATGGACA CTGCAAACCTCTTCGATTTTCTGACCTATC
LFShaplo	—	AB089203	ATAGTGGAGGGTCCTGAGCA ACACAACACTCAGTCTTACTTATT

RESULTS

Chromosomal mapping of LFS gene in A. cepa

From the genomic DNA of alien monosomic addition lines, the expected size amplicon was observed only in the alien monosomic addition line FF+5A (Figure 1).

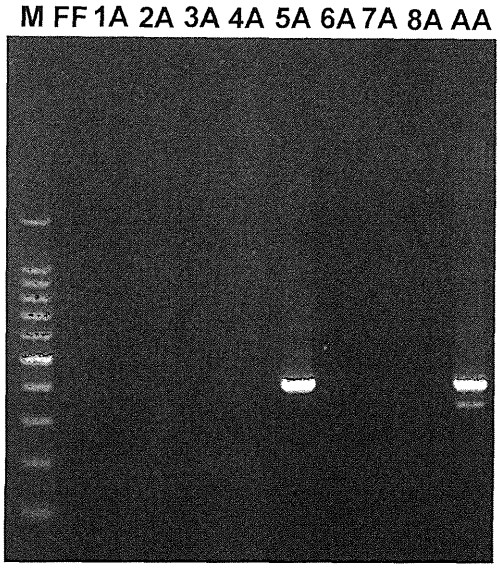


Figure 1 Amplification of LFS from AMALs.
M, molecular size marker (100bp ladder); FF, *A. fistulosum*; 1A-8A, eight different *A. fistulosum* – shallot monosomic additions; AA, shallot

Therefore, we assigned LFS of *A. cepa* to chromosome 5A. Genetic mapping of polymorphisms detected by heteroduplex analysis of LFS amplicons in the *A. cepa* x *A. roylei* interspecific cross revealed co-segregation with chromosome 5 markers (Figure 2).

The LFS marker was linked to several other markers developed to onion ESTs, most notably the SNP marker ACP052. This was designed to partial ESTs showing homology to N terminal regions of group I sucrose transporters (BRAUN and SLEWINSKI 2009). The RFLP marker API66C-E5 mapped in the 'BYG15-23 x AC43' intraspecific onion population (KING *et al.* 1998) was revealed by the cDNA probe API66 (GenBank

Accession BE205593.1) which is homologous to sucrose transporter group I and III proteins. It is not known yet whether ACP052 and API66 markers target the same gene or linked duplications. We previously showed close linkage of API66 markers to a well-supported QTL affecting bulb dry matter (GALMARINI *et al.* 2001; MARTIN *et al.* 2005; MASUZAKI *et al.* 2007; MCCALLUM *et al.* 2007), suggesting that the LFS loci are in close linkage with this QTL.

Masamura *et al.* (2011) showed that anonymous regulatory genes related to the regulation of the expression of the ACSO biosynthesis pathway were located on chromosome 5.

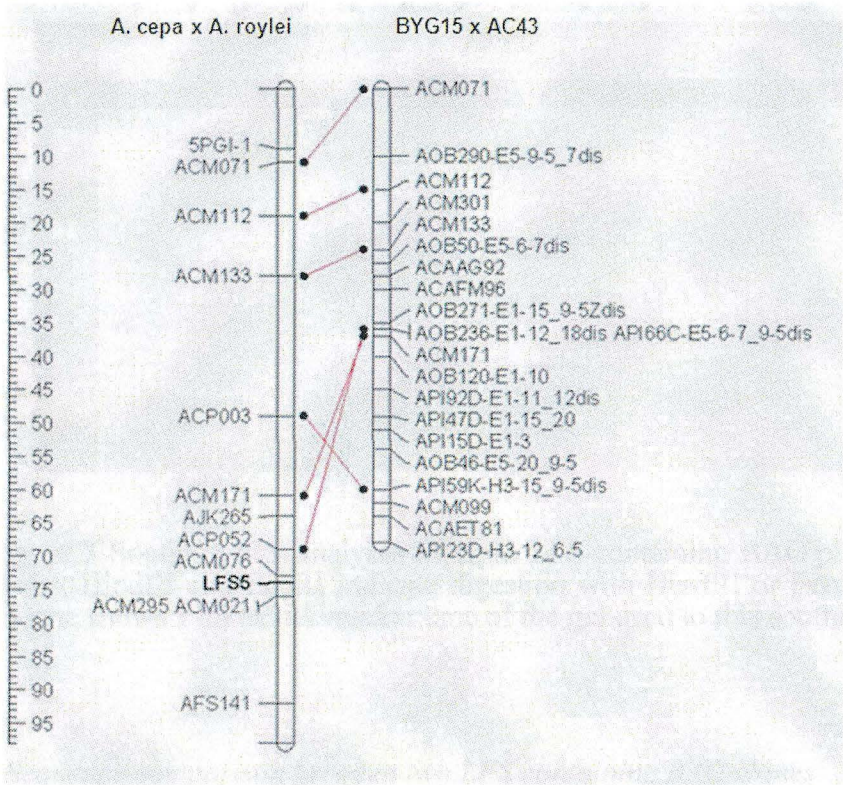


Figure 2 Genetic mapping of LFS heteroduplex marker LFS5 to chromsome 5 in the *A. cepa* x *A. roylei* population and alignment with the onion linkage map (BYG15 x AC43) of Martin *et al.* (2005).
The scale denotes the recombination distance in Kosambi units. AFLP markers in the interspecific map have been deleted for clarity.

Selection and characterization of BAC clones containing LFS gene

PCR screening of the BAC library revealed eight positive clones, from a total of 48,000 clones corresponding to 0.32 genome equivalent (SUZUKI *et al.* 2002). Southern blot analysis of the positive clones was carried out by using a DIG-labeled LFS probe (Figure 3). In *Hin*dIII digestion, all eight clones showed signal in similar size but *Eco*RI digestion, showed a larger size in 4F10/155 compared to the other seven clones. This observation suggested that there are at least two LFS loci in onion genome.

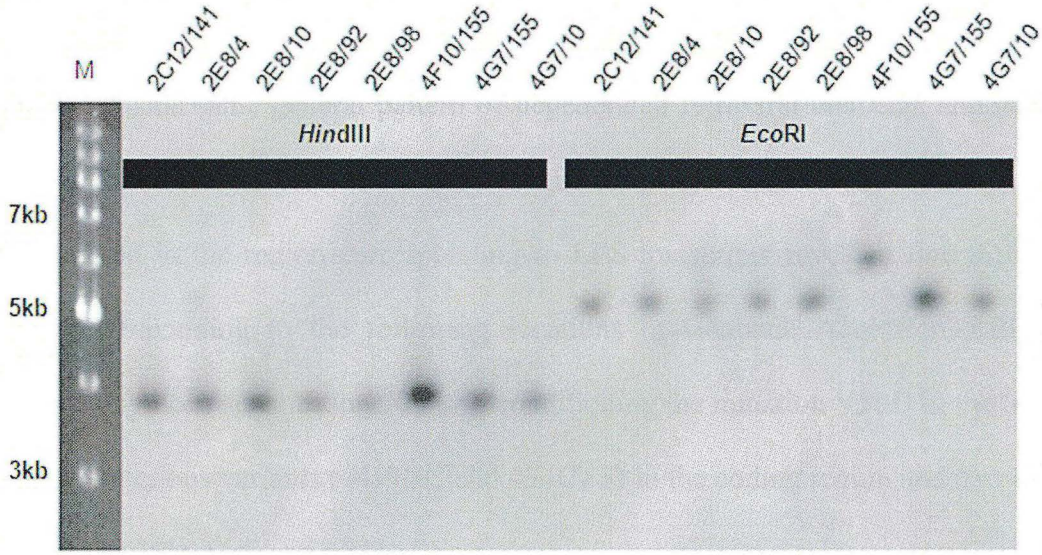


Figure 3 Southern blot analyses of eight LFS-containing BAC plasmid with LFS probe. HindIII and EcoRI indicate digestion with HindIII or EcoRI, respectively. M lane shows 1 kb ladder marker lane of the gel used to this southern blot.

Sequence comparison between two LFS containing BAC clones:

Sanger and 454 shotgun sequencing of clones 2E8/10 and 4F10/155 provided contigs of 19 and 7 which were longer than 500 bp respectively. From Sanger shotgun analysis of 2E8/10, a LFS ORF sequence was detected in a contig containing 8,623 bp and sequence of 20,577 bp (from upper 13 kbp to lower 7 kbp of LFS ORF) was built by assembling

contigs with 5 - 6x coverage (total of 138,394 bp) of 4F10/155, an LFS ORF sequence was observed in the longest contig containing 71,391 bp.

Nucleotide sequence comparison of the approximately 20 kbp contigs from both BAC exhibited high similarity over 10 kbp flanking the intronless LFS gene. Notable areas of lower similarity within this region were the region immediately upstream of the LFS gene, including its promoter, and a 1.5 kbp insertion in 2E8/10 showing BLASTX similarity to polyproteins (Figure 4). Both contigs shared a 445 bp direct repeat in 1 kbp upstream and a region of homology to other onion BAC sequences in 8–14 kbp upstream of LFS. Regions 3–6 kbp downstream showed BLASTX homology to plant *Ty1*-copia like elements. This same general pattern of degenerated retroviral elements and transposons was reported in earlier onion BAC sequencing by Jakše *et al.* (2008). Sequence comparison in the region corresponding to LFS transcripts revealed five differences at sites corresponding to the following locations in Genbank Accession AB094593: a non-synonymous substitution 203A>T conditioning the mutation V50D in the translation; two synonymous variants (441T>C and 474C>T) in the coding region and two variants in the 3'UTR (596A>G and 654T>C).

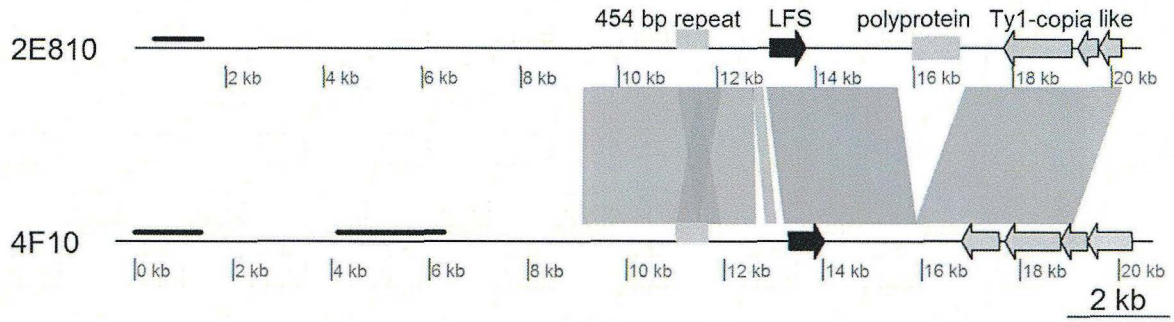


Figure 4 Sequence similarity and annotation of LFS flanking regions of 2E8/10 and 410/155. Gray-scale shading denotes sequence similarity over 80% detected by BLASTN alignment of sequences. Black arrow indicates intronless LFS ORF. Gray side-bar denotes a 454 bp direct tandem repeat. Gray bar denotes insertion in 2E810 sequence with BLASTX homology to polyproteins and gray arrows denote regions with BLASTX similarity to Ty1-copia like sequences. Black lines denote regions with BLASTN similarity to other onion BAC sequences.

Chromosomal localizations of BAC clones

The DIG-labeled BAC clones, 2E8/10 and 4F10/155 were hybridized to the mitotic metaphase chromosome preparation of *A. cepa*. By using the Cot-100 fraction to block the repetitive sequence hybridization, both BAC clones generated distinct signals from a single pair of somatic metaphase chromosomes (Figure 5A). The karyotype analysis revealed that signals were located on the proximal region of the long arm of chromosome 5 (Figure 5B). The identity of this chromosome was established on the basis of its chromosome size and the position of its centromere (Relative chromosome length 12.7±1.0, Centromere index 48.7±0.7; de Vries 1990). Each BAC's location from the centromere was measured (20 chromosomes from 10 metaphases per each BAC) and their relative positions were estimated. The position of 2E8/10 was 0.31±0.03 and the position of 4F10/155 was 0.32±0.03 (Figure 5C). No statistically significant difference was found between the position of hybridization signals in the two BAC clones (Student *t*-test, td = 0.195, tst = 2.04, n = 40, $P = 0.05$).

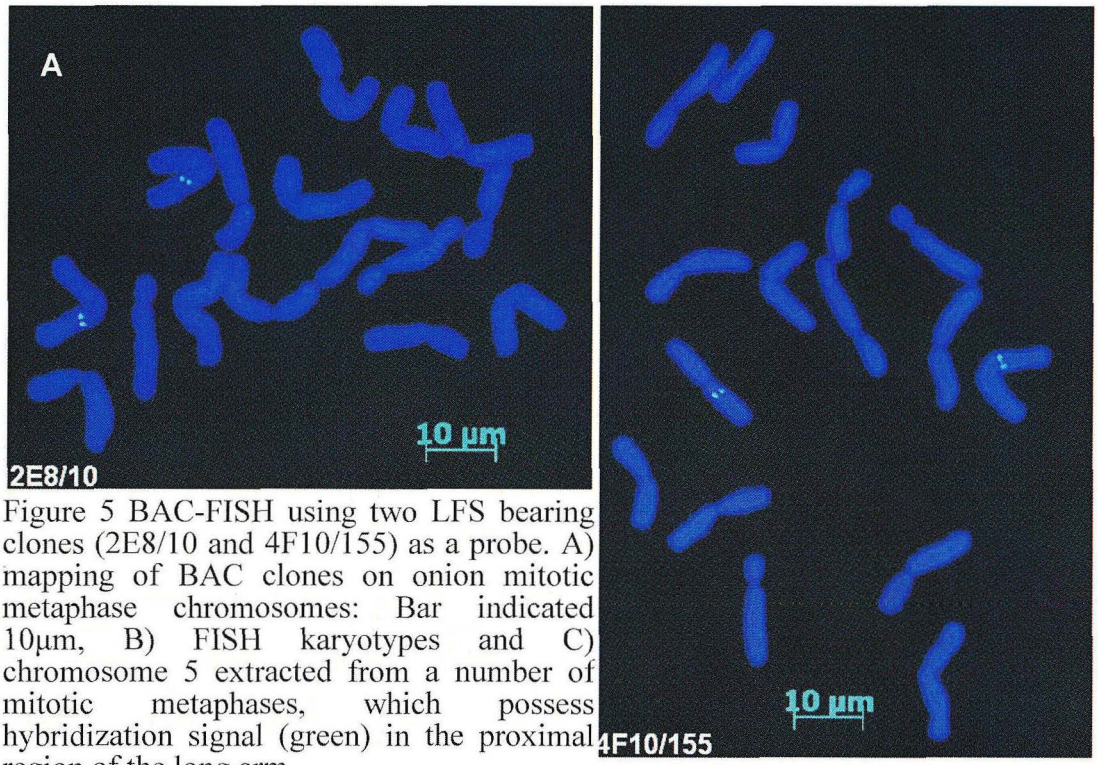
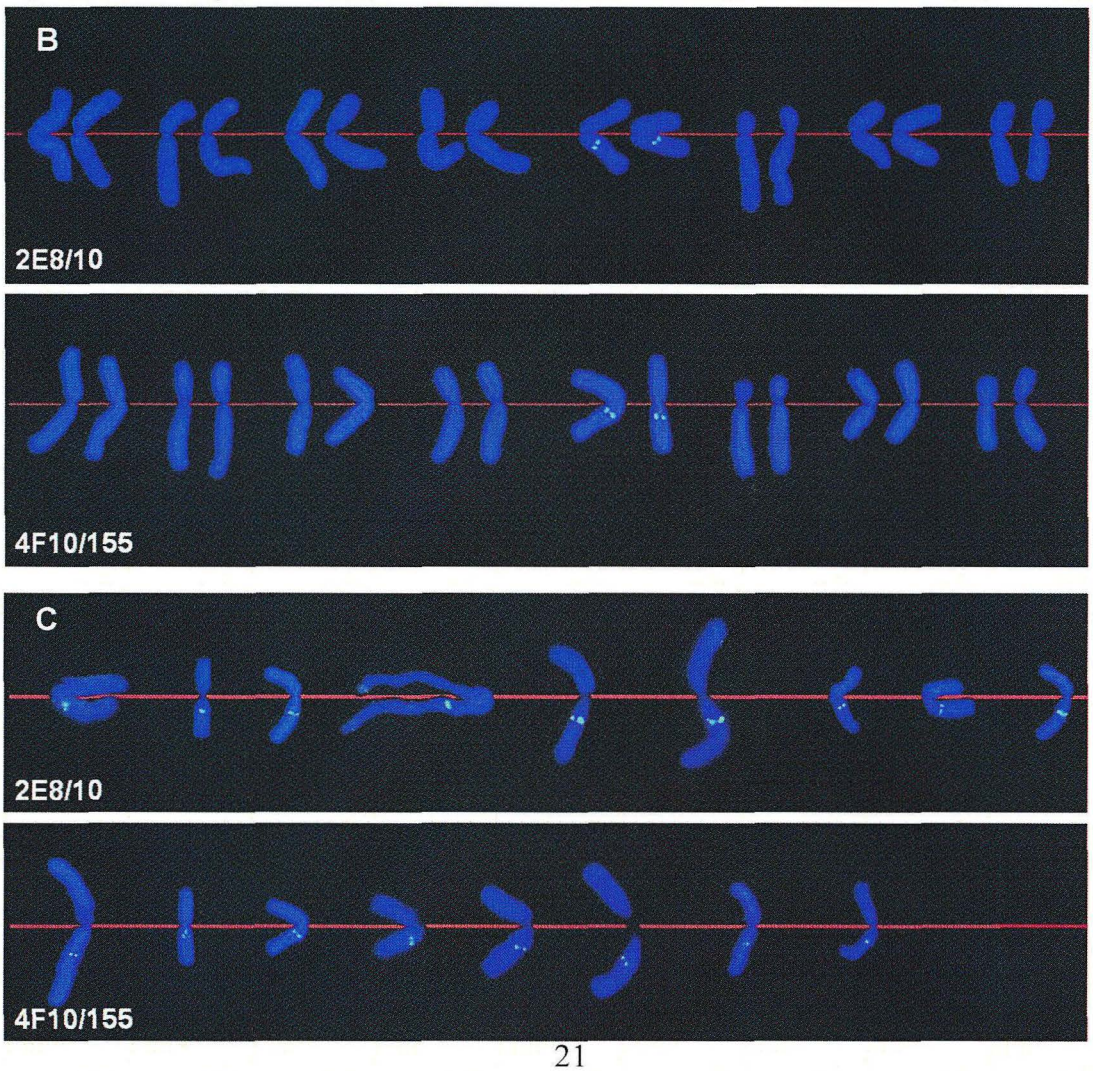


Figure 5 BAC-FISH using two LFS bearing clones (2E8/10 and 4F10/155) as a probe. A) mapping of BAC clones on onion mitotic metaphase chromosomes: Bar indicated 10µm, B) FISH karyotypes and C) chromosome 5 extracted from a number of mitotic metaphases, which possess hybridization signal (green) in the proximal region of the long arm.



Sequence variation in LFS amplicons and transcripts

Mapping leaf transcript 454 reads from the doubled haploid line 'CUDH2150' to *A. cepa* LFS cDNA sequences (Genbank accessions AB094593, AB089203) with stringent criteria revealed 119 matching reads containing three variants at intermediate frequencies (Table 2). The same variants were also observed in 'Nasik Red' and formed two haplotypes. In a previous study (McCallum *et al.* 2008) we surveyed allelic variation at multiple SSR loci and failed to detect any heterozygosity in 'CUDH2150', suggesting that the observed variants are transcripts from duplicated LFS genes. BLAST searches of Genbank EST division revealed that in addition to ESTs with the same haplotype as AB094593 ('haplotype 1'), three onion ESTs exhibited the alternate haplotype ('haplotype 2') at these three sites (accession numbers CF451348, FK935151, FK936343).

Table 2 Counts of 454 cDNA reads from 'CUDH2150' and 'Nasik Red' cDNA 454 cDNA sequencing classified by variants.

Site[a]	AB089203	Variant	Total Depth		Variant frequency %	
			CUDH2150	Nasik Red	CUDH2150	Nasik Red
439	T	C	106	37	41	49
594	A	G	75	39	48	23
652	T	C	62	31	50	26

a Site positions are relative to Genbank Acc AB089203.

Sanger sequencing of LFS amplicons from a range of diverse onion germplasm and *A. roylei* revealed these sequence variants as well as rarer ones (Figure 6). The most commonly observed haplotype was that matching AB094593 ('haplotype 1'), and this was amplified from all onion populations surveyed. Trees based on alignment of cDNA and genomic sequences (Figure 7) revealed clear clustering of onion sequences into two groups corresponding to haplotypes 1 and 2 observed in 454 sequencing and an outgroup containing other *Allium* species. The variant distinguishing AB089203 from other LFS

sequences (A570T) was not observed in any other reads suggesting this is a PCR or sequencing error. Four singleton reads with single base differences to the haplotype 1 sequence were observed, one of which (HQ738863) conditioned a non synonymous M > I mutation. By contrast, several well-supported haplotypes were observed within the 'haplotype 2' group corresponding to BAC 2E8/10. And a non-synonymous variant (HQ738883, V > G at 166) was also found in the group. Average nucleotide diversity was higher in the haplotype 2 group ($\pi = 0.0042$) than in haplotype 1($\pi = 0.0028$).

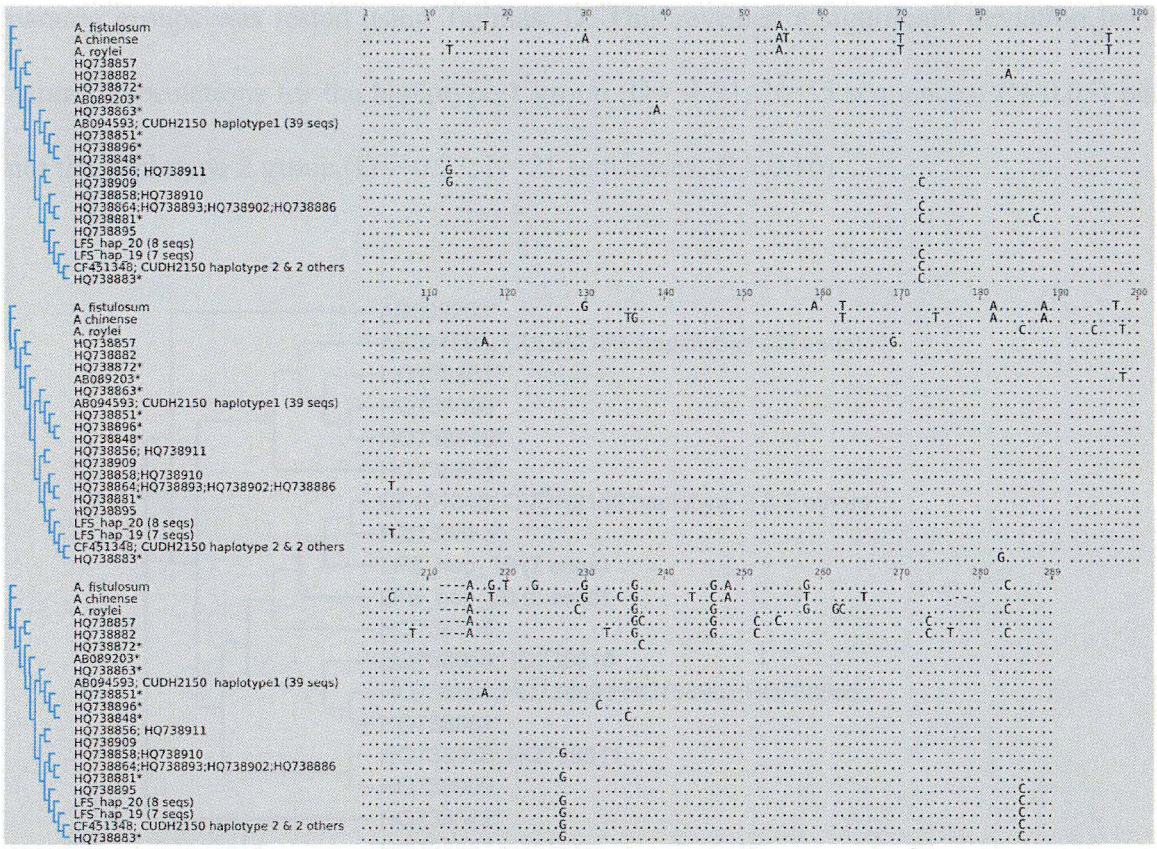


Figure 6 Alignment of cDNA sequences from onion and related Allium species with haplotypes observed in region corresponding to bases 370-658 of GB accession AB089203 in genomic PCR amplicons of LFS genes from diverse onion germplasm. Accession numbers marked by asterisk denotes haplotypes based on sequence variants only supported by a single read, which may represent PCR errors.

The out-group containing related *Allium* species also contained two sequences amplified from *A. cepa* 'Faridpuri' and 'AC43' (HQ738857, HQ738882,). Notably, these sequences exhibit a deletion corresponding to bases 578–581 of AB089203 and other variants in the 3' UTR that observed in LFS cDNA sequences from *A. roylei, A. chinense* and *A. fistulosum.* HQ738882 contains a premature stop (G450A) within the putative coding region. These may represent rare or ancestral alleles but could also be products amplified from other LFS loci.

A test of the hypothesis of selective neutrality in this sequence region in the two paralog groups was tested using Tajimas' D. This revealed a significant deviation from neutral expectations for the haplotype 1 group (D=−1.91; N=55 sequences; $P< 0.05$) but not the haplotype 2 group (D=−0.23; N=21 sequences; $P > 0.10$)

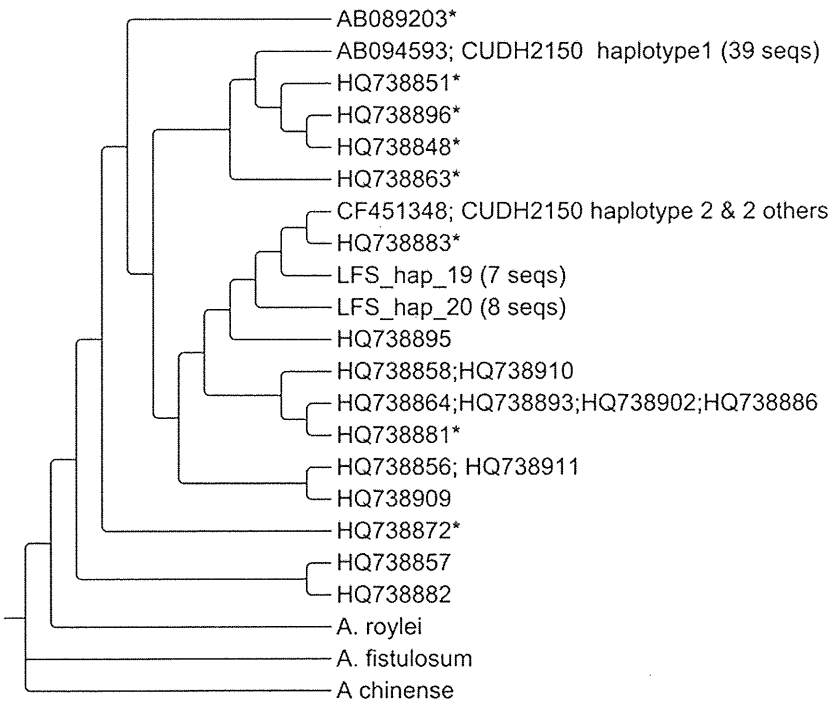


Figure 7 Cladogram of LFS sequences based on consensus maximum likelihood tree of LFS sequence alignments with 100 bootstraps

Existence of distinctive LFS haplotypes in the FF+5A

The A594G and T652C variants distinguishing the two haplotype groups condition *Nsi*I and *Sph*I restriction sites respectively and we used this to test for existence of distinctive haplotypes in the FF+5A. This revealed both *Nsi*I and *Sph*I digestion of the LFS product amplified from this line but no evidence for double digestion (Figure 8). This confirms that the two putative paralogous loci are both located on onion chromosome 5.

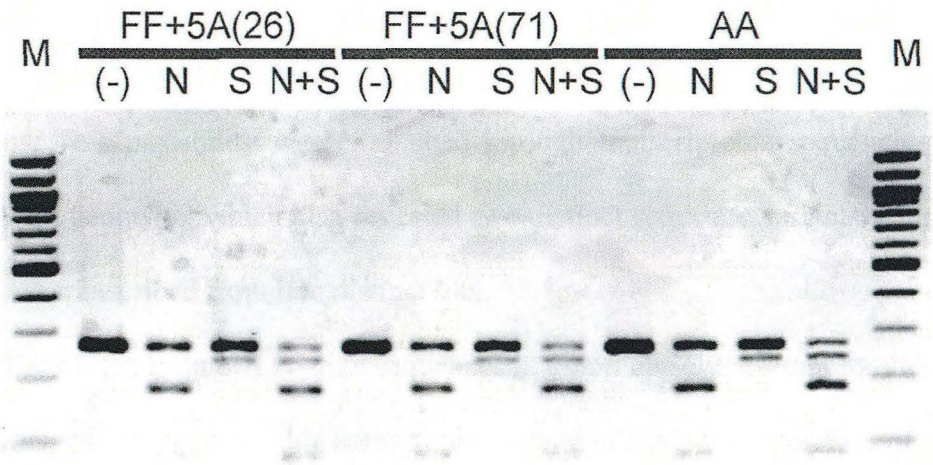


Figure 8 Restriction analysis of LFS PCR products from chromosome 5 AMALs. Lanes 1-4 FF+5A (plant number 26); Lanes 5-8 FF+5A (71); Lanes 9-12 AA ('Chiang Mai'). (-), N, S, and N+S indicate uncut PCR product, NsiI digested, SphI digest and double digest by NsiI + SphI. M shows 100 bp ladder marker. The two AMALs were the sibs obtained from a single cross.

DISCUSSION

By the use of multiple approaches this study has determined that LFS in onion is transcribed from at least two loci and that they are localized on chromosome 5. Southern blotting of eight LFS bearing clones revealed two distinct *Eco*RI RFLP patterns but assignments using alien addition lines and mapping analysis in an inter-specific cross

placed the gene only on chromosome 5. Furthermore, BAC-FISH study showed co-localization of these BACs. Under the optical limit of a 1.4 numerical aperture for a conventional microscope objective, a maximum distance of only 0.2 μm can be resolved. De Jong *et al.* (1999) reported that mitotic metaphase FISH can resolve 4–5 Mbp in a tomato. The level of metaphase chromosome condensation in *Allium* is about five times higher than in tomato metaphase chromosomes (KHRUSTALEVA and KIK 2001). Thus, 2E8/10 and 4F10/155 clones can be distant from each other up to 25 Mbp and still be located in the same position on the mitotic metaphase chromosome. Because onion is highly heterozygous, a possible interpretation of these results is heterozygosity at the LFS locus. This possibility was eliminated through transcriptome sequencing in a doubled haploid genotype, which also revealed two distinct sequence variants. This confirms that LFS is transcribed from two distinct loci.

The observation of LFS-like sequences in onion closely resembling those from related *Allium* species may provide some further scope to study the evolution of LFS in *Allium*. These sequences may be rare or ancestral alleles at either of the major loci but could also be cross-amplified from other LFS loci. It is possible that such sequences could represent ancestral copies of the gene family, and indicate that gene duplication is still ongoing in onion genome. The LFS gene is small and intronless. In plants, small and intronless genes are frequently plant- or lineage-specific (JAIN *et al.* 2008). It is therefore plausible that LFS is such an example of a novel, genus-specific gene family that has arisen recently and provided a strong selective advantage. A survey by PCR across diverse germplasm suggested lower nucleotide diversity in haplotype group 1 in the region surveyed compared to the haplotype group 2 corresponding to BAC 2E8/10 locus. This is suggestive of possible differences in selective constraints but wider sampling of

nucleotide diversity is required to test the role of purifying selection and gene conversion on these loci.

Gene duplication is known to be an important source of evolutionary innovation and adaptation (DES MARAIS and RAUSHER 2008). Mapping studies have suggested extensive duplication in onion. (KING *et al.* 1998) reported that 21% of 91 cDNA probes of RFLP detected more than one segregating RFLP, of which 53% were unlinked, 47% were linked. These findings indicated gene duplication has occurred with high frequency in onion genome, the ratio of the linked duplicated RFLP were higher than other plants. Our results demonstrate that LFS genes exist as functional linked duplicated genes, implying a low possibility for obtaining null mutants through mutant screening. However, BAC sequence analysis and PCR experiment to distinguish two haplotypes indicated that 2E8/10 was 'haplotype 2' and 4F10/155 was 'haplotype 1', and showed a large difference in upper flanking 1.2 kbp region of LFS between the BACs. These observations of variation in the promoter regions suggest that there may be differential expression between the paralogs. Kim *et al.* (2005) speculated similarly in pink onion. More detailed surveys of paralog-specific expression might reveal mutants with lower LFS activity. The identification of the complete genomic sequence surrounding the two paralogs will now enable wider surveys for natural mutants by deep sequencing, like the variant (HQ738882) possessing premature stop we identified in 'Faridpuri'. Using such genotypes as a material for further mutant induction or breeding, we may achieve production of LFS null onion. A non-synonymous variant found in this article, HQ738883 (V > G) may have enzymatic activity equivalent to wild-type LFS as we have shown a deletion of nine C-terminal amino acids (160–169) had no effect on the enzymatic activity (MASAMURA *et al.* 2012c). It remains unclear whether the non-synonymous

variant HQ738863 (M > I) has enzymatic activity. However, because LFS catalyzes an intra-molecular H^+ substitution reaction, this variant which changes between non-polar amino acids might retain the enzymatic activity of wild type onion LFS. For the same reason, we speculate that the variant (V50D) found in 4F10/155 has a low potential for loss of enzymatic activity.

It has been reported that a well-supported QTL affecting bulb dry matter and a putative sucrose transporter gene are located on chromosome 5 (GALMARINI *et al.* 2001; MARTIN *et al.* 2005; MASUZAKI *et al.* 2007; MCCALLUM *et al.* 2007). Dry matter is an important trait affecting firmness and storability of onion bulb, and therefore survival under human or natural selection. It is notable that LFS is located also in this region, sufficiently close that we would expect some effects on LFS through linkage drag. Consequently this region of chromosome 5 could be an important domestication region in the onion genome and it would be fruitful for further functional and crop evolutionary studies.

These findings provide a highly relevant case study of a duplication leading to multiple functional loci encoding a gene with important adaptive trait in onion.

Chapter III. IDENTIFICATION OF AMINO ACID RESIDUES ESSENTIAL FOR ONION LACHRYMATORY FACTOR SYNTHASE ACTIVITY

INTRODUCTION

While many people appreciate the nutritional and organoleptic values of onions, few enjoy the tear-inducing sensation produced while cutting the bulbs. The substance causing this irritation, a lachrymatory factor (LF) was identified as propanethial *S*-oxide

by Brodnitz and Pascale more than 40 years ago (BRODNITZ and PASCALE 1971).

LF belongs to the group of thioaldehyde *S*-oxides, and to date, only three other naturally occurring thioaldehyde *S*-oxides have been reported: (*Z*,*Z*)-*d*,*l*-2,3-dimethyl-1,4-butanedithial *S*,*S*'-dioxide from onion (BLOCK and BAYER 1990); (*Z*)-phenylmethanethial *S*-oxide from *Petiveria alliacea* (KUBEC *et al.* 2003); and (*E*)-/(*Z*)-butanethial *S*-oxide from *Allium siculum* (KUBEC *et al.* 2010) (Figure 9).

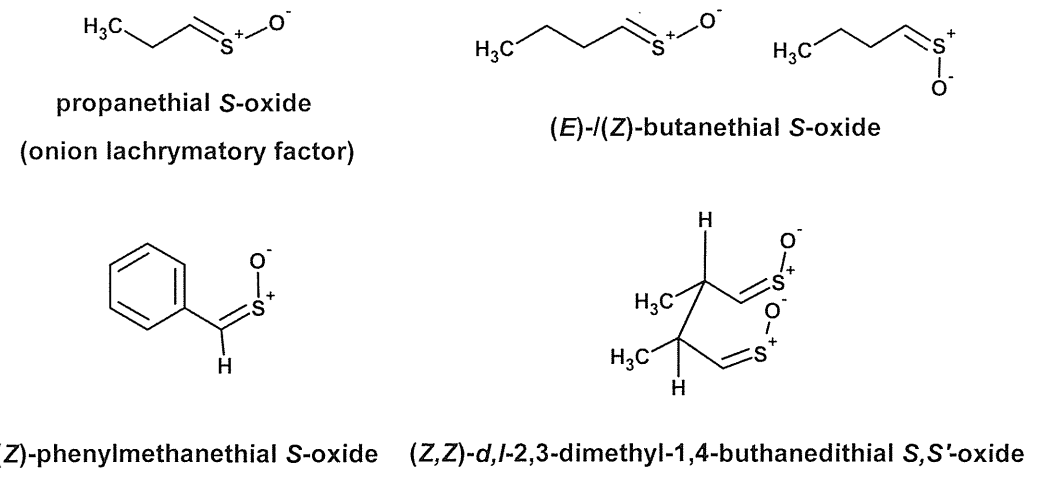


Figure 9 Naturally occurring thioaldehyde *S*-oxides

Onion LF has been widely believed to be non-enzymatically produced from 1-propenylsulfenic acid, a putative reaction product of alliinase and *trans*-(+)-*S*-(1-propenyl)-L-cysteine sulfoxide (*trans*-PRENCSO) that is most abundant in onion (BLOCK *et al.* 1979; BLOCK 1992; SHEN and PARKIN 2000; YAMANE *et al.* 1994). Contrary to this belief, we have discovered a novel enzyme, lachrymatory factor synthase (LFS), that catalyses LF formation after the action of alliinase on *trans*-PRENCSO [(*E*)-PRENCSO], and we have cloned its cDNA (IMAI *et al.* 2002) (Figure 10). Three isozymes, LFS-1, LFS-2, and LFS-3, have been isolated from the onion bulb. Analyses of

their N-terminal amino acid sequences indicated these three proteins to be respectively lacking 16, 14 and 12 amino acids when compared to predicted , to the predicted open reading frame (169 amino acids) of onion LFS cDNA (AB089203) (IMAI *et al.* 2008).

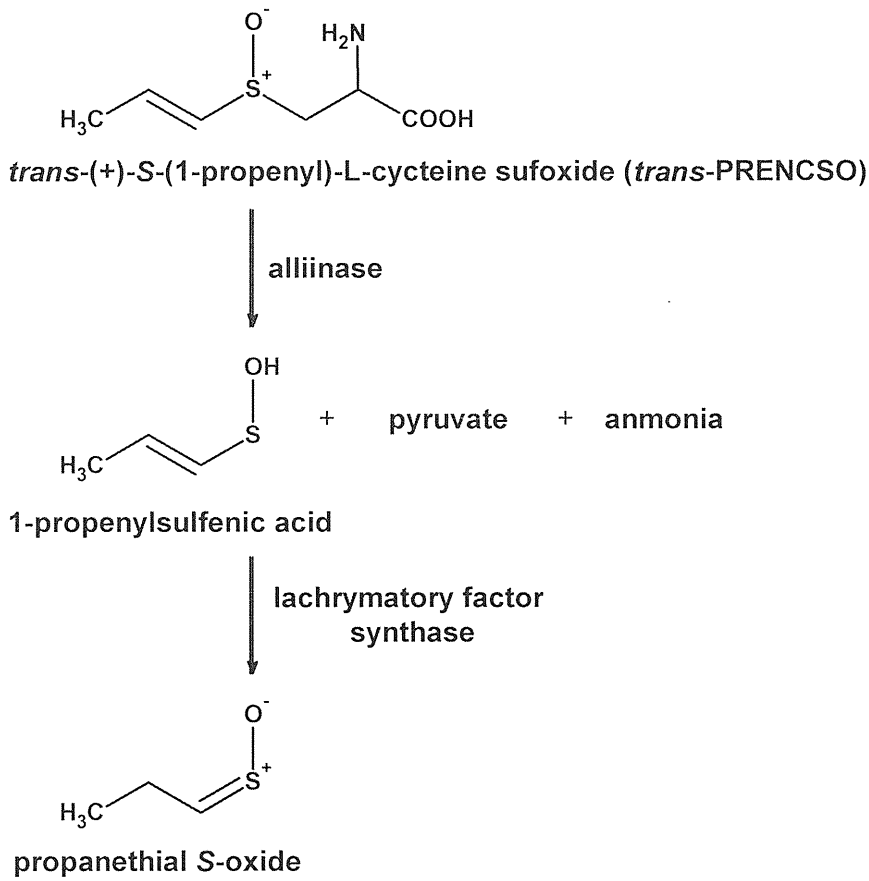


Figure 10 The pathway of onion lachrymatory factor synthesis

LF, the catalytic product of LFS, is a rare compound in nature, and at present, no other enzyme that catalyzes reactions similar to the reaction leading to LF formation have been reported. Musah *et al.* have reported an enzyme that catalyzes the formation of a sulfine LF, (*Z*)-phenylmethanethial *S*-oxide, from the roots of *Petiveria alliacea* (MUSAH *et al.* 2009), but the reaction catalyzed by this enzyme differs considerably from that catalyzed by onion LFS. While onion LF is

produced through an intra-molecular proton (H^+) substitution reaction (BLOCK *et al.* 1996), *Petiveria* LF is produced by a dehydrogenation reaction. Additionally, a database survey failed to identify any known functional domains exhibiting high sequence homology with onion LFS (IMAI *et al.* 2008).

We therefore attempted to identify the catalytic amino acid residues and developed a three-dimensional (3D) model structure of LFS by using SWISS-MODEL as the first step to elucidate the unique catalytic reaction mechanism of this enzyme.

MATERIALS AND METHODS

Materials

Shallot (*Allium ascalonicum*), Japanese bunching onion (*Allium fistulosum*), rakkyo (*Allium chinense*), leek (*Allium porrum*), and elephant garlic (*Allium ampeloprasum*) were purchased from a local market. Garlic alliinase was purified according to the method described by Nock and Mazelis (1989). We prepared *trans*-PRENCSO from onion bulbs according to the procedure described by Shen and Parkin (2000). All the reagents used were of analytical grade.

Obtaining the LFS genes from Allium *species*

Total RNA was extracted and purified from 100 mg of frozen powder of the fresh edible parts of *A. ascalonicum*, *A. fistulosum*, *A. chinense*, *A. porrum*, and *A. ampeloprasum* by using an RNeasy plant mini kit (QIAGEN, Hilden, Germany) according to the manufacturer's protocol. We used the total RNA as the template and synthesized cDNA by using a Ready-To-Go T-primed First-Strand Kit (GE Healthcare

UK Ltd., Buckinghamshire, UK). To determine the nucleotide sequence of the full-length LFS cDNA, we performed a rapid amplification of cDNA ends (RACE) analysis at both the 3′ and 5′ ends by using the LFS-specific primers and either a 3′-end adaptor primer included in the cDNA synthesis kit or a 5′-anchor primer supplied with a 5′-RACE kit (5′-RACE System for Amplification of cDNA Ends, Version 2.0 (Invitrogen Corp., Carlsbad, CA, USA)). All the RACE products were sequenced using an ABI Prism 310 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA), and their nucleotide sequence data were assembled to build the full-length LFS cDNA.

After the ORFs of the LFS cDNAs had been amplified by using primers designed at the initiation codon of the ORF and the adaptor primer of the cDNA synthesis kit, cDNAs were cloned in the pGEM-T Easy TA-cloning vector (Promega Corporation, Madison, WI, USA). All insertions were confirmed by sequencing the multi-cloning site (MCS) region of the TA-cloning vector. The primers used to search for LFS cDNA in *Allium* species are listed in Table 3a.

Preparation of truncated onion LFS cDNA

Six truncated onion LFS genes lacking the 5′ or 3′ sequence were generated by PCR from onion LFS cDNA cloned in the TA-cloning vector, using primers listed in Table 3b. The PCR products were cloned into the pGEM-T Easy TA-cloning vector and sequenced to confirm that the insertions were of the correct length and were in the correct location.

Preparation of onion LFS cDNA mutants

Site-directed mutagenesis of the onion LFS cDNA cloned in pGEX4T-3 GST fusion

expression vector (GE Healthcare UK Ltd., Buckinghamshire, UK) was achieved by using the Gene-Tailor™ site-directed mutagenesis system (Invitrogen Corp., Carlsbad, CA, USA), according to the manufacturer's protocol. All mutant cDNAs were verified by sequencing. The Primer sets used in the site-directed mutagenesis are listed in Table 3c.

Table 3a list of the primers for searching for LFS cDNA in allium species

We used the primers listed in this table to obtain LFS cDNA from various allium plant in this study.

	primer	allium plants
3'RACE		
adaptor primer		
	5'-TGG AAG AAT TCG CGG CCG CAG-3' (NotI primer)	
lfs sequence specific primers		
	5'-TGG AGG GTC CTG AGC ACA AG-3'	*A. fistulosum A.chinense A.ascalonicum* [1]
	5'-GGI GCI MGI AAR TGG-3' (E2-1N primer)	*A.ascalonicum* [1]
	5'-GAA TTT TGG GCC AAG GAG AAG CTG G-3'	*A. ampelopras A. porrum*
5'RACE		
anchor primer		
	5'- GGC CAC GCG TCG ACT AGT ACG GGI IGG GII GGG IIG -3'	
lfs sequence specific primers		
	5'- CTC TTC GAT TTT CTG ACC TAT CTC AGT AGC -3'	*A. fistulosum A.chinense A.ascalonicum*
	5'- TCC TCG TAC CCT GTA AAA CAC TCA G -3'	*A. ampeloprasum*
	5'- CAC ACA GCA TCA CAA ATT GAC-3'	*A. porrum*
LFS ORF region amplified from cDNA [2]		
lfs sequence specific primers		
	5'-ATG GAG CTA AAT CCT GGT GCG C-3'	*A. fistulosum*
	5'-ATG GAG CAA AAT TCT GGT ACG C-3'	*A.chinense*
	5'-ATG GAG CTA AAT CCT GGT GCA C-3'	*A.ascalonicum*
	5'-ATG GCG CAA AAT CCT GGT GTG C-3'	*A. porrum*
	5'-ATG ATG ACA TAT CCT GGA AAT CG-3'	*A. ampeloprasum*

Table 3b The primers for preparation of truncated onion LFS cDNA

We used the primers listed in this table to prepare terminal truncated LFS cDNA from onion LFS cDNA.

	primer	mutants
truncated onion LFS [4]		
N-terminal deletion mutant		
	5'-GGCAAAGTCCATGCTTTGCTTCC-3'	*LFS/23-169 mutant*
	5'-AATACAAAGCCAGAGCAAGCATGG-3'	*LFS/31-169 mutant*
	5'-ATGCCTTCGTTGTCAGTCTGTG-3'	*LFS/51-169 mutant*
	5'-TTCAAGATCAAGCACTGCAAACCTC-3'	*antisense primer*
C-terminal deletion mutant		
	5'-GATAGTGCTAACGGAGCTCGAAAATGG-3'	*sense primer*
	5'-TCAACCTATCTCAGTAGCCC-3'	*LFS/13-160 mutant*
	5'-TCAGATATACTTGCACTGAAAAGACC-3'	*LFS/13-140 mutant*
	5'-TCATTGCATGGTAGCCGTG-3'	LFS/13-119 mutant

[1] Nested PCR was applied to *A. ascalonicum* lfs cDNA. E2-1N primer was used in 1st step pcr.

[2] Not I primer was used as a reverse primer in all amplification.

[4] For subcloning of amplified DNA fragments into the expression vector from the TA-cloning vector, EcoR I recognition sequences were added to all of primers used.

[4] In addition, the reverse primers that were amplified in the C-terminal deletion mutants were designed to contain the terminal codon at their 5'-end.

Table 3c The primers for preparation of site-directed mutant onion LFS cDNA

We used the primers listed in this table to prepare amino acid substituted LFS cDNA from onion LFS cDNA.

forward primer	reverse primer	position of substituted amino acids in onion LFS	native amino acids	mutaint amino acids
5'-TGCTTTGCTTCCAAATACACTGCCACAGCAAGCATGGAC-3'	5'-TGTATTTGGAAGCAAAGCATGGACTTTGC-3'	33, 35	KPE	LPQ
5'-CAAGCATGGACACTACTACTGGACTTTATTAACC-3'	5'-TAGTAGTGTCCATGCTTGCTCTGGCTTTGT-3'	42	K	L
5'-GCATGGACACTACTAAAGAACTTTATTAACC-3'	5'-TTTAGTAGTGTCCATGCTTGCTCTGGC-3'	43	D	N
5'-CTTTATTAACCTTCACCTGGTCATGCCTGCTTTGTCAGTCTGTG-3'	5'-GTGAAGGTTAATAAAGTCTTTTAGTAGTG-3'	49, 53	KVMPS	LVMPA
5'-GGTCATGCCTTCGTTGGCTGTCAGCGAACTGGTAGAAG-3'	5'-CAACGAAGGCATGACCTTGTGAAGGT-3'	55, 57	SVC	AVS
5'-GCCTTCGTTGTCAGTCTGTCAACTGGTAGAAG-3'	5'-ACAGACTGACAACGAAGGCATGACCTTGTG-3'	58	E	Q
5'-TCGTTGTCAGTCTGTGAACTGGTACAAGGTGAGGCCA-3'	5'-ACAGACTGACAACGAAGGCATGACCTTGTG-3'	61	E	Q
5'-GCCAATGTTGTTGGTAGTGTTCTGTACGTTCTGGGTATAATGC-3'	5'-ACCAACAACATTGGCCTCACCTTCTAC-3'	69, 71, 74	CVRYVK	SVLYVL
5'-GCCAATGTTGTTGGTAGTGTTCGCTACG-3'	5'-ACCAACAACATTGGCCTCACCTTCTAC-3'	69	C	S
5'-GCCAATGTTGTTGGTTGTGTTCTGTACGTTAAAGG-3'	5'-ACCAACAACATTGGCCTCACCTTCTAC-3'	71	R	L
5'-GGTTGTGTTCGCTACGTTCTGGGTATAATGC-3'	5'-AACGTAGCGAACACAACCAACAACATTGGC-3'	74	K	L
5'-GTTAAAGGTATAATGCTCCAAATACAACAGCAATTTTGGGCCA-3'	5'-CATTATACCTTTAACGTAGCGAACAC-3'	78, 81, 82, 83	HPIEEE	LPIQQQ
5'-GAATTTTGGGCCAAGCAGAAGCTGGTGGCGCTGAACAATAAGAACATG-3'	5'-CTTGGCCCAAAATTCCTCTTCTATTG-3'	88, 94	EKLVALD	QKLVALN
5'-GAATTTTGGGCCAAGCAGAAGCTGGTGG-3'	5'-CTTGGCCCAAAATTCCTCTTCTATTG-3'	88	E	Q
5'-GAAGCTGGTGGCGCTGAACAATAAGAACATG-3'	5'-CAGCGCCACCAGCTTCTCCTTGGCCC-3'	94	D	N
5'-TTTACTGAGTGTTTTGCTGGGTACCAGGATTTCACGGCTACCATG-3'	5'-AAAACACTCAGTAAAAATATAACTGTAGC-3'	109, 112, 114	TGYEDY	AGYQDF
5'-ACCATGCAAATAGTGCAGGGTCCTCAGCTCCTGGGAAGTAGAT-3'	5'-CACTATTTGCATGGTAGCCGTGTAATC-3'	122, 125, 126, 127	EGPEHK	QGPQLL
5'-GAGCACAAGGGAAGTCTGTTTGACTGGGCTTTTCAGTGCAAG-3'	5'-ACTTCCCTTGTGCTCAGGACCCTCCAC-3'	130, 134	RFDWS	LFDWA
5'-CACAAGGGAAGTAGATTTAACTGGTCTTTTC-3'	5'-AAATCTACTTCCCTTGTGCTCAGGACCC-3'	132	D	N
5'-GACTGGTCTTTTCAGAGCTTGTATATCCAGGGTATGACTG-3'	5'-CTGAAAAGACCAGTCAAATCTACTTCC-3'	137, 138, 141	CKYIE	SLYIQ
5'-CTGCAGCATTGGGCTACTCAAATAGGTCAGA-3'	5'-AGTAGCCCAATGCTGCAGAATCTCGGTGAA-3'	158	E	Q

Protein production and purification

For each *Allium* species, the full-length LFS cDNA from each *Allium* species, the cDNA corresponding to the truncated onion LFS mutants and site-directed mutants were inserted into the pGEX-4T-3 GST fusion expression vector (GE Healthcare UK Ltd, Buckinghamshire, UK) and transformed into *Escherichia coli* strain BL-21-Gold for protein expression. Centrifuged supernatants of the cell lysates prepared from IPTG-induced *E. coli* cells were used to assay the LFS enzyme activity and determine whether the recombinant protein possessed LF-forming activity. The specific activities of R71L, E88Q, and a mutant with concurrent substitutions at three positions (T109A, E112Q, and Y114F) were compared with those of LFS-3, a wild-type onion LFS. These LFS-GST fusion proteins were purified from the supernatant by using a GST-tag and subsequent

thrombin cleavage to remove the GST sequence. The Protein content was measured by the Bradford method, using BSA as a standard (BRADFORD 1976).

Assay of LFS enzyme activity

Purified garlic alliinase (40 μl, 50 U/ml) was mixed with 10 μl of an assay sample diluted with a 50 mM potassium phosphate buffer (pH 6.5) in a 1.5-ml Eppendorf tube. The *trans*-PRENCSO substrate (20 μl, 20 mg/ml), which had been purified from onion bulbs, was added to this mixture, and the mixture was allowed to stand for three min at room temperature. The LFS enzyme activity was quantified in arbitrary units based on the peak area of LF in HPLC analysis (KAMOI *et al.* 2008; EADY *et al.* 2008). One micro liter of this reaction mixture was applied to a C18 reversed-phase HPLC column (Pegasil ODS 4.6 mm Φ × 250 mm Sensyu Kagaku, Tokyo, Japan) that had been pre-equilibrated with 30% (v/v) acidic methanol (pH 3.3). The column was developed at 35°C at a flow rate of 0.6 ml/min, the eluate being monitored by its absorbance at 254 nm. The retention time of LF, at around 9 min, was determined by a direct injecting a standard solution of authentic LF which had been prepared by fractionation with reversed phase HPLC after diethyl ether extraction from onion juice. The purity of the LF had been confirmed by ^{1}H-NMR spectroscopy. A linear correlation between the amount of LFS and the peak area was confirmed at 20–100 mV·sec (r^2 = 0.99) by using diluted samples of recombinant onion LFS protein. The assay sample was diluted with the same buffer before HPLC analysis so that the peak area was maintained at 20–100 mV·sec.

Prediction of the secondary structure of Allium *LFS*

The secondary structures of the deduced amino acid sequences were predicted using

PSIpred (cms.cs.ucl.ac.uk/typo3/.../PSIPRED/psipred_help.html) (JONES 1999; BRYSON *et al.* 2005).

Homology modeling of LFS

BLASTP searches were made for LFS/23–169 by SWISS-MODEL. The abscisic acid receptor, PYL2, from *Arabidopsis thaliana* showed the closest homology with LFS/23–169 among the proteins with a 3D structure present in the Protein Database. A structural model of LFS was generated with the SWISS-MODEL server by using the X-ray crystallographic structure of PYL2 (PDB code: 3KDH) as the template, and then visualized with MOLMOL (SCHWEDE *et al.* 2003; ARNOLD *et al.* 2006; GUEX and PEITSCH 1997)

RESULTS

LFS cDNAs in lachrymatory Allium *species*

We isolated in this study LFS cDNAs from five *Allium* species: *A. ascalonicum*, *A. fistulosum*, *A. chinense*, *A. porrum*, and *A. ampeloprasum*. Two slightly different sequences were obtained from *A. fistulosum* and *A. porrum,* so we prepared eight LFS cDNAs from lachrymatory alliaceous plants, including one from *A. cepa* (IMAI *et al.* 2002). The sequences obtained were deposited in the DDBJ database under the accession numbers listed in Table 4. Alignment of the deduced amino acid sequences of the eight LFS

Table 4 LFS cDNA sequence accecssion numbers from *allium* species

Allium species	comom name	DDBJ Accession
A. fistulosum	Japanese bunching onion	AB094590
		AB094591
A. chinense	rakkyo	AB094592
A. ascalonicum [a)]	shallot	AB094593
A. porrum [b)]	leek	AB094594
		AB094595
A. ampeloprasum	elephant garlic	AB094596
A. cepa [c)]	onion	AB089203

[a)] *A. cepa* is used as a synonym

[b)] *A. ampeloprasum* is used as a synonym

[c)] onion lfs cDNA sequence published previously (Imai *et al.* 20

cDNAs showed high sequence homology, with the lowest (65.9%) homology found between *A. porrum* and *A. ampeloprasum* (Figure 11), although these have been commonly classified as the same or related species on the basis of their morphological characteristics (VAN DER MEER and HANELT 1990). It is notable that 96 of the 169 amino acid residues in the full-length onion LFS were conserved in all determined LFS sequences.

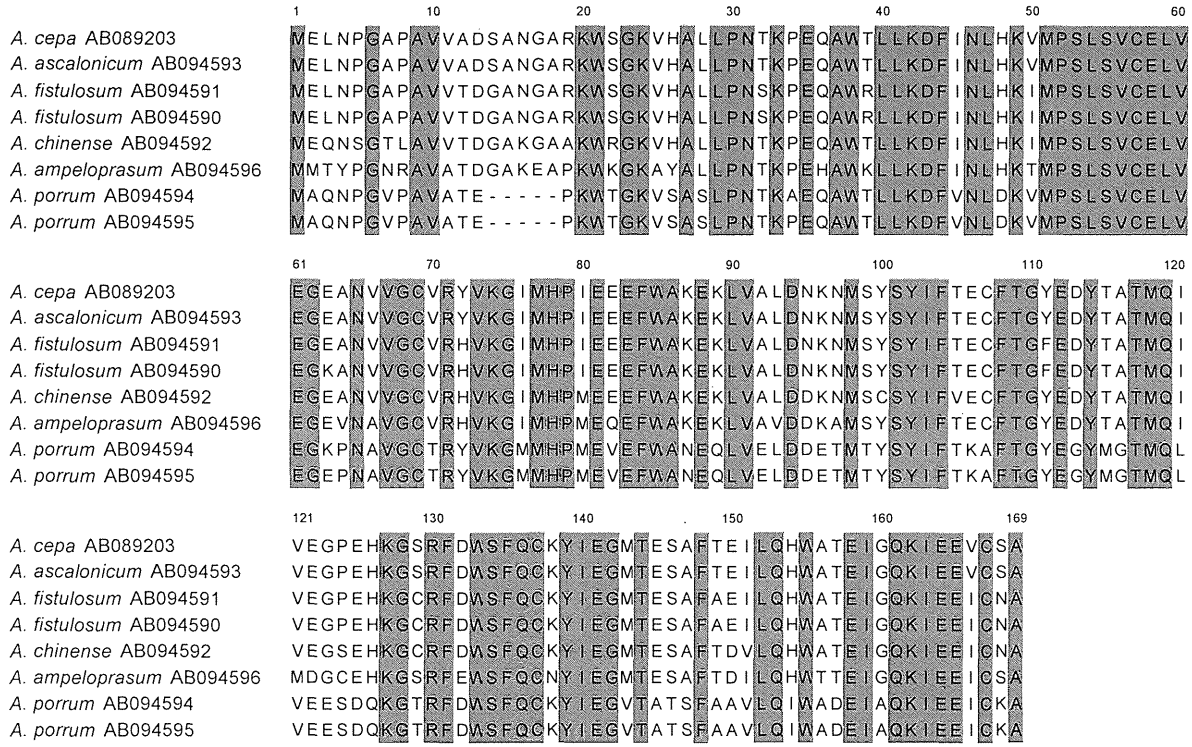


Figure 11 Comparison of the deduced amino acid sequences of LFS from several *Allium* plants
The numbering is from the N terminus methionine of the onion LFS. amino acid residues conserved across all the LFSs investigated are shown in gray.

All LFSs deduced from amino acid sequences of cDNAs from lachrymatory alliaceous plants were predicted to have a similar secondary structure (Figure 12). However we were unable to allocate a specific region that could be the catalytic domain for the LFS activity by sequence comparison alone, since the conserved amino acids were evenly distributed over the entire length of the LFS sequences.

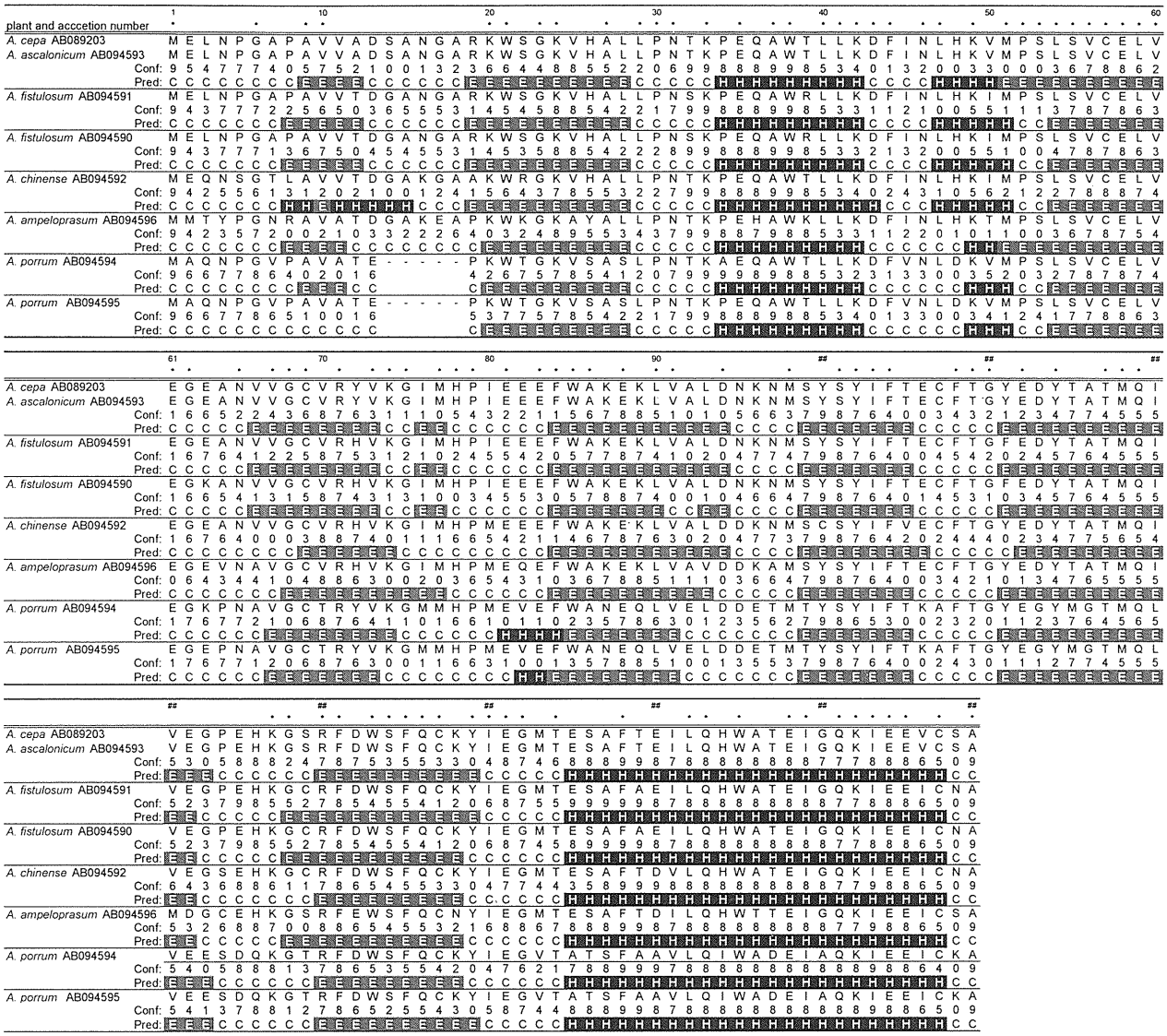


Figure 12 Predicted secondary structure of *Allium* LFS proteins
The numbering is the same as shown in Fig. 3 and * indicates amino acid residues conserved across all the LFSs investigated. "Pred" indicates the predicted secondary structure; H = helix, E = sheet, C = coil. Helix regions and sheet regions are shown in black and gray, respectively. "Conf" indicates confidence of prediction where 0 = low and 9 = high.

LFS enzyme activity of the truncated onion LFS proteins

A series of six deletion mutants of the onion LFS cDNA lacking N-terminal or C-terminal regions (Figure 13) were expressed in *E. coli*, and their activities were measured. Only the LFS/13–160 mutant protein lacking nine amino acids (residues 161–169) at the C-terminus, and the LFS/23–169 mutant protein lacking residues 1–22 at

the N-terminus were found to retain enzymatic activity (their specific activity were 30.6 mV·sec/ng of protein (LFS/13-160) and 310.8 mV·sec/ng of protein (LFS/23–169)). Four other deletion mutant proteins, LFS/13–140, LFS/13–119, LFS/31–169, and LFS/51–169, were obtained as insoluble inclusion bodies and did not show LFS activity. Dissolving the inclusion bodies with 8 M urea and a subsequent refolding treatment by gradual removing the urea failed to restore the enzymatic activity. These results suggest that residues 23–160 of onion LFS were required for enzymatic activity.

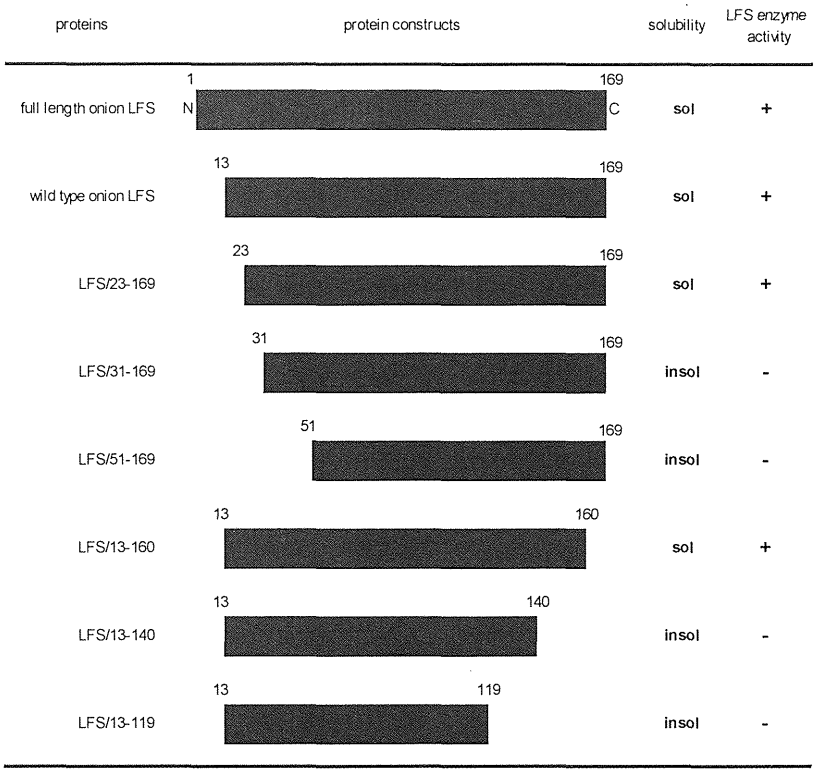


Figure 13 Schematic representation of onion LFS terminal deletion mutants, their solubility, and LFS enzyme activity
The first bar shows the full-length onion LFS (169 a.a.), the sequence of which was deduced from onion LFS cDNA (AB089203). The second bar shows LFS-3, a wild-type onion LFS isolated from the onion bulb lacking 12 amino acids in the N-terminus of the full-length onion LFS. N and C beside the first bar indicate N-terminus and C-terminus of the full-length LFS, respectively. The third to the fifth bar represent N-terminal deletion mutants, and the sixth to the eighth bar represent C-terminal deletion mutants. LFS/m-n represents the amino acids in each mutant protein. The numbers at both the ends of each bar indicate the amino acids positions from the N-terminus of full-length LFS. Solubility and LFS enzyme activity of each deletion mutant are shown on the right.

Effects of site-directed mutagenesis on the LFS enzymatic activity of onion LFS protein

We substituted amino acid residues selected from residues 23–160 of the onion LFS by site-directed mutagenesis to determine which amino acid residues were important for enzymatic activity. Since LFS was considered to catalyze an intra-molecular proton (H^+) substitution reaction in the substrate, 1-propenylsulfenic acid (MUSAH *et al.* 2009), we speculated that amino acids with an acidic side chain (i.e., glutamic acid and aspartic acid) would be primary candidates for the catalytic residues of LFS. We also selected those amino acids identified as catalytic residues in other enzymes (i.e., arginine, cysteine, lysine, histidine, threonine, serine, and tyrosine) as secondary candidates (GUTTERIDGE and THORNTON 2005). Each of these target amino acids was replaced by an amino acid with a side chain of similar size and properties; i.e., glutamic acid with glutamine, aspartic acid with asparagine, cysteine with serine, arginine with leucine, lysine with leucine, histidine with leucine, threonine with alanine, serine with alanine, and tyrosine with phenylalanine.

Twenty-one mutagenesis experiments were conducted with amino acid substitutions at 33 positions (Table 5). The mutagenesis produced simultaneous mutations at multiple sites in ten cases. The results of experiments for determining the enzymatic activity of the mutant proteins are summarized in Table 5. Two mutants, one with three simultaneous substitutions (C69S, R71L, and K74L) and the other with two simultaneous substitutions (E88Q and D94N) showed no LFS enzymatic activity. We then tried to determine which of the five substitutions (C69S, R71L, K74L, E88Q and D94N) were responsible for the loss of the enzymatic activity. Only the two mutations, R71 to L71 and E88 to Q88, resulted in complete loss (more than a 500-fold reduction) of enzymatic activity, enabling

R71 and E88 to be identified as the potential catalytic residues. Simultaneous substitutions of T109A, E112Q, and Y114F also showed very low activity (2.8 mV·sec /ng of protein), equivalent to approximately 1/300 of wild-type LFS (880 mV·sec /ng protein), suggesting an important role of these three amino acids in the enzymatic activity.

Table 5 Amino acids substitution in onion LFS and LFS enzyme activity

position of subustituted amino acids in onion LFS[a)]	native amino acids[b)]	mutant amino acids[c)]	LFS enzyme activity
33-35	KPE	LPQ	+
42	K	L	+
43	D	N	+
4953	KVMPS	LVMPA	+
5557	SVC	AVS	+
58	E	Q	+
61	E	Q	+
69-74	CVRYVK	SVLYVL	-
69	C	S	+
71	R	L	-
74	K	L	+
78-83	HPIEEE	LPIQQQ	+
88-94	EKLVALD	QKLVALN	-
88	E	Q	-
94	D	N	+
109-114	TGYEDY	AGYQDF	+/-
122-127	EGPEHK	QGPQLL	+
130-134	RFDWS	LFDWA	+
132	D	N	+
137-142	CKYIE	SLYIQ	+
158	E	Q	+

[a)] number indicates the substituted amino acids position from N-terminus of full length onion LFS.

[b)] Red character indicates target amino acids.

[c)] deduced amino acids of site-directed mutant protein are shown.

Homology modeling of LFS/23–169

The BLASTP searches with SWISS-MODEL for LFS/23–169 enabled nine of the top ten proteins in the search results to be classified as plant pathogenesis-related proteins of class 10 (PR-10), and the top three matched proteins were regulatory components of the abscisic acid receptor protein that functions in early abscisic acid signaling. These top three proteins, identified as pyrabactin resistance 1 (PYR1) and PYR1-like proteins (PYLs), belong to PR-10, indicating that LFS could be classified in the structural family of PR-10.

Homology modeling was performed for LFS/23–169; its enzymatic activity was equivalent to that of LFS-3. A 3D model of LFS/23–169 was built on the basis of a template structure of PYL2 (PDB code: 3KDH) (YIN *et al.* 2009; KLINGLER *et al.* 2010). Since LFS shares low identity (18.8 %) with that of PYL2, we compared the secondary elements of LFS in model structures with those determined by NMR experiments. Despite the relatively low sequence conservation (ca.20% similarity), the secondary structure elements of the modeled structure showed good agreement with those determined by the NMR secondary shifts (Ohashi, W. *et al.,* in preparation), indicating that the model was acceptable and reliable.

The modeled structure of LFS/23–169 contained six anti-parallel β-strands and two α-helices. Although a short α-helix was present between α1 and β1 in the template structure, the corresponding region was somewhat disordered and could not form the typical helical structure in the LFS model. The PR-10 family uses a topologically conserved cavity for ligand recognition (KLINGLER *et al.* 2010) which is formed by concave β-sheet and α-helices. This cavity was also found in the LFS model structure. R71 and E88 were respectively located in β2 and β3 strand, and close to each other in the

cavity (Figure 14). Three important residues for the enzymatic activity (T109, E112, and Y114) were also located in the cavity, E112 and Y114 being located on β6 strand, and T109 being positioned between the β5 and β6 strands.

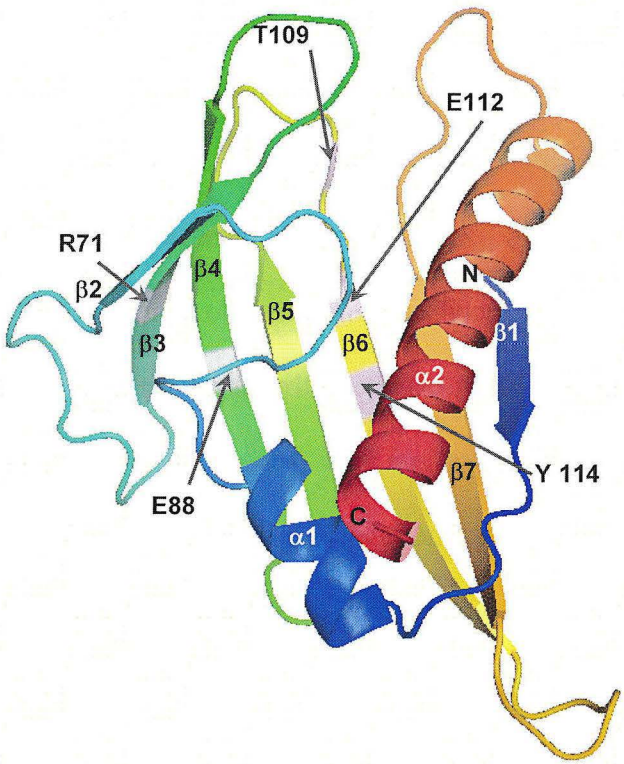


Figure 14 Three-dimensional model of LFS/23–169 of *Allium cepa*
Arg 71 and Glu 88 are shown in gray and Thr 109, Glu 112 and Tyr 114 are shown in pink.

DISCUSSION

Despite our initial expectation that some domains relevant to the enzymatic activity would be commonly conserved among LFSs from *Allium* plants, a comparison of their cDNA and amino acid sequences showed that all the LFSs studied were very similar to each another over their entire length; we were therefore unable to identify any consensus regions for the LFS enzymatic activity. The most conspicuous difference found among the sequences was that LFS of leek (*A. porrum*) lacked five amino acids (13–18) in the

N-terminus of the full-length LFS of onion (Figure 11). This suggests that this region was not critical for enzymatic activity and was consistent with those of deletion mutant experiments.

The results of deletion mutant experiments in this study suggest that both sequences from residues 23 to 30 relative to the N-terminus end and from residues 10 to 29 relative to the C-terminus end of onion LFS were essential for the structural stability of LFS, and as a result, were involved in the expression of LFS enzymatic activity., A similar but more striking observation has been reported in C-terminal deletion for bacterial chloramphenicol acetyltransferase (CAT), where deletion of only two C-terminal helix amino acid residues that were not involved in substrate binding, catalysis, or inter-subunit interaction provoked aggregation or destabilization of the native state, resulting in aggregation and/or nearly complete loss of enzymatic activity (ROBBEN *et al.* 1993). Added to this, a detailed analysis of the structure of ABA receptors showed that the level of helix-grip, which was important for binding to ABA on the ABA receptor, was formed by interactions among all motifs in the protein, particularly the N-terminal β-strand being close to the C-terminal helix in 3D structure (MELCHER *et al.* 2009; MIYAZONO *et al.* 2009; NISHIMURA *et al.* 2009). The N-terminal β-strand might therefore play an important role for stabilizing the level of helix-grip in LFS, deleting the motif inducing instability of the helix grip and resulting in inactivation of the enzyme.

Amino acid substitutions at R71 and E88 unambiguously showed these two amino acids to be essential for LFS activity. Consistent with our initial speculation, there is a possibility that glutamic acid (E88) interacted with the proton of the substrate. However, we cannot speculate on the role played by arginine (R71) in the catalytic reaction of LFS. Arginine and an amino acid bearing a carboxyl side chain, such as aspartic acid or

glutamic acid, when placed close to each other in an enzyme, can interact with each other and stabilize the charged form of each residue (GUTTERIDGE and THORNTON 2005). It is therefore possible that charge-dependent stability will result from coordination between R71 and E88 in onion LFS. Moreover, the charge(s) could be involved in the polarizing of 1-propenylsulfenic acid, a putative substrate of LFS. Negatively charged E88 might act as a neutrophile and catalyze the conversion of sulfenic acid to LF.

Although the functions of the residues for LFS activity remain unclear, the results from homology modeling indicate that these residues were located in a cavity functioning as a ligand-binding site in other PR-10 super family proteins, the structures of which have already been elucidated (LIU and EKRAMODDOULLAH 2006).

A survey of *Allium* LFSs in the GenBank showed that these sequences did not have high similarity with other sequences with defined functions, suggesting that they would constitute a novel group of proteins with a novel function. However, the BLASTP search and subsequent homology modeling showed LFS belonging to PR-10 superfamily, despite the low sequence similarity of LFS to PYL2. This protein superfamily consists of many proteins that have been purified from various flowering plants, the proteins in this superfamily exhibit a common structure but low sequence similarity (LIU and EKRAMODDOULLAH 2006; RADAUER *et al.* 2008). PR-10 proteins seem to be involved in plant defense against both pathogen attack and environmental stress. The speculation that *Allium* LFSs also play roles in the plant defense response to predators through LF formation implies that they perform functions similar to those of other PR-10 proteins in a broad sense.

Chapter IV. DETAILED ANALYSIS OF REACTION MECHANISM OF ONION LACHRYMATORY FACTOR SYNTHASE

INTRODUCTION

Although onions have significant nutritional and organoleptic values, it is not uncommon to experience a "crying" sensation when cutting and chopping onion bulbs. The substance that causes this irritation is called the lachrymatory factor (LF) and was identified as propanthial *S*-oxide by Brodnitz and Pascale more than 40 years ago(BRODNITZ and PASCALE 1971). The LF is a rare compound belonging to the group of thioaldehyde *S*-oxides. To date, only three other naturally occurring thioaldehyde *S*-oxides have been reported (BLOCK and BAYER 1990; KUBEC *et al.* 2003, 2010).

Onion LF has been widely believed to be non-enzymatically produced from (*E*)-1-propenylsulfenic acid, a putative reaction product of alliinase and (*E*)-(+)-*S*-(1-propenyl)- L-cysteine sulfoxide [(*E*)-PRENCSO)] that is most abundant in onion (BLOCK *et al.* 1979; BLOCK 1992; SHEN and PARKIN 2000; YAMANE *et al.* 1994). Contrary to this belief, we have discovered lachrymatory factor synthase (LFS), that is an essential enzyme for the formation of the LF after alliinase acts on (*E*)-PRENCSO and cloned from an onion bulb (*Allium cepa*) (IMAI *et al.* 2002).

Based on chemical analysis of freshly prepared onions crushed in D_2O it has been speculated that the LFS catalyzes the intra-molecular H^+ substitution reaction between atom 1 (oxygen) and atom 4 (carbon) in (*E*)-1-propenylsulfenic acid (Figure 15) (BLOCK *et al.* 1996). However, this crude work does not exclude the possibility that other unknown reactions may be involved in LF formation. Moreover, a detailed mechanism of this enzyme has not yet been published, which likely results from (*E*)-1-propenylsulfenic

acid instability.

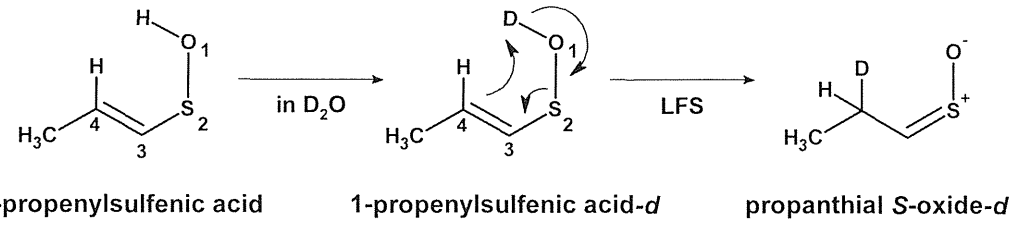


Figure 15 Speculated reaction pathway from *(E)*-1-propenyl sulfenic acid to LF.

In this study, we directly examined whether the LFS catalyzes the intra-molecular H^+ substitution reaction. We reacted purified garlic alliinase, onion LFS, and (*E*)-PRENCSO in D_2O. Substrate selectivity of the LFS for (*E*)- and (*Z*)-PRENCSO was also examined to provide corroborative evidence about the reaction mechanism.

MATERIALS AND METHODS

Materials

Garlic alliinase was purified according to the method described by Nock and Mazelis (1989). Recombinant onion LFS protein was prepared following published protocols (MASAMURA *et al.* 2012c). (*E*)-PRENCSO was purified from onion bulb by using the protocol described by Shen and Parkin (2000). After PRENCSO was synthesized according to the method of Lancaster and Kelly (1983), the (*E*)- and (*Z*)-forms were separated by reversed-phase HPLC (Pegasil ODS [20 mm i. d. × 250 mm], pH 3.3 H_2O (TFA), 5 ml/min, 230 nm).

LF formation in D_2O and GC-MS analysis

A small amount (2.5 mg) of (*E*)-PRENCSO was resolved with 1 ml D_2O (Sigma-Aldrich, 151882, 99.9 atom % deuterium) or deionized H_2O. Paper pads were blotted with garlic alliinase (7.59 units) and excess onion LFS and lyophilized. The resulting lyophilized paper pads were placed on the bottom of a 20 ml gas chromatography (GC) headspace vial and 90 μl of the PRENCSO solution was added. The vial was closed with a lid immediately and maintained at room temperature for approximately 5 min. The headspace of the vial was applied to the 5975 MSD MS system (Agilent Technologies) equipped with a 7890A GC system (Agilent Technologies). GC separation was achieved using a 60 m × 250 μm × 0.25 μm column (J&W 122-7062). The carrier gas was helium at a flow of 1.2 ml/min, and the column temperature program included an increase of 10°C/min from 40°C to 210°C. The detector temperature was held constant at 230°C and total ion chromatograms and mass spectra were analyzed with the electron-impact mode. Selected ion monitoring (SIM) was used to identify the incorporation of deuterium.

Substrate selectivity of onion LFS

(*Z*)-PRENCSO, which does not exist in onion, was used as a substrate instead of (*E*)-PRENCSO (a naturally existing form) in the *in vitro* model reaction system of LF formation described in Masamura *et al.* (2012c). The resulting reaction mixture was applied to HPLC to determine whether LF was produced from (*E*)-PRENCSO.

RESULTS

One hydrogen was exchanged to one deuterium in LF which was formed in D_2O

As shown in Figure 16, when LF formation proceeded in D_2O, an m/z peak of 91 was detected. This result strongly suggested that the exchangeable proton in (*E*)-1-propenylsulfenic acid was replaced with deuterium, resulting in the production of one deuterated LF (*m/z* = 91). Thus, LFS catalyzes the intra-molecular H^+ substitution reaction. In the D_2O reaction, deuterium-labeled propanal was also detected. Although the pathway of desulfuration is unclear, propanal is a known decomposition product of the LF (BLOCK 2010).

Onion LF was not produced from (Z)-PRENCSO

The amount of LF produced by each substrate [(*E*)-PRENCSO and (*Z*)-PRENCSO] was directly compared. Although alliinase has been reported to hydrolyze both PRENCSOs equally (FREEMAN and WHENHAM 1975), the degraded products of (*Z*)-PRENCSO by alliinase did not lead to LF formation (Figure 17). This result indicates that the LFS has a high selectivity to the substance with which it can react, and supports the findings of our deuterium incorporation experiment.

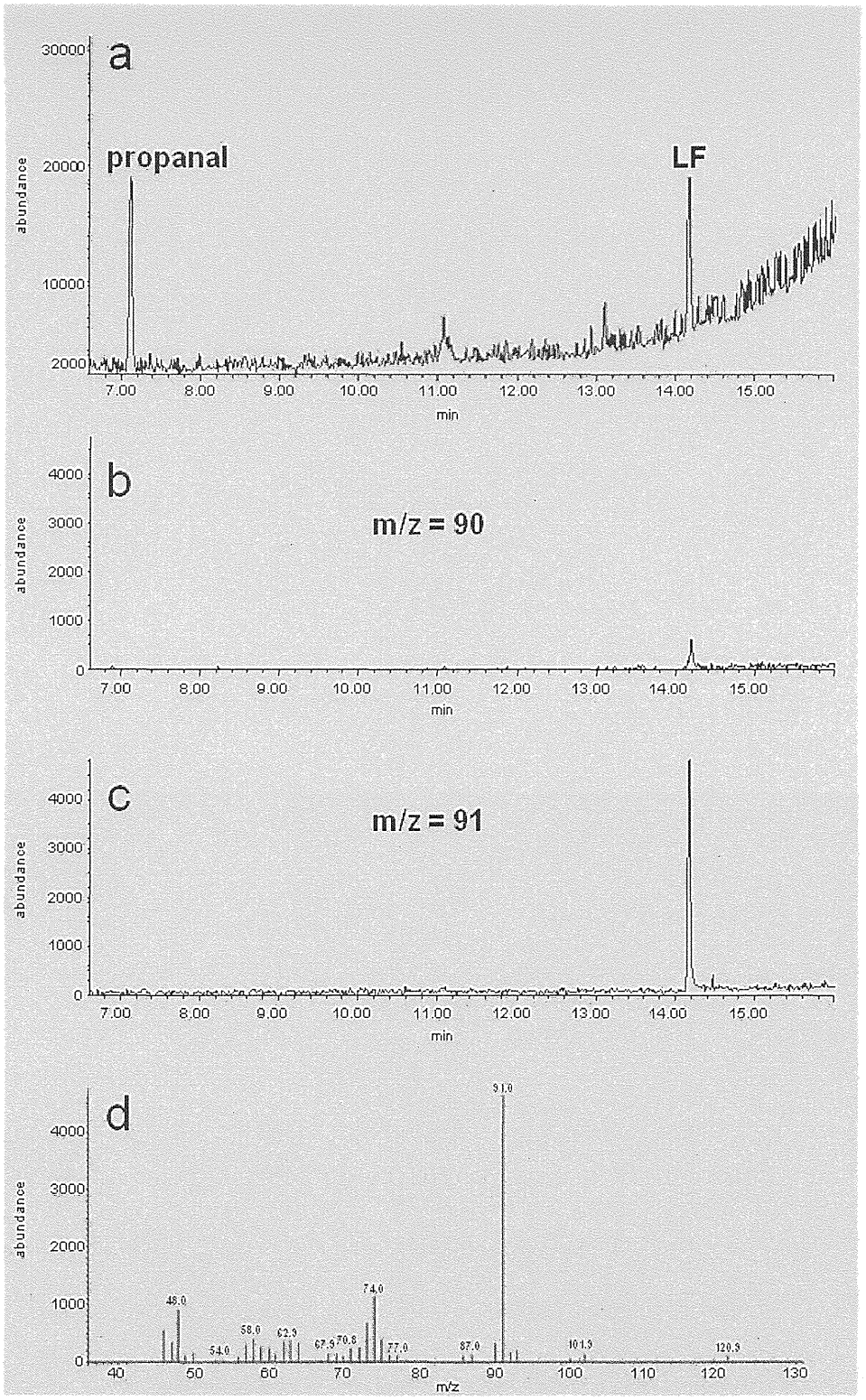


Figure 16 GS-MS analysis of head space of the reaction mix
Analysis of the reaction mix in D_2O by total ion chromatography (a) and selected ion monitoring (SIM) of *m/z* of 90 (b) and 91 (c). Mass spectra of the peak detected in the SIM of *m/z* of 91 (d).

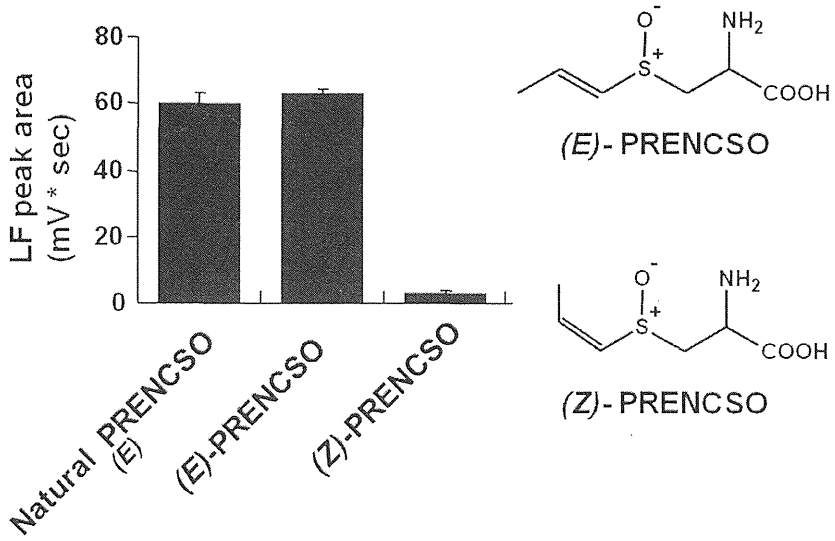


Figure 17 Substrate selectivity of onion LFS
LF-forming activity was indicated as LF peak area with according to Masamura *et al.* 2012

DISCUSSION

In this study, we showed that the reaction occurred following onion tissue disruption was reproduced by the model consisting only alliinase, LFS, and (*E*)-PRENCSO in physiological condition. Moreover, in (*Z*)-propenylsulfenic acid, the hydrogen connecting with atom 4 (carbon) is too far from the exchangeable proton connecting with atom 1 (oxygen) to facilitate the 1−4 rearrangement reaction. From these results, we propose that onion LFS should be classified as an (*E*)-1-propenylsulfenic acid isomerase.

In addition, a high degree of deuterium enrichment (87%) was observed. The deuterium-labeled LF may help further understand the mechanism of thiosulfinate formation following LF formation.

LFS cDNAs have also been cloned from five other lachrymatory *Allium* species (*A.*

ampeloprasum, *A. cepa Aggregatum group*, *A. chinese*, *A. fistulosum*, and *A. porrum*), and all resulting recombinant proteins showed LFS enzymatic activity (MASAMURA *et al.* 2012c). Strong sequence homology was observed in these LFS cDNAs (IMAI *et al.* 2008; MASAMURA *et al.* 2012c), suggesting that the reaction mechanism shown in this study is widely used among lachrymatory *Allium* species.

Chapter V. GENERAL DISCUSSION

Allium cepa L., which includes the bulb onion and shallot, is one of the most important crops cultivated worldwide. Although onions are appreciated for their nutritional and organoleptic values, they do cause tearing when they are cut or chopped. The tearing property conferred by LF is the most distinctive attribute of onion. LF formation from 1-propenylsulfenic acid is mediated by LFS (IMAI *et al.* 2002). The discovery of LFS has proven that the formation of the LF is under the control of a specific single gene and has led to the possibility of producing a tearless onion through approaches that involve GM, mutagenesis, and traditional breeding. Onion with low LFS activity might be low in pungency, tearless, and health enhancing.

In the present studies, to contribute to genetic mutagenesis and molecular breeding of LFS-suppressed onions, the genomic organization of the onion LFS gene and enzymatic properties of its protein product were thoroughly analyzed. LFS in *A. cepa* is transcribed from at least two loci, which are localized on chromosome 5. The onion LFS gene was classified broadly into two haplotypes according to three SNPs. Arg71 and Glu88 of the LFS protein were identified as amino acid residues that are indispensable to LFS enzymatic activity. We assigned the three-dimensional structure of LFS by homology

modeling. The model showed LFS belonging to the PR-10 superfamily. The reaction mechanism of LFS was isomerization from (*E*)-1-propenylsulfenic acid to LF. Onion LFS can react with the degraded products of (*E*)-PRENCSO by alliinase, but not with those of (*Z*)-PRENCSO.

This general discussion encompasses an advanced study of the production of LFS-null onions and a detailed explanation of LFS. It is based on the findings in this study regarding the possibility of the production of LFS-null onion from the perspective of the copy number of the LFS gene, gene duplication in the onion genome, the importance of chromosome 5 according to the results of onion breeding, the conservation of LFS in *Allium*, and the catalytic mechanism of LFS.

(1) Possibility of mutagenesis and breeding approaches for generating an LFS-null plant

Estimation of the copy number and distribution of target gene loci in the genome is important to assess the possibility of success of mutagenesis screening for obtaining a plant possessing mutation(s) in the gene. As reported in Chapter II, the onion LFS gene is transcribed from at least two loci, and they are localized narrowly on chromosome 5. This finding suggests that there is a low possibility for obtaining null mutants via mutant screening. However, two haplotypes showed a large difference in the upper flanking 1.2 kbp (corresponding to the promoter region) of LFS. These observations of variation in the promoter regions suggest that differential expression may exist between the paralogs. Therefore, first, we should compare the expression level between the paralogs. In parallel with the analysis of differential expression levels, a high-density linkage map around the region where LFS loci were assigned will be required to select the LFS-null plant.

Another strategy for obtaining low-LFS onion is deep sequencing of the key functional region of the LFS-coding region corresponding to Arg71–Glu88 in diverse onion germplasm and mutation-induced population.

Moreover, as reported in Chapter IV, because LF does not form from (*Z*)-PRENCSO, the screening of the onion variety producing and storing (*Z*)-PRENCSO could be an alternative approach to produce a low-pungency, tearless, and health-enhancing onion.

(2) **Gene duplication in the onion genome**

The nuclear genome of the onion is characterized by its great size, 17 pg or 15 Gbp per 1 C (ARUMUGANATHAN and EARLE 1991; BENNETT and LEITCH 1995; RICROCH *et al.* 2005), one of the largest among cultivated plants. Along with a modest chromosome number (2n = 2x = 16), onion is frequently used for cytological research and education. The enormous DNA content and low chromosome number of onion suggests that genomic evolution in onion has occurred as a result of intrachromosomal duplication. Mapping studies have suggested extensive duplication in onion. King *et al.* (1998) reported that 21% of the 91 cDNA probes of RFLP detected more than one segregating RFLP, of which 53% were unlinked and 47% were linked. These findings indicate that gene duplication has occurred with high frequency in the onion genome, and the ratio of linked duplicated RFLP was higher than that in other plants. As reported in Chapter II, LFS was mapped on a marker linkage map. At least two LFS loci were located narrowly on the chromosome. This result will open the possibility for marker-assisted selection of a tearless onion. In addition, The findings in Chapter II suggests that as the onion linkage map becomes more precise, it will show that many genes exist as duplicated genes within a narrow region of a single chromosome.

(3) The importance of chromosome 5 in the context of onion breeding

Bulbing in leaf sheaths is observed only in a monosomic addition-carrying chromosome 5A of the shallot (FF+5A), but not in other monosomic addition lines. (SHIGYO *et al.* 1996; MASUZAKI *et al.* 2007).

A QTL affecting bulb dry matter and a putative sucrose transporter gene are located on chromosome 5 (GALMARINI *et al.* 2001; MARTIN *et al.* 2005; MASUZAKI *et al.* 2007). These findings suggested that carbohydrate composition and transportation would affect bulb formation and that the key gene(s) would be located in onion chromosome 5. As described in Chapter II, this study revealed that an essential gene for another important property of onion, production of LF, was also located on chromosome 5. These observations suggest that linkage drag causes some effects on LFS. Consequently, this region of chromosome 5 could be an important domestication region in the onion genome, and it could be valuable to pursue further functional and crop evolutionary studies.

Addition to this, a distinct characteristic of ACSO biosynthesis in FF+5A was observed. In FF+5A, ACSOs and amino acid (Ser and Cys) accumulation levels remained high throughout the year, but no candidate genes involved in sulfur assimilation and ACSO synthesis were detected. Microarray analysis revealed that APSR gene expression was inhibited by 5A chromosome additions (MASAMURA *et al.* 2011). These results suggested that anonymous regulatory genes related to the up-regulation of the expression of *A. fistulosum* candidate genes in the ACSO biosynthesis pathway were located on chromosome 5.

(4) LFS as a new protein family

As reported in chapter III, the seven LFS cDNA sequences of five alliaceous species—shallot (*A. ascalonicum*: AB094593), Japanese bunching onion (*A. fistulosum*: AB094590, AB094591), rakkyo (*A. chinense*: AB094592), leek (*A. porrum*: AB094594, AB04595), and elephant garlic (*A. ampeloprasum*: AB09496)—were cloned, and their recombinant proteins were expressed in *Escherichia coli*. All the recombinant products showed LF-forming activity. In addition, strong homologies were observed in the LFS cDNA ORF sequences, including in the cDNA from bulb onion (*A. cepa*: AB089203) (IMAI *et al.* 2002). These results suggest that the LFS gene is highly conserved and primarily responsible for producing lachrymatory property among lachrymatory alliaceous plants. Moreover, GenBank searches did not yield any sequences homologous to these LFS cDNAs, indicating that they constitute a novel group of proteins. The distinct location of LFS only in lachrymatory alliaceous plants, as well as the highly conserved sequence of LFSs, suggests that the LFS gene family only developed in this group of plant species. Therefore, this gene family could be a suitable model to investigate the phylogenetic relationships among this group of plants.

(5) The catalytic reaction of LFS

Onion LF belongs to the group of thioaldehyde *S*-oxides in which to date only three other naturally occurring thioaldehyde *S*-oxides have been reported. Because onion LFS was the first reported enzyme involved in the production of thioaldehyde *S*-oxides, it is a novel protein also from the terms of its reaction product. As shown in chapter IV, GC-MS analysis of a model reaction mix in D_2O indicated that onion LFS catalyzed an intra-molecular H^+ substitution reaction to produce LF from (*E*)-1-propenylsulfenic acid.

From the results of chapter III, Arg71 and Glu88 were identified as the amino acids essential for LFS enzyme activity. Negatively charged Glu88 might act as neutrophiles and catalyze the conversion of sulfenic acid to LF. However, I cannot speculate on the role played by Arg71 in the catalytic reaction of LFS. It is possible that charge-dependent stability will occur as a result of coordination between Arg71 and Glu88 in onion LFS. Charge(s) could be involved in the polarization of 1-propenylsulfenic acid.

Although functions of the residues for LFS activity remain unclear, the results of homology modeling indicate that these residues are located in a cavity that functions as a ligand-binding site in other PR-10 superfamily proteins for which the structures have been elucidated (LIU and EKRAMODDOULLAH 2006).

LITTERATURE CITED

ALTSCHUL S. F., MADDEN T. L., SCHÄFFER A. A., ZHANG J., ZHANG Z., MILLER W., LIPMAN D. J., 1997 Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25**: 3389–3402.

AOYAGI M., KAMOI T., KATO M., SASAKO H., TSUGE N., IMAI S., 2011 Structure and Bioactivity of Thiosulfinates Resulting from Suppression of Lachrymatory Factor Synthase in Onion. Journal of Agricultural and Food Chemistry **59**: 10893–10900.

ARNOLD K., BORDOLI L., KOPP J., SCHWEDE T., 2006 The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics **22**: 195–201.

ARUMUGANATHAN K., EARLE E., 1991 Nuclear DNA content of some important plant

species. Plant Molecular Biology Reporter **9**: 208–218.

BENNETT M. D., LEITCH I. J., 1995 Nuclear DNA Amounts in Angiosperms. Annals of Botany **76**: 113 –176.

BLOCK E., 1992 The Organosulfur Chemistry of the GenusAllium - Implications for the Organic Chemistry of Sulfur. Angewandte Chemie International Edition in English **31**: 1135–1178.

BLOCK E., 2010 *Garlic and other alliums: the lore and the science.* RSC Pub., Cambridge, UK.

BLOCK E., BAYER T., 1990 (Z,Z)-d,l-2,3-Dimethyl-1,4-butanedithial S,S'-dioxide: a novel biologically active organosulfur compound from onion. Formation of vic-disulfoxides in onion extracts. Journal of the American Chemical Society **112**: 4584–4585.

BLOCK E., GILLIES J. Z., GILLIES C. W., BAZZI A. A., PUTMAN D., REVELLE L. K., WANG D., ZHANG X., 1996 Allium Chemistry: Microwave Spectroscopic Identification, Mechanism of Formation, Synthesis, and Reactions of (E, Z)-Propanethial S-Oxide, the Lachrymatory Factor of the Onion (Allium cepa). Journal of the American Chemical Society **118**: 7492–7501.

BLOCK E., PENN R. E., REVELLE L. K., 1979 Flash vacuum pyrolysis studies. 7. Structure and origin of the onion lachrymatory factor. A microwave study. Journal of the American Chemical Society **101**: 2200–2201.

BRADFORD M. M., 1976 A Rapid and Sensitive Method for the Quantitation of

Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. Anal. Biochem. **72**: 248–254.

BRAUN D. M., SLEWINSKI T. L., 2009 Genetic control of carbon partitioning in grasses: roles of sucrose transporters and tie-dyed loci in phloem loading. Plant Physiol. **149**: 71–81.

BRODNITZ M. H., PASCALE J. V., 1971 Thiopropanal S-oxide: a lachrymatory factor in onions. Journal of Agricultural and Food Chemistry **19**: 269–272.

BRYSON K., MCGUFFIN L. J., MARSDEN R. L., WARD J. J., SODHI J. S., JONES D. T., 2005 Protein structure prediction servers at University College London. Nucleic Acids Res. **33**: W36–38.

DRUMMOND A. J., ASHTON B., BUXTON S., CHEUNG M., COOPER A., 2010 *Geneious v5.3, Available from http://www.geneious.com/4 January 2011, date last accessed.*

EADY C. C., KAMOI T., KATO M., PORTER N. G., DAVIS S., SHAW M., KAMOI A., IMAI S., 2008 Silencing onion lachrymatory factor synthase causes a significant change in the sulfur secondary metabolite profile. Plant Physiol. **147**: 2096–2106.

FAOSTAT, 2007

FREEMAN G. G., WHENHAM R. J., 1975 The use of synthetic (±)-s-1-propyl-L-cysteine sulphoxide and of alliinase preparations in studies of flavour changes resulting from processing of onion (allium cepa l.). Journal of the Science of Food and Agriculture **26**: 1333–1346.

FRITSCH R. M., FRIESEN N., 2002 Evolution, domestication and taxonomy. In: Rabinowitch HD, Currah L (Eds.), *Allium crop science recent advances*, CABI Pub., Wallingford, Oxon; New York, NY, pp. 5–30.

GALMARINI C. R., GOLDMAN I. L., HAVEY M. J., 2001 Genetic analyses of correlated solids, flavor, and health-enhancing traits in onion (Allium cepa L.). Mol. Genet. Genomics **265**: 543–551.

GRIFFITHS G., TRUEMAN L., CROWTHER T., THOMAS B., SMITH B., 2002 Onions—A global benefit to health. Phytotherapy Research **16**: 603–615.

GUEX N., PEITSCH M. C., 1997 SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis **18**: 2714–2723.

GUINDON S., GASCUEL O., 2003 A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. **52**: 696–704.

GUTTERIDGE A., THORNTON J. M., 2005 Understanding nature's catalytic toolkit. Trends Biochem. Sci. **30**: 622–629.

GUY L., KULTIMA J. R., ANDERSSON S. G. E., 2010 genoPlotR: comparative gene and genome visualization in R. Bioinformatics **26**: 2334–2335.

HANELT P., 1990 Taxonomy, evolution, and history. In: Brewster JL, Rabinowitch HD (Eds.), *Onions and allied crops*, CRC Pr., Boca Raton, Fla.

HAVEY M. J., 2002 Genome organization in Allium. In: Rabinowitch HD, Currah L (Eds.), *Allium crop science recent advances*, CABI Pub., Wallingford, Oxon;

New York, NY, pp. 59–79.

HEUSDEN A. W. VAN, OOIJEN J. W. VAN, VRIELINK-VAN GINKEL R., VERBEEK W. H. J., WIETSMA W. A., KIK C., 2000a A genetic map of an interspecific cross in Allium; based on amplified fragment length polymorphism (AFLPTM;) markers. TAG Theoretical and Applied Genetics **100**: 118–126.

HEUSDEN A. W. VAN, SHIGYO M., TASHIRO Y., VRIELINK-VAN GINKEL R., KIK C., 2000b AFLP linkage group assignment to the chromosomes of Allium cepa L. via monosomic addition lines. TAG Theoretical and Applied Genetics **100**: 480–486.

IMAI S., MASAMURA N., TSUGE N., TOMOTAKE M., SAWADA H., NAGATA T., KUMAGAI H., 2008 Lachrymatory Factor Synthase (LFS), a Long Overlooked Enzyme in Onion. J. Clin. Biochem. Nutr **43**: 211–214.

IMAI S., TSUGE N., TOMOTAKE M., MASAMURA N., SAWADA H., 2005 *The production of a lachrymatory factor in onion is mediated by a novel enzyme in Sulfur transport and assimilation in plants in the post genomic era*. Backhuys, Leiden.

IMAI S., TSUGE N., TOMOTAKE M., NAGATOME Y., SAWADA H., NAGATA T., KUMAGAI H., 2002 Plant biochemistry: An onion enzyme that makes the eyes water. Nature **419**: 685–685.

JAIN M., KHURANA P., TYAGI A., KHURANA J., 2008 Genome-wide analysis of intronless genes in rice and Arabidopsis. Functional & Integrative Genomics **8**: 69–78.

JAKŠE J., MEYER J., SUZUKI G., MCCALLUM J., CHEUNG F., TOWN C., HAVEY M., 2008

Pilot sequencing of onion genomic DNA reveals fragments of transposable elements, low gene densities, and significant gene enrichment after methyl filtration. Molecular Genetics and Genomics **280**: 287–292.

JIANG J., GILL B. S., WANG G. L., RONALD P. C., WARD D. C., 1995 Metaphase and interphase fluorescence in situ hybridization mapping of the rice genome with bacterial artificial chromosomes. Proceedings of the National Academy of Sciences **92**: 4487 –4491.

JONES D. T., 1999 Protein secondary structure prediction based on position-specific scoring matrices. J. Mol. Biol. **292**: 195–202.

JONG H. J. DE, FRANSZ P., ZABEL P., 1999 High resolution FISH in plants – techniques and applications. Trends in Plant Science **4**: 258–263.

KAMOI T., EADY C. C., IMAI S., 2008 Gene silencing assessment for genes from recalcitrant or poorly studied plant species. Plant Biotechnology Reports **2**: 199–206.

KHRUSTALEVA L. I., KIK C., 2001 Localization of single-copy T-DNA insertion in transgenic shallots (Allium cepa) by using ultra-sensitive FISH with tyramide signal amplification. The Plant Journal **25**: 699–707.

KHRUSTALEVA L. I., MELO P. E. DE, HEUSDEN A. W. VAN, KIK C., 2005 The Integration of Recombination and Physical Maps in a Large-Genome Monocot Using Haploid Genome Analysis in a Trihybrid Allium Population. Genetics **169**: 1673 –1685.

KIM S., YOO K.-S., PIKE L. M., 2005 Development of a codominant PCR-based marker for allelic selection of the pink trait in onions (Allium cepa), based on the insertion mutation in the promoter of the anthocyanidin synthase gene. TAG Theoretical and Applied Genetics **110**: 573–578.

KING J. J., BRADEEN J. M., BARK O., MCCALLUM J. A., HAVEY M. J., 1998 A low-density genetic map of onion reveals a role for tandem duplication in the evolution of an extremely large diploid genome. TAG Theoretical and Applied Genetics **96**: 52–62.

KLINGLER J. P., BATELLI G., ZHU J.-K., 2010 ABA receptors: the START of a new paradigm in phytohormone signalling. J. Exp. Bot. **61**: 3199–3210.

KOORNNEEF M., FRANSZ P., JONG H. DE, 2003 Cytogenetic tools for;Arabidopsis thaliana. Chromosome Research **11**: 183–194.

KUBEC R., CODY R. B., DANE A. J., MUSAH R. A., SCHRAML J., VATTEKKATTE A., BLOCK E., 2010 Applications of direct analysis in real time-mass spectrometry (DART-MS) in Allium chemistry. (Z)-butanethial S-oxide and 1-butenyl thiosulfinates and their S-(E)-1-butenylcysteine S-oxide precursor from Allium siculum. J. Agric. Food Chem. **58**: 1121–1128.

KUBEC R., KIM S., MUSAH R. A., 2003 The lachrymatory principle of Petiveria alliacea. Phytochemistry **63**: 37–40.

KUHL J. C., CHEUNG F., YUAN Q., MARTIN W., ZEWDIE Y., MCCALLUM J., CATANACH A., RUTHERFORD P., SINK K. C., JENDEREK M., PRINCE J. P., TOWN C. D., HAVEY

M. J., 2004 A Unique Set of 11,008 Onion Expressed Sequence Tags Reveals Expressed Sequence and Genomic Differences between the Monocot Orders Asparagales and Poales. The Plant Cell Online **16**: 114 –125.

LAMB J. C., DANILOVA T., BAUER M. J., MEYER J. M., HOLLAND J. J., JENSEN M. D., BIRCHLER J. A., 2007 Single-Gene Detection and Karyotyping Using Small-Target Fluorescence in Situ Hybridization on Maize Somatic Chromosomes. Genetics **175**: 1047 –1058.

LANCASTER J. E., KELLY K. E., 1983 Quantitative analysis of the S-alk(en)yl-L-cysteine sulphoxides in onion (Allium cepa L.). Journal of the Science of Food and Agriculture **34**: 1229–1235.

LIBRADO P., ROZAS J., 2009 DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics **25**: 1451 –1452.

LIU J.-J., EKRAMODDOULLAH A. K. M., 2006 The family 10 of plant pathogenesis-related proteins: Their structure, regulation, and function in response to biotic and abiotic stresses. Physiological and Molecular Plant Pathology **68**: 3–13.

LOMBARD K. A., GEOFFRIAU E., PEFFLEY E. B., 2004 Total Quercetin Content in Onion: Survey of Cultivars Grown at Various Locations. HortTechnology **14**: 628 –630.

MARAIS D. L. DES, RAUSHER M. D., 2008 Escape from adaptive conflict after duplication in an anthocyanin pathway gene. Nature **454**: 762–765.

MARTIN W., MCCALLUM J., SHIGYO M., JAKSE J., KUHL J., YAMANE N., PITHER-JOYCE M., GOKCE A., SINK K., TOWN C., HAVEY M., 2005 Genetic mapping of expressed sequences in onion and in silico comparisons with rice show scant colinearity. Molecular Genetics and Genomics **274**: 197–204.

MASAMURA N., AOYAGI M., TSUGE N., KAMOI T., IMAI S., 2012a Proton Transfer in a Reaction Catalyzed by Onion Lachrymatory Factor Synthase. Bioscience, Biotechnology, and Biochemistry **76**: 1799–1801.

MASAMURA N., MCCALLUM J., KHRUSTALEVA L., KENEL F., PITHER-JOYCE M., SHONO J., SUZUKI G., MUKAI Y., YAMAUCHI, N., SHIGYO M., 2012b Chromosomal Organization and Sequence Diversity of Genes Encoding Lachrymatory Factor Synthase in Allium cepa L. Genes|Genomes|Genetics **2**: 643–651.

MASAMURA N., OHASHI W., TSUGE N., IMAI S., ISHII-NAKAMURA A., HIROTA H., NAGATA T., KUMAGAI H., 2012c Identification of Amino Acid Residues Essential for Onion Lachrymatory Factor Synthase Activity. Bioscience, Biotechnology, and Biochemistry **76**: 447–453.

MASAMURA N., YAGUCHI S., ONO Y., NAKAJIMA T., MASUZAKI S., IMAI S., YAMAUCHI N., SHIGYO M., 2011 Characterization of Amino Acid and S-alk(en)yl-L-cysteine Sulfoxide Production in Japanese Bunching Onion Carrying an Extra Chromosome of Shallot. Journal of the Japanese Society for Horticultural Science **80**: 322–333.

MASUZAKI S., SHIGYO M., YAMAUCHI N., 2006a Direct comparison between genomic constitution and flavonoid contents in Allium multiple alien addition lines reveals chromosomal locations of genes related to biosynthesis from dihydrokaempferol

to quercetin glucosides in scaly leaf of shallot (Allium cepa L.). TAG Theoretical and Applied Genetics **112**: 607–617.

MASUZAKI S., SHIGYO M., YAMAUCHI N., 2006b Complete assignment of structural genes involved in flavonoid biosynthesis influencing bulb color to individual chromosomes of shallot (Allium cepa L.). Genes & Genetic Systems **81**: 255–263.

MASUZAKI S., YAGUCHI S., YAMAUCHI N., SHIGYO M., 2007 Morphological characterisation of multiple alien addition lines of Allium reveals the chromosomal locations of gene(s) related to bulb formation in Allium cepa L. The Journal of Horticultural Science & Biotechnology **82**: 393–396.

MCCALLUM J., 2007 onion. In: Kole C (Ed.), *Genome Mapping and Molucular Breeding inPlants*, Vegetables. Springer, Berlin; New York, pp. 331–347.

MCCALLUM J., CLARKE A., PITHER-JOYCE M., SHAW M., BUTLER R., BRASH D., SCHEFFER J., SIMS I., HEUSDEN S. VAN, SHIGYO M., HAVEY M., 2006 Genetic mapping of a major gene affecting onion bulb fructan content. TAG Theoretical and Applied Genetics **112**: 958–967.

MCCALLUM J., PITHER-JOYCE M., SHAW M., KENEL F., DAVIS S., BUTLER R., SCHEFFER J., JAKSE J., HAVEY M., 2007 Genetic mapping of sulfur assimilation genes reveals a QTL for onion bulb pungency. TAG Theoretical and Applied Genetics **114**: 815–822.

MCCALLUM J., PITHER-JOYCE M., SHAW M., PORTER N. G., SEARLE B., MCMANUS M. T., HAVEY M. J., 2005 Genomic and biochemical analysis of onion (Allium

cepa L.) sulfur metabolism. Acta Hort. **688**: 75–84.

MCCALLUM J., THOMSON S., PITHER-JOYCE M., KENEL F., CLARKE A., HAVEY M. J., 2008 Genetic Diversity Analysis and Single-nucleotide Polymorphism Marker Development in Cultivated Bulb Onion Based on Expressed Sequence Tag–Simple Sequence Repeat Markers. Journal of the American Society for Horticultural Science **133**: 810 –818.

MEER Q. VAN DER, HANELT P., 1990 *Onion and Allied Crops volume III Biochemistry, Food Science and Minor Crops*. CRC Press Inc. Boca Raton, FL, USA.

MELCHER K., NG L.-M., ZHOU X. E., SOON F.-F., XU Y., SUINO-POWELL K. M., PARK S.-Y., WEINER J. J., FUJII H., CHINNUSAMY V., KOVACH A., LI J., WANG Y., LI J., PETERSON F. C., JENSEN D. R., YONG E.-L., VOLKMAN B. F., CUTLER S. R., ZHU J.-K., XU H. E., 2009 A gate-latch-lock mechanism for hormone signalling by abscisic acid receptors. Nature **462**: 602–608.

MIYAZONO K.-I., MIYAKAWA T., SAWANO Y., KUBOTA K., KANG H.-J., ASANO A., MIYAUCHI Y., TAKAHASHI M., ZHI Y., FUJITA Y., YOSHIDA T., KODAIRA K.-S., YAMAGUCHI-SHINOZAKI K., TANOKURA M., 2009 Structural basis of abscisic acid signalling. Nature **462**: 609–614.

MOGREN L. M., OLSSON M. E., GERTSSON U. E., 2006 Quercetin Content in Field-Cured Onions (Allium cepa L.) Effects of Cultivar, Lifting Time, and Nitrogen Fertilizer Level. Journal of Agricultural and Food Chemistry **54**: 6185–6191.

MOGREN L. M., OLSSON M. E., GERTSSON U. E., 2007 Quercetin content in stored onions (Allium cepa L.) effects of storage conditions, cultivar, lifting time and nitrogen fertiliser level. Journal of the Science of Food and Agriculture **87**: 1595–1602.

MUSAH R. A., HE Q., KUBEC R., 2009 Discovery and Characterization of a Novel Lachrymatory Factor Synthase in Petiveria alliacea and Its Influence on Alliinase-Mediated Formation of Biologically Active Organosulfur Compounds. PLANT PHYSIOLOGY **151**: 1294–1303.

NISHIMURA N., HITOMI K., ARVAI A. S., RAMBO R. P., HITOMI C., CUTLER S. R., SCHROEDER J. I., GETZOFF E. D., 2009 Structural mechanism of abscisic acid binding and signaling by dimeric PYR1. Science **326**: 1373–1379.

NOCK L. P., MAZELIS M., 1989 Lack of homology between the alliin lyases of garlic and onion. Phytochemistry **28**: 729–731.

OKAMOTO D., NOGUCHI Y., MURO T., MORISHITA M., 2006 Genetic Variation of Quercetin Glucoside Content in Onion (Allium cepa L.). Journal of the Japanese Society for Horticultural Science **75**: 100–108.

OOIJEN J. W. VAN, JoinMap ® 4, Software for the calculation of genetic linkage maps in experimental populations. Kyazma B.V., Wageningen,: Netherlands. **2006**.

PEARCE S., PICH U., HARRISON G., FLAVELL A., HESLOP-HARRISON J., SCHUBERT I., KUMAR A., 1996 The Ty1-copia group retrotransposons of Allium cepa are distributed throughout the chromosomes but are enriched in the terminal

heterochromatin. Chromosome Research **4**: 357–364.

PIJNACKER L. P., FERWERDA M. A., 1984 Giemsa C-banding of potato chromosomes. Can. J. Genet. Cytol. **26**: 415–419.

RADAUER C., LACKNER P., BREITENEDER H., 2008 The Bet v 1 fold: an ancient, versatile scaffold for binding of large, hydrophobic ligands. BMC Evol. Biol. **8**: 286.

RANDLE W. M., 1997 Onion Flavor Chemistry and Factors Infuencing Flavor Intensity. In: Risch SJ, American Chemical Society. Division of Agricultural and Food Chemistry (Eds.), *Spices: Flavor Chemistry and Antioxidant Properties*, American Chemical Society, Washington, DC.

RANDLE W. M., LANCASTER J. E., SHAW M. L., SUTTON K. H., HAY R. L., BUSSARD M. L., 1995 Quantifying Onion Flavor Compounds Responding to Sulfur Fertility-Sulfur Increases Levels of Alk(en)yl Cysteine Sulfoxides and Biosynthetic Intermediates. Journal of the American Society for Horticultural Science **120**: 1075 –1081.

RICE P., LONGDEN I., BLEASBY A., 2000 EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet. **16**: 276–277.

RICROCH A., YOCKTENG R., BROWN S. C., NADOT S., 2005 Evolution of genome size across some cultivated Allium species. Genome **48**: 511–520.

ROBBEN J., SCHUEREN J. VAN DER, VOLCKAERT G., 1993 Carboxyl terminus is essential for intracellular folding of chloramphenicol acetyltransferase. J. Biol.

Chem. **268**: 24555–24558.

ROGERS S. O., BENDICH A. J., 1988 *Extraction of DNA from plant tissues, pp. 1–10 in Plant Molecular Biology Manual A6 edited by S. B. Gelvin and R. A. Schilperoort. Kluwer Academic Publishers, Dordrecht, The Netherlands*. Kluwer Academic Publishers.

ROSE P., WHITEMAN M., MOORE P. K., ZHU Y. Z., 2005 Bioactive S-alk(en)yl cysteine sulfoxide metabolites in the genus Allium: the chemistry of potential therapeutic agents. Natural Product Reports **22**: 351.

SCHWEDE T., KOPP J., GUEX N., PEITSCH M. C., 2003 SWISS-MODEL: An automated protein homology-modeling server. Nucleic Acids Res. **31**: 3381–3385.

SHEN C., PARKIN K. L., 2000 In Vitro Biogeneration of Pure Thiosulfinates and Propanethial-S-oxide. Journal of Agricultural and Food Chemistry **48**: 6254–6260.

SHIGYO M., TASHIRO Y., ISSHIKI S., MIYAZAKI S., 1996 Establishment of a series of alien monosomic addition lines of Japanese bunching onion (Allium fistulosum L.) with extra chromosomes from shallot (A. cepa L. Aggregatum group). Genes & Genetic Systems **71**: 363–371.

STACK S. M., COMINGS D. E., 1979 The chromosomes and DNA of Allium cepa. Chromosoma **70**: 161–181.

STADEN R., 1996 The staden sequence analysis package. Molecular Biotechnology **5**:

233–241.

STEARN W. T., 1992 HOW MANY SPECIES OF ALLIUM ARE KNOWN? Curtis's Botanical Magazine **9**: 180–182.

SUZUKI G., DO G. S., MUKAI Y., 2002 Efficient Storage and Screening System for Onion BAC Clones. Breeding Science **52**: 157–159.

SUZUKI G., URA A., SAITO N., DO G. S., SEO B. B., YAMAMOTO M., MUKAI Y., 2001 BAC FISH analysis in Allium cepa. Genes & Genetic Systems **76**: 251–255.

SZINAY D., CHANG S.-B., KHRUSTALEVA L., PETERS S., SCHIJLEN E., BAI Y., STIEKEMA W. J., HAM R. C. H. J. VAN, JONG H. DE, KLEIN LANKHORST R. M., 2008 High-resolution chromosome mapping of BACs using multi-colour FISH and pooled-BAC FISH as a backbone for sequencing tomato chromosome 6. The Plant Journal **56**: 627–637.

TASHIRO Y., MIYAZAKI S., KANAZAWA K., 1982 On the shallot cultivated in the countries of southern Asia. Bull. Fac. Agr. **53**: 65–73.

YAGUCHI S., 2008 Biochemical and Genetic Studies on Functional Components in Allium Alien Addition Lines.

YAGUCHI S., MCCALLUM J., SHAW M., PITHER-JOYCE M., ONODERA S., SHIOMI N., YAMAUCHI N., SHIGYO M., 2008 Biochemical and Genetic Analysis of Carbohydrate Accumulation in Allium cepa L. Plant and Cell Physiology **49**: 730–739.

YAMANE A., YAMANE A., SHIBAMOTO T., 1994 Propanethial S-Oxide Content in Scallions (Allium fistulosum L. Variety Caespitosum) as a Possible Marker for Freshness during Cold Storage. Journal of Agricultural and Food Chemistry **42**: 1010–1012.

YIN P., FAN H., HAO Q., YUAN X., WU D., PANG Y., YAN C., LI W., WANG J., YAN N., 2009 Structural insights into the mechanism of abscisic acid signaling by PYL proteins. Nat. Struct. Mol. Biol. **16**: 1230–1236.

ZWICK M. S., HANSON R. E., ISLAM-FARIDI M. N., STELLY D. M., WING R. A., PRICE H. J., MCKNIGHT T. D., 1997 A rapid procedure for the isolation of Cot-1 DNA from plants. Genome **40**: 138–142.

SUMMARY

Molecular Biology and Biochemical Study of

Lachrymatory Factor Synthase in Allium cepa L. (onion)

Allium cepa L., which includes bulb onion (*A. cepa* L. Common onion group) and shallot (*A. cepa* L. Aggregatum group, 2n = 2x = 16, genome AA), is an essential species cultivated in many parts of the world. Bulb onion is suitable for cultivation in high- and mid-latitude climates, whereas shallot is cultivated best at low latitudes. The most distinctive attribute of onion is the tearing property conferred by its production of lachrymatory factor (LF; propanethial *S*-oxide). LF is a rare compound in nature, formed from 1-propenylsulfenic acid by lachrymatory factor synthase (LFS). A survey of *Allium* LFSs in the GenBank showed that these sequences did not have high similarity with other sequences with defined functions, suggesting that they would constitute a novel group of proteins with a novel catalytic function. Additionally, discovery of LFS has proven that the formation of LF is under the control of a specific single gene, which suggests a possibility of producing an onion that lacks the tearing property (tearless onion) via genetic modification (GM), mutagenesis, and traditional breeding. Thus, LFS is an important target for molecular breeding in onion. The present studies were conducted to reveal the genomic organization of the LFS gene and biochemical characterization of the LFS protein.

1) **Chromosomal organization and sequence diversity of onion lachrymatory factor synthase (LFS)**

LFS catalyzes the formation of lachrymatory factor (LF), one of the most distinctive

traits of bulb onion (*Allium cepa* L.). Therefore, we used LFS as a model for a functional gene in a huge genome, and examined the chromosomal organization of LFS in *A. cepa* by using multiple approaches. The first-level analysis completed the chromosomal assignment of the LFS gene to chromosome 5 of *A. cepa* by using a complete set of *A. fistulosum*-shallot (*A. cepa* L. Aggregatum group) monosomic addition lines. Subsequent use of an F_2 mapping population from the interspecific cross *A. cepa* × *A. roylei* confirmed the assignment of an LFS locus to this chromosome. Sequence comparison of two BAC clones bearing LFS genes, LFS amplicons from diverse germplasm, and expressed sequences from a doubled haploid line revealed variations that were consistent with duplicated LFS genes. Furthermore, a BAC-FISH study using the two BAC clones as probes showed that LFS genes were localized in the proximal region of the long arm of the chromosome. These results suggest that LFS genes in *A. cepa* are transcribed from at least two loci and that they are localized on chromosome 5.

2) Identification of essential amino acid residues for onion LFS activity

LFS, an enzyme essential for the synthesis of onion LF (propanethial *S*-oxide), was identified in 2002. This was the first reported enzyme involved in the production of thioaldehyde *S*-oxides via an intra-molecular H^+ substitution reaction, and we therefore attempted to identify catalytic amino acid residues of LFS as the first step for elucidating the unique catalytic reaction mechanism of this enzyme. Comparison of LFS cDNA sequences among lachrymatory *Allium* plants, deletion analysis, and site-directed mutagenesis enabled us to identify two amino acids (Arg71 and Glu88) that were indispensable to LFS activity. Homology modeling was performed for

LFS/23–169 on the basis of the template structure of pyrabactin resistance 1-like protein (PYL), which had been selected following a BLASTP search on SWISS-MODEL against LFS/23–169. In the modeled structure of LFS, a pocket corresponding to the ligand-binding site in PYL was identified; further, Arg71 and Glu88 were located in this pocket.

3) Detailed analysis of the reaction mechanism of onion LFS

We produced a single deuterated LF (*m/z* = 91) in a model reaction system comprising purified alliinase, LFS, and (*E*)-(+)-*S*-(1-propenyl)-L-cysteine sulfoxide [(*E*)-PRENCSO] in D_2O. These findings indicate that onion LFS is an (*E*)-1-propenylsulfenic acid isomerase.

ACKNOWLEDGEMENTS

I am most grateful to Dr. Masayoshi Shigyo, (Professor of Department of Biological and Environmental Sciences, Faculty of Agriculture, Yamaguchi University), for his valuable guidance, invaluable suggestions, advice, encouragement and insightful criticisms throughout this study.

I express my deep gratitude to Dr. Shinsuke Imai for his valuable guidance, helpful suggestions, advice, encouragement, and insightful criticisms throughout this study.

I sincerely thank Dr. Naoki Yamauchi (Professor, Department of Biological and Environmental Sciences, Faculty of Agriculture, Yamaguchi University) for his valuable assistance, kind support, and encouragement, which were invaluable to this research.

I would like to gratefully acknowledge Dr. Akiyoshi Fukamizu (Professor of Life Science Center, Tsukuba Advanced Research Alliance, University of Tsukuba) and Dr. Norio Ishida (Professor of Life Science and Bioengineering, Graduate School of Life and Environmental Science, University of Tsukuba) for their valuable assistance, kind support, advice, and constructive criticism in this study.

I would also like to thank Dr. Michihiko Kobayashi (Professor of Life Science and Bioengineering, Graduate School of Life and Environmental Science, University of Tsukuba) and Dr. Osamu Negishi (Assistant Professor of Life Science and Bioengineering, Graduate School of Life and Environmental Science, University of Tsukuba) for insightful discussions and scientific comments for this study.

I gratefully acknowledge Dr. Nobuaki Tsuge for his energetic, enthusiastic, and insightful participation in discussions, and for his scientific and technical advice, and suggestions throughout this work.

I thank Dr. John McCallum for his helpful advice and suggestions during this research and express my appreciation to Dr. Ludmila Khrustaleva for her performance of FISH experiments. I thank Dr. Shigenori Yaguchi for his help with the maintenance of the plants in the greenhouse.

I appreciate the support of all of my lab associates, who contributed directly or indirectly towards the completion of this study. I also thank House Foods Corporation for providing this valuable opportunity to conduct this research.

Finally, and most importantly, I would like to extend my deepest gratitude to my family for all their assistance, patience, support, inspiration, and encouragement over the course of this study.