Analyses of Genetic Diversity and Desirable Traits in Sesame (*Sesamum indicum* L., Pedaliaceae): Implication for Breeding and Conservation

Toan Duc Pham

(Phạm Đức Toàn)

Faculty of Landscape Planning, Horticulture and Agricultural Sciences
Department of Plant Breeding and Biotechnology
Alnarp

Doctoral Thesis
Swedish University of Agricultural Sciences
Alnarp 2011

Acta Universitatis Agriculturae Sueciae
2011:42

Cover: Sesame in the field, flowers, capsules, seeds, molecular markers, and cluster analysis
(Photo: Pham Duc Toan)

ISSN 1652-6880
ISBN 978-91-576-7586-6

Print: SLU Service/Repro, Alnarp 2011

Analyses of Genetic Diversity and Desirable Traits in Sesame (*Sesamum indicum* L., Pedaliaceae): Implication for Breeding and Conservation

Abstract

Sesame (*Sesamum indicum* L., Pedaliaceae) is a traditional oil crop cultivated in Vietnam and Cambodia. It is known as the king of oil seeds in Vietnam due to the high oil content (50-60%) in its seed. Yet, the insufficient genetic information regarding Vietnamese and Cambodian sesame populations is limiting the access to useful traits present among adapted landraces in this region. The purpose of this study was to characterize various sesame accessions to gain information that could help design strategies for future breeding program and conservation of this crop in the two countries. Morphological and molecular markers as well as oil content and quality analyses were employed to evaluate sesame accessions from different sources.

High genetic variation was found among populations of sesame collected in Vietnam and Cambodia. The two type of markers, morphological and molecular, were both useful in analyzing the extent of genetic diversity in sesame and the result of these analyses will help to better understand the genetic diversity and relationship within and among populations. Overall, the sesame accessions included in the study showed a correlation with their geographical origins such that accessions from the same region tended to have higher genetic similarity as compared to those from different regions. However, when cluster analysis was applied to evaluate the genetic relationship, some sesame accessions were found not to be grouped based on geographical origins. This contrasting result could perhaps be a result from the exchange over time, of sesame germplasm, between farmers across the regions.

The results from morphological and oil content analyses showed that several sesame accessions in Southern Vietnam and Cambodia displayed a good potential for high seed yield and oil content. Overall, the studies in this thesis provide important insights into the populations of sesame in Vietnam and Cambodia and constitute a set of useful background information that can be used as a basis for future breeding strategy and improvement of sesame in this region.

Keywords: Sesame, breeding, diversity, molecular markers, morphology, oil content, Vietnam, Cambodia

Author's address: Toan Duc Pham, SLU, Department of Plant breeding and Biotechnology, P.O. Box 101, 230 53 Alnarp, Sweden

E-mail: Toan.Pham@ltj.slu.se or phamductoan@hcmuaf.edu.vn

Dedication

To my family with gratitude for your unfailing love and support, and for believing in me through the years

Contents

List of Publications

This thesis is based on the work contained in the following papers, referred to by Roman numerals in the text:

I Toan Duc Pham, Tri Minh Bui, Gun Werlemark, Tuyen Cach Bui, Arnulf Merker, Anders S. Carlsson (2009). A study of genetic diversity of sesame (*Sesamum indicum* L.) in Vietnam and Cambodia estimated by RAPD markers. *Genetic Resources and Crop Evolution* 56(5):679-690.

II Toan Duc Pham, Mulatu Geleta, Tri Minh Bui, Tuyen Cach Bui, Arnulf Merker and Anders S. Carlsson (2011). Comparative analysis of genetic diversity of Sesame (*Sesamum indicum* L.) from Vietnam and Cambodia using agro-morphological and molecular markers. *Hereditas* 148(1):28-35

III Toan Duc Pham, Thuy-Duong Thi Nguyen, Anders S. Carlsson and Tri Minh Bui (2010). Morphological evaluation of sesame (*Sesamum indicum* L.) varieties from different origins. *Australian Journal of Crop Science* 4(7):498-504

IV Toan Duc Pham, Mulatu Geleta, Aman Bonaventure, Tri Minh Bui, and Anders S. Carlsson. Study of oil content and fatty acid composition of sesame (*Sesamum indicum* L.) in Vietnam and Cambodia (*Manuscript*)

The contribution of Toan Duc Pham to the papers included in this thesis was as follows:

I Planned, collected the samples, carried out all experimental work, evaluated data, analysed data and wrote the manuscript in cooperation with co-authors

II Planned, carried out all experimental work, evaluated data, analysed data and wrote the manuscript in cooperation with co-authors

III Planned, set up experiment, carried out all experimental work, evaluated data, analysed data and wrote the manuscript in cooperation with co-authors

IV Planned, evaluated data, analysed data and wrote the manuscript in cooperation with co-authors

Abbreviations

PCR	Polymerase chain reaction
DNA	Deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
dNTP	Deoxynucleotide triphosphates
RAPD	Random Amplification Polymorphic DNA
SSR	Simple sequence repeat
UPGMA	Unweighted pair group method with arithmetic mean
NTSYS	Numerical Taxonomy System
UFA	Unsaturated fatty acid
PUFA	Polyunsaturated fatty acid

1 Introduction

Sesame (*Sesamum indicum* L.) is an annual plant that belongs to the Pedaliaceae family (Fig 1). It is considered to be the oldest of the oilseed plants and has been under cultivation in Asia for over 5000 years (Bisht *et al.*, 1998). Nowadays, it is widely cultivated as an oil-crop in tropical and subtropical climates. Renowned for its high oil content with seeds that can contain up to 60% oil, sesame is known as the king of oil seeds in Vietnam. The oil has a composition that provides good health benefits including high levels of unsaturated fatty acids (80%) and antioxidants. Possibly for this reason, sesame oil is widely considered to prevent diseases of different kinds.

Fig.1 Sesame (A) plant with capsules along the stem and (B) plant at flowering stage

Beside food, sesame also finds its uses in application areas such as pharmaceutics, industrial, and as biofuel. Sesame has been distributed and cultivated in various ecological regions of Vietnam and Cambodia for hundreds of years.

1.1 Taxonomy and cytogenetics

The genus *Sesamum* consists of many species and the most cultivated is *Sesamum indicum* L. (Ashri, 1998). According to Kobayashi *et al.* (1990), 36 species have been identified of which 22 species have been found in Africa, five in Asia, seven in both Africa and Asia, and one species each in Crete and Brazil. There are three cytogenetic groups of which 2n = 26 consist of the cultivated *S. indicum* along with *S. alatum, S. capense, S. schenckii, S. malabaricum*; 2n = 32 consist of *S. prostratum, S. laciniatum, S. angolense, S. angustifolium;* while *S. radiatum, S. occidentale, S. schinzianum* belong to 2n = 64. Mainly due to the difference in chromosomal numbers across the three cytotaxonomic groups, there is limited cross compatibility among the species. Therefore, it has been difficult to transfer desirable characteristics such as drought tolerance, pest, and resistance to diseases, from wild relatives into cultivated sesame (Carlsson *et al.*, 2008).

1.2 Origin and geographical distribution

The origin of sesame has been a major subject of discussion with the African or the Indian subcontinent as the two suggestions. Hiltebrandt (1932) considered Africa as the original home of sesame, since this continent seems to be a center of sesame diversity as seen from the high number of wild species found. Bedigian (1981) argues that owing to the wide genetic diversity in Africa it is reasonable to assume that this subcontinent is the primary center of origin. India would then be thought of as a secondary center for sesame. On the other hand, India is generally held as the subcontinent where sesame was first domesticated and then spread to other places in the world such as Africa, the Far East, China and Americas along trade routes Bedigian (2004). Today, sesame is widely grown as an oilseed crop in India, China, Korea, Japan, Turkey, Thailand, Vietnam and Cambodia as well as on the American and African continents (Fig 2).

1.3 Growth habit

Sesame is a short-day plant and is normally self-pollinated, although cross-pollination ranging from 5 to over 50% occurs (Rheenen, 1980; Pathirana, 1994). It is an erect herbaceous annual plant that has two growth

characteristics indeterminate and determinate, with the plants reaching heights of up to two meters. Most varieties show an indeterminate growth habit, which is shown as a continuous production of new leaves, flowers and capsules as long as the environment remains suitable for growth (Carlsson *et al.*, 2008). The growth period range from 70 to 150 days depending on the variety and the conditions of cultivation (Ashri, 1998). However, sesame varieties in Vietnam and Cambodia have a growth period ranging from 70 to 100 days, with the first flower produced within 4-5 weeks after planting.

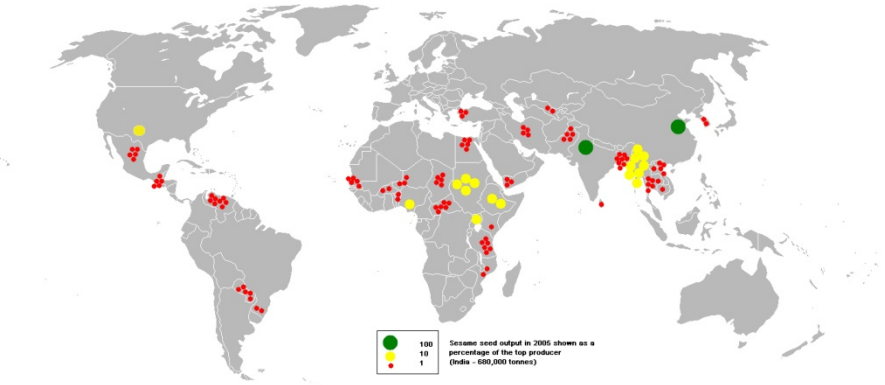

Fig.2 Map of sesame distribution in the world. Source: wikipedia.org

1.4 Agro-ecology, pests and diseases

Sesame can grow well in many ecological regions of tropical and subtropical climate, and also grows in most kind of soils. However, it thrives best on well-drained soil with a moderate fertility and a pH between 5.5 and 7.0. Sesame is highly drought tolerant, and it can adapt and produce seed well under fairly high temperatures (Ashri, 1998). However, moisture levels before planting and flowering have great impact on the seed yield. Sesame requires a minimum of 300–400 mm of rainfall per season, but is sensitive to wet condition and has a very low salt tolerance (Carlsson *et al.*, 2008).
Green peach aphid (*Myzus persica*), white fly (*Bemisia argentifolium, Bemisia tabacci*), cabbage looper (*Pieris rapae*) and army worm (*Cupis unipuncta*) are example of pests that attack sesame. Bacteria and fungi such as bacterial leaf blight (*Pseudomonas sesami*), root rots (*Fusarium oxysporum, Phytophtora parasitica, Rhizoctonia bataticola*) and leaf spot (*Cercospora sesamicola, Cercospora sesame, Corynespora cassiicola, Alternaria sesami*) also causes problem for sesame (Tripathi *et al.*, 1998; Ojiambo *et al.*, 1999; Langham & Wiemers, 2002; John *et al.*, 2005; Choudhary *et al.*, 2006).

1.5 Advantage of sesame cultivation

Beside advantages such as drought tolerance and acceptance of a range of different soils, sesame stimulates a flora of beneficial soil microbes and reduces the nematode populations, particularly the root knot nematodes (*Meloidogyne arenaria, Meloidogyne incignita*) that attack peanuts and cotton. Cotton following sesame in a crop rotation scheme show a significantly reduced root-rot caused by *Phymatotrichopsis omnivora* (Langham *et al.*, 2008). Sesame is also an excellent soil builder because of the large amounts of root biomass that are left to decay underground after harvest. The soil is very mellow after sesame cultivation and therefore requires less work to prepare for the next crop in rotation (Langham *et al.*, 2008).

1.6 Status of production of sesame

India and China are the world's largest producers of sesame, followed by Myanmar, Sudan, Uganda, Ethiopia, Nigeria, Tanzania, Pakistan and Paraguay (FAOSTAT, 2008). In 2008, the total world production was about 3.54 million tons that was grown on 7.42 million hectares. In recent years, sesame has had a low ranking in the world production of edible oil crops. The low ranking of sesame among oil crops may be attributed to several factors including low seed yield and strong competition from other oil crops such as soybean, sunflower, peanut.

In Vietnam, sesame was grown on about 45000 hectares, and produced about 22000 tons of seed (FAOSTAT, 2008). Most of the sesame was grown in crop rotation in small-scale production. Presently, sesame could not compete with other crops such as tobacco, soybean, groundnut, rice, sweet potato, sugarcane and cassava. Due to aforementioned factors and others such as lack of improved varieties and a low income for the farmers as a consequence from cultivating low yielding landraces, sesame has therefore been given less attention from the farmer.

1.7 Oil content and fatty acid composition

Among oil crops, sesame is one of the highest in oil content. Generally, the oil content in sesame ranges from 34 to 63% (Yermanos *et al.*, 1972; Ashri, 1998; Baydar *et al.*, 1999; Uzun *et al.*, 2002; Were *et al.*, 2006b). Genetic and environmental factors influence the oil content and fatty acid compositions in sesame (Carlsson *et al.*, 2008). Late maturing cultivars are reported to have higher oil content than early cultivars (Yermanos *et al.*, 1972) and the indeterminate cultivars accumulated more oil than determinate ones (Uzun *et al.*, 2002). Sesame contains high levels of antioxidants such as sesamol, sesamin, sesamolin and sesaminol. Therefore,

sesame oil is called the queen of the vegetable oils because of its antioxidants (ASGA, 2011). The composition of sesame oil consist of mainly four fatty acids (palmitic–C16:0, stearic–C18:0, oleic–C18:1 and linoleic–C18:2), while other fatty acids appear in very small amounts (Ashri, 1998). Even though different fatty acid compositions of sesame oil have been reported the major fatty acids are oleic and linoleic acids (Kamaleldin *et al.*, 1992). Yermanos *et al.* (1972) reported oleic acid level ranging from 32.7% to 58.2% and linoleic acid from 27.3% to 59.0%. Were *et al.* (2006b), reported variation in the quantities of palmitic, stearic, oleic and linoleic acids, with palmitic and stearic acids ranging from 7.2 to 9.6% and 3.7 to 5.6%, respectively, while high levels of oleic acid and linoleic acid ranged from 31.6 to 41.9% and 42.9 to 53.9%. Linolenic acid was found but in very small amount (0.5%).

1.8 Economic importance of sesame

Although sesame seeds are used as an ingredient in many different food supplies, a major part of the sesame seed production is processed into oil and meal (Morris, 2002). Sesame oil is, as mentioned before, an excellent vegetable oil because of its high contents of antioxidants such as sesamin, sesamol and sesamolin and its fatty acid composition (Suja *et al.*, 2004). The antioxidants make the oil very stable and it has therefore a long shelf life (Chung *et al.*, 2004; Suja *et al.*, 2004). In sesame oil, oleic (C18:1) and linoleic (C18:2) acids are the predominant fatty acids and make up more than 80% of the total fatty acids. The high levels of unsaturated (UFA) and polyunsaturated fatty acids (PUFAs) increase the quality of the oil for human consumption (Nupur *et al.*, 2010). Moreover, high level of PUFAs in sesame oil is claimed to reduce blood cholesterol, high blood pressure and play an important role in preventing atherosclerosis, heart diseases and cancers (Ghafoorunissa, 1994; Hibasami *et al.*, 2000; Miyahara *et al.*, 2001). People in Vietnam use sesame oil for cooking fish, meat, and frying vegetable.

Sesame seed flour has a high protein content, with high levels of the essential amino acids methionine and tryptophan, contains about 10 to 12% of oil and has three times more calcium than milk (Morris, 2002). The meal that remains after oil extraction from sesame seeds is an excellent feed for poultry and livestock. A potential problem, though, with sesame when used for human consumption is that it contains one of the top food allergens. Its allergenicity emanates from protein sources, such as the 14 kDa 2S albumin precursor, which belongs to one of four protein families known to be allergenic and several oleosins (15 and 17 kDa), also known to cause hypersensitivity (Wolff *et al.*, 2003; Leduc *et al.*, 2006). Therefore, food

products that contain sesame as an ingredient are now required to be labeled as potential allergenic in Europe and Canada (Carlsson *et al.*, 2008).

1.8.1 Use of sesame for pharmaceutical purposes

Due to the fact that sesame seeds are rich in PUFAs, as well as vitamin and minerals such as calcium, magnesium and phosphorus, its oil has health benefits. Therefore, sesame oil has been used as medicine or for pharmaceutical purposes (Bedigian, 2004). Sesame oil contains vitamin E and several important antioxidants such as sesaminol and sesamolinol, that are believed to promote the integrity of body tissues in the presence of oxidizing compounds (Morris, 2002). Cooney *et al.* (2001) found that intake of food made from sesame seed, which contain gamma-tocopherol, significantly elevated its levels in the serum. Gamma-tocopherol has been shown to positively influence vitamin E activity that is believed to prevent cancer and heart disease (Cooney *et al.*, 2001). Sesame oil is used as an pharmaceutics aid (Jellin *et al.*, 2000) and as such used to remedy gum disease, treat toothaches, relieve anxiety or insomnia, or used as an antibacterial mouthwash by Chinese and Indian in the past (Annussek, 2001; Morris, 2002).

1.8.2 Non-food application

Sesame oil has found use in several non-food applications such as in paints, cosmetics, solvents and soap (Bedigian, 2004). In Africa, sesame flowers have been used to prepare perfumes (Morris, 2002). The antioxidant sesamin is used as a synergist for pyrethrum or rotenone insecticides and increases the toxicity of insecticides when spayed against flies (Haller *et al.*, 1942). Hasan *et al.* (2000) reported that chlorosesamone extracted from roots of sesame has antifungal properties and was used as a fungicide. Sesame oil has been investigated as a source for biodiesel and found to give a product with fuel properties in parity with mineral diesel but with superior environmental performance (Ahmad *et al.*, 2010).

1.9 Genetic diversity and biotechnology studies in sesame

Genetic diversity of crops plays an important role in sustainable development and food security (Esquinas-Alcazar, 2005), as it allows the cultivation of crops in the presence of various biotic and abiotic stresses. It is also important for selection of parents that can be used in plant breeding programs. Information on genetic diversity is important when working to improve crop varieties. Genetic diversity is studied by using various methods such as morphological, biochemical and molecular markers. Morphology has been a primary tool to estimate genetic differences among sesame genotypes. Several studies based on morphological markers have found a high genetic

diversity in sesame populations (Bedigian & Harlan, 1986; Ganesh & Thangavelu, 1995; Bisht *et al.*, 1998; Arriel *et al.*, 2007). However, morphological markers have limitations in their ability to estimate genetic diversity because of strong influence from environmental factors, which make them highly dependent on the cultivation conditions. Molecular markers overcome this limitation. Molecular markers techniques such as amplified fragment length polymorphism (AFLP), random amplified polymorphic DNA (RAPD), inter simple sequence repeats (ISSR) and simple sequence repeats (SSR) have been widely used in genetic diversity studies in sesame. Low genetic diversity (0.14–0.21) among groups were reported by Laurentin & Karlovsky (2006) in 32 sesame germplasm collections using amplified fragment length polymorphism (AFLP). Kim *et al.* (2002) reported that a ISSR-based study of sesame from Korea and other countries revealed a low level of polymorphism and the genetic distance between populations ranged from 0 to 0.25. In contrast, Bhat *et al.* (1999) reported a high genetic diversity of 0.35 among 58 sesame accessions from the Indian subcontinent (36 accessions) and other countries (22 accessions). Ercan *et al.* (2004) using RAPD markers reported a genetic diversity among 38 Turkish sesame accessions ranging from 0.14 to 0.40. Similarly, a high genetic diversity of sesame in Vietnam and Cambodia was found using RAPD, as estimated by percent polymorphism (83%) and genetic distance coefficient (0.03-0.31) (Pham *et al.*, 2009). Microsatellites are one of the most commonly used molecular markers to determine the genetic diversity in crop species. However, only a few studies used microsatellites to evaluate genetic diversity in sesame (e.g. Dixit *et al.*, 2005; Wei *et al.*, 2008). Molecular markers have been developed to identify morphological traits of sesame such as growth habit and closed capsule trait (Uzun *et al.*, 2003; Uzun & Cagirgan, 2009). Moreover, biotechnology techniques such as *in vitro* regeneration and genetic transformation have also been developed for sesame (Were *et al.*, 2006a; Yadav *et al.*, 2010).

1.10 Conservation and breeding of sesame

The purpose of conservation is to conserve plant genetic resources for potential future usage, and as such it should support basic research and improvement of crops. Core collections of sesame germplasm have been established by the Oil Crop Research Institute of the Chinese Academy of Agricultural Sciences, and the National Bureau of Plant Genetic Resources (NBPGR) of India in collaboration with the International Plant Genetic Resources Institute (IPGRI). Thus, NBPGR maintains 6658 accessions of sesame where 4136 are indigenous and 2522 are from exotic sources (Bisht *et al.*, 1998). The Gene Bank of Rural Development Administration (RDA) located in Suwon, Korea have collected 7698 sesame accessions, that consist

of 3538 exotic collections, 2660 indigenous collections, 1072 improved genetic stocks and 428 others (Kang *et al.*, 2006). In addition, conservation efforts of sesame have been done by other organization and gene banks in the world. For example the United States Department of Agriculture (USDA), Agricultural Research Service (ARS), Plant Resources Conservation Unit (PGRCU) has conserved 1226 sesame accessions originating from Europe, Africa, Asia, North America and South America (Morris, 2009).

Although, sesame is considered to be the oldest of the oilseed plants and has been under cultivation in Asia for over 5000 years (Bisht *et al.*, 1998), in comparison with many other oil crops it has a low seed yield with a world average of about 477 kg ha^{-1} and about 500 kg ha^{-1} in Vietnam (FAOSTAT, 2008). The low seed yield of sesame is a consequence of a lack in breeding attention (Ashri, 1998). Sesame production is also limited by pests, diseases, lack of uniform maturity of capsules, and seed shattering (Langham & Wiemers, 2002).

Screening for resistance to diseases was done among sesame from Kenya and two cultivars, SPS045 and SIK013, were found to be resistant to angular leaf spot (*Cercospora sesamicola*) and white leaf spot (*Cercospora sesame*), respectively while another cultivar SIK031 was resistance to both of those diseases (Nyanapah *et al.*, 1995). In another study sesame mutants were screened for resistance to *Fusarium* wilt disease (Silme & Cagirgan, 2010), and 25 sesame genotypes were evaluated for their reaction to *Fusarium oxysporium* f.sp. *sesame*. It was reported that four genotypes, Birkan, Cambidi, WS-143 and WS-313 were classified as resistant (R), and genotype Birkan was released as a cultivar for high seed yield and resistant to *Fusarium* wilt disease. It was suggested that genotype Birkan could be used as a source for resistance to *Fusarium* wilt disease in sesame breeding program. Efforts to reduce shattering in sesame (non-dehiscent sesame) in order to improve seed yield and make this crop suitable for mechanical harvesting has been undertaken in the United States (Langham & Wiemers, 2002). Molecular markers linked to traits such as closed capsule (or non-dehiscent capsule) and growth habit were developed to identify sesame cultivars with the desired traits at an early developmental stage (Uzun *et al.*, 2003; Uzun & Cagirgan, 2009). Recently, the exploitation of genetic diversity and heterosis has been another approach to improve the seed yield as well as some other traits in sesame varieties. Crossing experiments of different sesame lines resulted in hybrid vigor, particularly in seed yield. (Gaikwad *et al.*, 2009; Banerjee & Kole, 2010; Durai *et al.*, 2010; Parameshwarappa & Salimath, 2010; Prajapati *et al.*, 2010). Agro-morphological evaluation of sesame to select good varieties has resulted in varieties that can reach over 1 ton of seed ha^{-1} (Furat & Uzun, 2010; Langham *et al.*, 2010). However,

those sesame varieties should be grown in their recommended specific region, because sesame is very sensitive to climatic conditions such as temperature, day length and related humidity. Sesame produces no seed or a very low seed yield if it is grown under unfavorable environmental conditions (Pham *et al.*, 2010). In general, the breeding strategies for sesame are similar to those applicable in other crops and include higher yields, improved plant architecture, length of growing season, resistance to diseases and pests (Ashri, 1998). Specific objectives for sesame breeding vary with the level of technology and were summarized by Ashri (1998) as follows: (1) high and stable seed yield of good quality under a wide range of environmental conditions, (2) resistance to water logging, drought, salinity, pests, diseases, shattering and other abiotic stresses, (3) increased number of capsules per leaf axil, full seed-set without aborted ovules, and (4) uniform plant type, rapid growth, good adaptation to varying environment conditions and seasons. These objectives are applicable to all sesame producing countries including Vietnam and Cambodia.

2 Objectives of the study

The main objective of this thesis was to study the phenotypic and genotypic genetic diversity of sesame (*Sesamum indicum* L.). The specific objectives were to:

- characterize the agro-morphological and molecular genetic diversity in the sesame germplasm
- compare and understand the degree of congruence between the agro-morphological and molecular markers in genetic diversity analyses
- evaluate the genetic relationship of sesame accessions from different agro-ecological regions
- evaluate morphological and phenotypic traits of sesame varieties from different origins based on morphological character
- determine the seed oil content, oil quality and seed yield potential of different sesame accessions for their future use in the breeding programs.

3 Materials and methods

3.1 Plant material

The sesame (*Sesamum indicum* L.) germplasm used in this study were collected from different regions in Vietnam and Cambodia. The Vietnamese accessions consisted of collections from the three regions North, Central and South Vietnam and the Cambodian accessions included accessions from Kandal, Kampongcham, Kampongthum provinces. In addition, five other sesame accessions were received from Dr. Beatrice Angiyo Were, Moi University, Kenya.

3.2 DNA extraction

DNA was extracted from young leaves of 3- to 4-weeks old seedlings using a modified protocol described in Warwick & Gugel (2003). DNA pellet was washed twice in ethanol (75%) and allowed to dry for 20-30 min. After dissolving the DNA pellet in 30 µl TE (10 mM Tris pH 8.0, 1 mM EDTA pH 8.0) buffer containing RNAse (2 µl of 20 mg/ml RNAse mixed in 1 ml TE), concentration of DNA was measured using a NanoDrop spectrophotometer (NanoDrop Technologies, Inc, USA). DNA samples were stored at -20°C, until further use.

3.3 RAPD assay

The PCR-RAPD procedure for sesame was carried out as described by Williams *et al.* (1990). Amplification reaction was performed in a volume of 25 µl containing 1X PCR reaction buffer (10 mM Tris-HCl, pH 8.3 and 50 mM KCl), 0.2 mM dNTP (each of dATP, dTTP, dCTP, dGTP), 3 mM $MgCl_2$, 1 U Taq polymerase (Sigma, Germany), 0.2 µM primer (Operon Biotechnologies) and about 50–60 ng DNA template. The PCR reaction

was carried out in a thermal cycler (GeneAmp PCR Applied Biosystems 9700, Singapore) using 5 min at 94°C for denaturation followed by 40 cycles of 60 sec at 94°C, 90 sec at 35°C, and 180 sec at 72°C. The cycles were followed by a final extension period of 10 min at 72°C. The electrophoresis of PCR products, gel staining, visualization and photographing were carried out as described in paper I & II.

3.4 Development of SSR markers for sesame

3.4.1 Microsatellite mining and primer design

A total of 3699 sesame EST sequences were downloaded from the NCBI Expressed Sequence Tags (ESTs) database. The MSATCOMMANDER 0.8.1 software (Faircloth, 2008) was used to screen each sequence for microsatellites. The criteria for minimum number of repeat motifs were six repeats for dinucleotides and four repeats for tri-, tetra- and pentanucleotides. After redundant sequences were removed from the mined microsatellite containing sequences, 108 sequences were selected. These sequences were thoroughly evaluated to screen sequences that are suitable for primer design. SSR primers were then designed using the Primer3 primer designing program (Rozen & Skaletsky, 2000) for selected sequences. A total of 31 primer-pairs were designed and tested for PCR amplification, of which seven primer-pairs were used for final data analysis. Additional data are being generated using the remaining 24 primer-pairs.

3.4.2 SSR reaction

The PCR reaction was conducted in a volume of 25 µl containing 25 ng genomic DNA, 0.3 µM forward and reverse primers, 2 mM $MgCl_2$, 0.3 mM dNTPs, 1 U Taq polymerase (Sigma, Germany) and 1X PCR buffer (10 mM Tris-HCl, pH 8.3 and 50 mM KCl). The PCR reaction was carried out using the Eppendorf AG 5345-016209 (Germany) thermo-cycler using the following thermo-cycling conditions: 1 cycle of 95°C for 3 min, followed by a touchdown -1°C/cycle of 6 cycles of 94°C for 30 sec, 62-68°C for 30 sec, 72°C for 45 sec, and then 32 cycles of 94°C for 30 sec, 55-61°C for 30 sec, 72°C for 45 sec and a final extension step of 72°C at 20 min. The annealing temperature was changed based on *Ta* values of each SSR primer-pairs. The PCR products were separated through capillary electrophoresis on the ABI 3100 genetic analysis (Applied Biosystems, Singapore) at Genomics Core Facility, Göteborg University. The Peak scanner software v1.0 (Applied Biosystems) was used to size peak patterns using the internal Genescan-500 LIZ size standard, and Sizing default PP analysis method.

3.5 Field trial experiment

The field experiments were carried out at Can Tho and Tien Giang provinces, in Mekong delta region, in the South-Western part of Vietnam, where the annual temperature fluctuates between 25°C and 35°C and average annual rainfall measures up to 1500 mm. The experimental field layout and the measurement of agronomic traits were described in detail in paper II & III.

3.6 Data analysis

In the case of RAPD, the amplified fragments were treated as dominant markers and each locus was considered as a bi-allelic with one amplifiable and one null-allele. Data was scored as 1 for the presence and 0 for the absence of a DNA band for each locus across all individuals. In the case of SSR, each peak is considered as an allele at a co-dominant locus and the genotype of each individual at each locus was recorded and used for various genetic diversity analyses. Alleles were also binary coded as 1 or 0 for their presence or absence, respectively in each genotype, and this data was combined with RAPD data for cluster analyses. Dendrograms were constructed based on unweighted pair group method with arithmetic mean (UPGMA) clustering method using the computer software NTSYSpc version 2.10 (Rohlf, 2000). Shannon diversity index of phenotypic diversity (H'), Nei's gene diversity (H), expected heterozygosity (*He*), observed heterozygosity (*Ho*) and F-statistics (*Fis, Fit, Fst*) were computed with the POPGENE software version 1.31 (Yeh & Boyle, 1997). The bootstrap analysis was done using FreeTree program (Pavlicek *et al.*, 1999). The TreeView version 1.6.6 (Page, 1996) was used to view the tree. The Minitab version 15.0 (Minitab, USA) was used to analyse the significant differences as ANOVA and for cluster analysis of morphological data.

4 Results and Discussion

4.1 Molecular genetic diversity of sesame (paper I and SSR results)

Due to several advantages molecular markers have been widely used to estimate genetic diversity in crop species including sesame. RAPD and microsatellite markers were used to evaluate the extent of genetic diversity among Vietnamese and Cambodian sesame accessions. RAPD technique was chosen as it does not rely on prior knowledge of DNA sequences, is comparatively cheap and simple (Bhat *et al.*, 1999; Ercan *et al.*, 2004; Salazar *et al.*, 2006). Analysis of genetic diversity of sesame using ten RAPD primers resulted in overall diversity of 0.513 and 0.340, respectively, when estimated using Shannon index as H'_{sp} and Nei's gene diversity estimate (Nei, 1978) with modification of Lynch & Milligan (1994) as H_t. The overall within population genetic variation estimated with Shannon diversity index as H'_{pop} and Nei' gene diversity as H_s was 0.234 and 0.160, respectively. The extent of genetic variation in each population was estimated using both Shannon diversity and mean gene diversity as H'_{loci} and $H_{GD,}$ respectively, which is the mean value across the whole loci. H'_{loci} ranged from 0.11 to 0.48, and H_{GD} ranged from 0.08 to 0.33.

In the case of SSR data, expected heterozygosity (*He*) and observed heterozygosity (*Ho*) ranged from 0.12 to 0.48 (mean, 0.29) and 0.00 to 0.30 (mean, 0.15), respectively. The estimate of gene flow (Nm) was 0.38 on average. The *Fis, Fit, Fst* were also estimated to determine the genetic structure of the accessions. The mean *Fis, Fst* and *Fit* were 0.19, 0.51 and 0.40, respectively. The SSR based analyses revealed that average values per accession for percent polymorphism, *He* and *Ho* were 48.9, 0.18 and 0.14, respectively. The expected heterozygosity *He* (0.29) obtained in this study is lower than that obtained by Dixit *et al.* (2005) for SSR markers on Korean sesame (0.69). The low *He* in this study can be partly explained by the fact that the microsatellites were developed from ESTs, which are generally less

variable than unexpressed regions of the genome. Given the fact that our analyses were based on only four SSR primers, these values may slightly change when new data is added. However, these data can be used to compare and identify accessions with high genetic variation. The data generated with four newly developed polymorphic microsatellite markers were combined with RAPD data generated using 10 primers (paper I) in order to evaluate the extent to which the clustering pattern of accessions is affected by the addition of microsatellite data (Fig. 3)

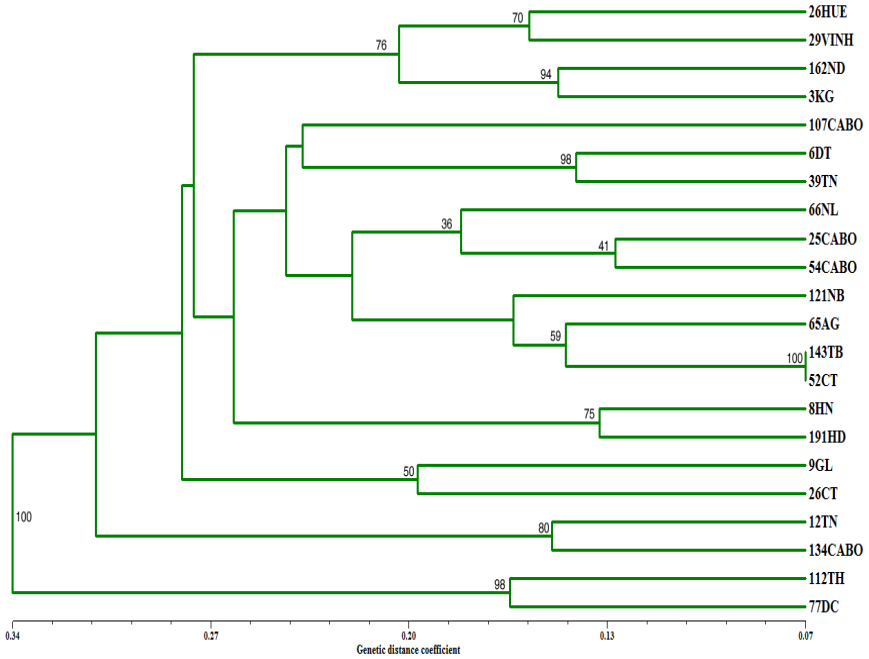

Fig.3 UPGMA based dendrogram generated using a combined data of RAPD and microsatellite markers. Numbers near the branches are bootstrap values.

The dendrogram obtained using combined data of RAPD and SSR markers is quite similar to the cluster generated from RAPD data (paper I). Combined analysis revealed that the genetic distance coefficient between accessions ranged 0.07 to 0.34, which is close to that of the RAPD data (0.03 to 0.31). In this combined analysis, most of the accessions were grouped in similar clusters as that of clusters generated based on RAPD data; such as 112TH, 77DC; 143TB, 52CT; 12TN, 134CABO; 26HUE, 29VINH. A high variation in genetic distance between accessions was found in both Vietnamese and Cambodian sesame accessions (Fig.3). Although sesame is mainly self-pollinated, some authors have reported levels of out-

crossing between 5–60% in this species (Rheenen, 1980; Yermamos, 1980; Pathirana, 1994). Out-crossing plant species tend to present between 10 and 20% of the genetic variation between populations (Hamrick & Godt, 1989) and the remaining 80-90% of the total genetic variation within populations. Hence, some degree of out-crossing could explain the high genetic diversity observed in the studied sesame accessions. The results of this study are in agreement with the previous studies based on morphological, agronomic traits and molecular markers which have reported high genetic diversity in sesame germplasm (Ashri, 1998; Bhat *et al.*, 1999; Ercan *et al.*, 2004; Salazar *et al.*, 2006).

4.2 Morphology and congruence between morphological and molecular genetic diversity (paper II)

Analysis of genetic diversity in crop species using more than one methods helps to better understand the levels of genetic variation and the genetic structure of populations when compared to the results obtained using only one method, as each method has its own advantages and disadvantages. A high variation in plant height, and number of branches/plant was found among the studied accessions. In some accessions, branches were few or absent whereas in other accessions branches appeared at an early vegetative growth phase, at about 20–30 days after planting. A high variation was also found in several other agronomic traits such as capsules per plant and seeds per capsule. The flower color of sesame accessions in this study was observed to be predominantly white. Leaves and capsules were varying much in their size and shape.

A significant positive correlation between different markers used in genetic diversity analyses has been reported in various crop species, such as maize, *Zea mays* L. (Kantety *et al.*, 1995; Pejic *et al.*, 1998), barley, *Hordeum vulgare* L. (Russell *et al.*, 1997), soybean, *Glycine max* (L.) Merr., (Powell *et al.*, 1996). In the present study, a relatively high correlation ($r=0.88$, $P<0.001$) was found between the agro-morphological and RAPD data in terms of relationship between accessions. The two methods are in agreement in grouping 7 out of the 12 accessions in to two groups in this study: (1) DT, TN, AG, and (2) HD, TH, EKD, KPC (Table 1). The correspondence between agro-morphological and RAPD markers also has been reported in species such as *Andrographis paniculata* (Burm.f.) Nees (Lattoo *et al.*, 2008) and *Trifolium pretense* L. (Greene *et al.*, 2004). The range of agro-morphological data based genetic distance between pairs of accessions was wider (0.02 to 0.47) than that of RAPD based data (0.06 to 0.27), but the mean genetic distance between accessions estimated based on the two methods were similar (0.23 and 0.22, in that order). However, the clustering pattern of the accessions at country and regional level showed

some discrepancies (Table 1). Many previous studies have reported that genetic diversity obtained from agro-morphological markers was not entirely similar to RAPD markers (Black-Samuelsson *et al.*, 1997; Talebi *et al.*, 2008). The difference in results between the agro-morphological and RAPD markers might occur if agro-morphological similarity was due to different combinations of alleles producing similar phenotypes (Johns *et al.*, 1997). Differences could also occur if a single or few genes controlled the expression of agro-morphological traits that RAPD markers fails to detect (Steiner & de los Santos, 2001). The mismatch between diversity, based on RAPD and agro-morphological traits, might be that most of the agronomic traits are controlled by polygenes and these traits are highly influenced by environmental conditions (Dey *et al.*, 2006). Linhart & Grant (1996) indicated that the discordance might also be due to differences in evolutionary rates between agro-morphological characters and characters originating from selectively neutral, non-coding DNA region if the agro-morphological characters have adaptive value and the molecular markers are selectively neutral. In addition, several other factors may affect the estimation of genetic diversity and relationship between individuals and populations, such as type and number of markers used, distribution of markers in the genome, and the nature of evolutionary mechanisms underlying the variance measured (Powell *et al.*, 1996). A significant relationship between molecular markers and morphology could perhaps be obtained if the markers were linked to the morphological traits under study (Black-Samuelsson *et al.*, 1997; Persson & Gustavsson, 2001).

Table 1 Clustering patterns of 12 sesame accessions based on RAPD and agro-morphological markers

Molecular RAPD markers				Agro-morphological markers		
Cluster	nrA [(a)]	Name of accessions	Origin	nrA [(a)]	Name of accessions	Origin
I	2	VINH, ND	N-VN	5	TN, AG, VINH, DT, SKD	S-VN, N-VN & Cambo
II	2	SKD, KPT	Cambo	1	CT	S-VN
III	4	DT, TN, CT, AG	S-VN	6	HD, EKD, ND, KPC, TH, KPT	N-VN & Cambo
IV	4	HD, TH, EKD, KPC	N-VN & Cambo	---	---	---

[(a)]nrA: number of accession

N-VN: Northern Vietnam; S-VN: Southern Vietnam; Cambo: Cambodia

4.3 Genetic relationship of sesame accessions between ecological regions in Vietnam and Cambodia (paper I, II)

A high genetic variation was found in sesame accessions as presented in paper I & II. Generally, the accessions were correlated with their geographical origin, and in most cases, those from the same region had a relatively close genetic relationship. However, cluster analyses showed that some accessions did not group together with other accessions from the same geographical region. One possible reason for this is the exchange of sesame varieties between farmers across regions. The close relationship between some sesame accessions from Northern Vietnam and Southern Vietnam might be due to gene flow along with the movement of immigrants from the Northern to Southern Vietnam after Vietnam-American war, as the immigrants carried sesame seeds from different regions in the North to new geographical regions in the South for cultivation. Similarly, the close genetic relationship between Vietnamese and Cambodian sesame might be partly due to germplasm exchange and free trade between farmers across Vietnam-Cambodia border. Natural distribution of genetic diversity in plant species depends on its isolation, habitat alteration due to climatic change or its evolution, ecological, geographical factors, and more often on human breeding activities (Rao & Hodgkin, 2002). However, Wei *et al.* (2008) suggested that ecological and geographical factors did not play an important role in the sesame evolution, therefore, the most important factor affecting the current sesame genetic structure was human activities (Laurentin & Karlovsky, 2006). This is similar to previous work on safflower by Khan *et al.* (2009) who reported that the genetic diversity showed no relationship to geographical origin as a consequence from human selection and genetic drift. Overall, the genetic distance among populations may or may not correlate with the geographical distance between them in some species depending on natural and artificial factors involved in shaping the population genetic structure of the species (Stankiewicz *et al.*, 2001). For example, the positive correlation between genetic distance and geographical regions was found in niger populations (Geleta, 2007) but not in safflower (Khan *et al.*, 2009).

4.4 Phenotype evaluation and phenotypic relationship of sesame varieties from different origins (paper III)

Sesame varieties in this experiment were from Northern Vietnam, Southern Vietnam, Cambodia, and Kenya as well as from Tanzania, El Salvador and India. All varieties grew well but with a clear difference in plant development between those varieties originating from El Salvador, Tanzania and Kenya and the others. Those varieties (Fig.4 A, B, C) developed a large biomass with big trifoliate leaves and robust stem and had a much-delayed switch into growing lanceolate leaves and a longer juvenile period before the start of flowering in comparison with Vietnamese or Cambodian varieties. The Vietnamese and Cambodian varieties showed similar growth and development (Fig.4 D, E). The Indian variety much resembled the Cambodian and Vietnamese varieties in development but developed smaller leaves and capsules (Fig.4 F).

The early flowering groups were Vietnamese, Cambodian and Indian varieties, which opened their first flower within 24-31 days after sowing. In contrast, the El Salvadoran and Kenyan varieties showed the first flowering 40 days or more, after sowing, and the Tanzanian variety started flowering 79 days after sowing. With the exception of the Kenyan, Tanzanian, El Salvadoran varieties that only developed a large biomass, this study showed a positive relationship between plant height, number of branches and seed production. These results were in agreement with previous observations by Muhammad & Stephen Dorairaj (1964), Gupta & Gupta (1977) and Pathak & Dixit (1992) who reported a positive relationship between plant height, capsules per plant and seed yield. The present study found that capsules per plant and seeds per capsule directly influenced the seed yield component of sesame. There was a positive relationship between number of seeds per capsule, capsule number per plant and seed yield per plant. Studies by Majumdar *et al.* (1987) and Reddy & Haripriya (1992) also reported a highly significant positive correlation between number of seeds per capsule, number of capsules per plant and seed yield per plant. The Kenyan, Tanzanian and El Salvadoran varieties gave a lower seed yield than varieties from India, Vietnam and Cambodia. Several factors may contribute to this observation, including environmental factors such as temperature (day/night), day length, light intensity, precipitation, altitude and latitude. Photosynthesis is influenced by various biotic and abiotic stresses during grain-filling, therefore, change in photosynthesis capacity is a major limiting factor for yield and all yield components (Beheshti & Fard, 2010).

The cluster analysis for phenotypic relationship between accessions showed that the accessions grouped into two main clusters (Fig. 5). The first group was composed of Kenyan, Tanzanian, El Salvadoran sesame. The second group consisted of Indian, Cambodian and Vietnamese sesame. There were defined groups according to geographic regions as Africa, Asia. Based on

dendrogram of cluster analysis, sesame from Africa (MT2, McBlack, Lungalunga) had a close relationship to the West sesame (El Salvador - Exel) and sesame from India (Indian) has a close association to the Orient sesame including Vietnam (CT, TH) and Cambodia (KPT).

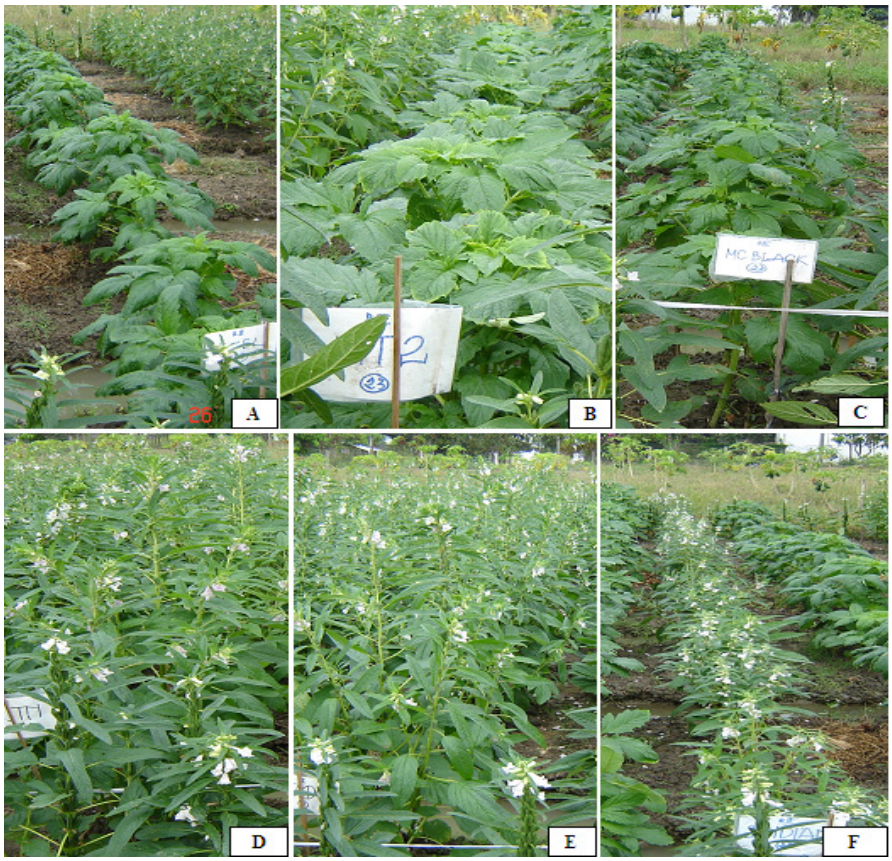

Fig.4 Morphological phenotypes of sesame varieties from different origins, A: El Salvadoran sesame (Exel); B: Tanzanian sesame (MT2); C: Kenyan sesame (McBlack); D: Vietnamese sesame (TN); E: Cambodian sesame (SKD); F: Indian sesame (Indian)

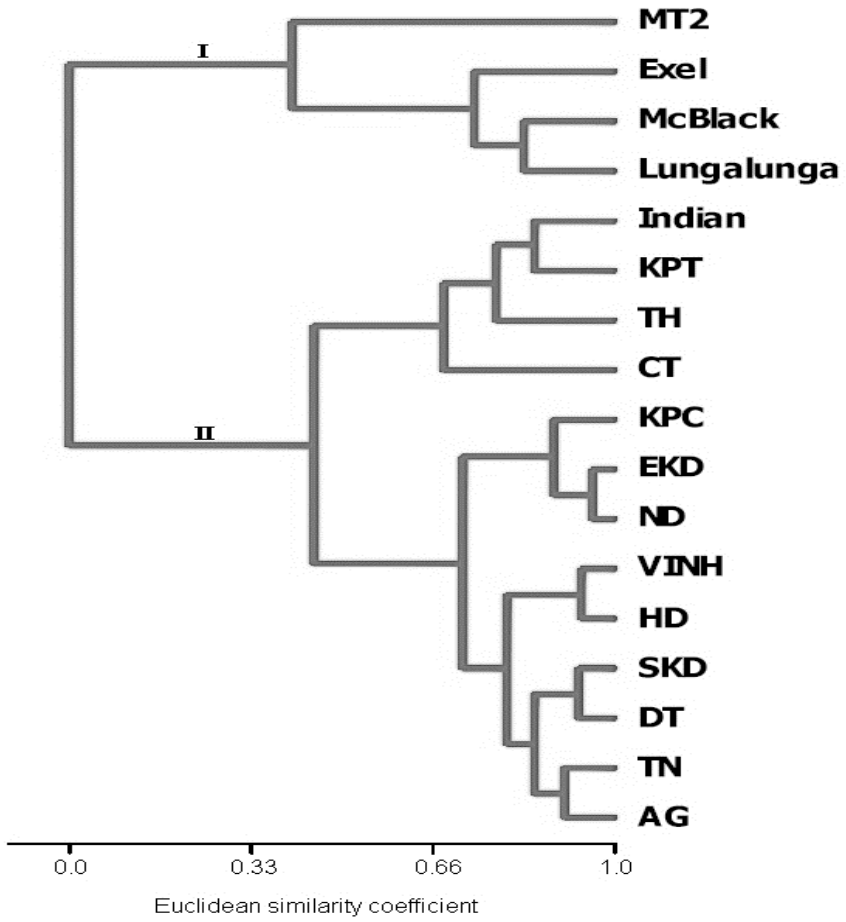

Fig.5 Phenogram of 17 sesame accessions from different origins generated based on morphological data

4.5 Seed oil content, fatty acids and seed yield potential of sesame (paper III, IV)

All Vietnamese and Cambodian sesame accessions had high oil content ranging from 47.2 to 55.6%, with a mean value among all accessions of 50.4% (paper IV). The percentage of the two major fatty acids oleic (18:1) and linoleic (18:2) was 37.65% and 45.67%, respectively, while other fatty acids were found in lower amounts such as palmitic (16:0) 9.33%, stearic (18:0) 5.3% and linolenic (18:3) 0.32%. Based on the average oil content at regional level, the order of sesame accessions from low to high was Southern Vietnam < Northern Vietnam < Cambodia. It should be pointed

out that a few accessions from Northern Vietnam displayed very high oil content, although on average the Northern Vietnam accessions did not show the highest oil content. For example, accession 112TH from Northern Vietnam showed highest oil content (54.8%), whereas the lowest oil content (47.6%) belongs to accession 134CABO from Cambodia. The range in oil content among the accessions in this study (47.2-55.6%) was more or less similar to previous studies such as 47.3-54.2% (Kamaleldin *et al.*, 1992), 42.2-52.1% (Eltinay *et al.*, 1976), 40.4-59.8% (Yermanos *et al.*, 1972) and 53.2-58.5% (Azeez & Morakinyo, 2011).

Seed yield traits determine the seed production in sesame cultivation. When a number of accessions from Southern Vietnam, Northern Vietnam and Cambodia were investigated in the field, a few varieties with good seed yield potential were identified, each having seed production reaching over 1 ton ha^{-1} (paper III). Accessions with high seed yield included those from Cambodia (SKD, EKD, and KPC) and some varieties from Southern Vietnam (AG, DT and TN). The results in paper III showed high seed yield potential in several varieties from Cambodia and Southern Vietnam that could be used for future breeding activities to improve seed yield. When estimating the production from the different types of sesame in cultivation, the variety is of course the number one important factor, and followed by the cultivation technique used. Besides providing high-yielding varieties to the Vietnamese farmers there is also a need for transferring information on improved sesame cultivation techniques with regards to irrigation, fertilizer and plant density. Such combined efforts could give positive result on the seed yield potential of the sesame grown in this region as also reported by others (Ucan *et al.*, 2007; Langham *et al.*, 2010; Ucan & Killi, 2010).

4.6 Development of microsatellite markers for sesame (*Sesamum indicum* L.)

Microsatellite markers have been developed and widely used in analysis of genetic diversity of crop species. All SSR primers developed for this study were evaluated for amplification using eight sesame accessions. The amplified products were separated on 1.5% agarose gel and Hyres 52S-denaturing polyacrylamide gel (Gelcampany GmbH-code nr1001-28). For accurate measurement of fragment size of amplified products, 12 primers were labeled with four different fluorescent dyes (6-FAM, NED, VIC, PET). The PCRs were carried out using Eppendorf AG 5345-016209 thermo-cyclers (Germany). PCR products were separated by capillary electrophoresis on the ABI 3100 genetic analyzer (Applied Biosystems). The Peak scanner software v1.0 (Applied Biosystems) was used to determine the

size of the peaks using the internal Genescan-500 LIZ size standard and sizing default PP analysis method. Out of the seven microsatellite primer-pairs used to generate the final data, three primer-pairs produced monomorphic microsatellites whereas the remaining primer-pairs amplified polymorphic loci (Table 2). Thus, four polymorphic primers were used for genetic diversity analyses of the 22 sesame accessions, comprising accessions from Vietnam and Cambodia. The seven primer-pairs amplified a total of 11 alleles, and the number of alleles per polymorphic loci was only two. The difference in allele size in polymorphic loci ranged from 161-422 bp. The expected heterozygosity (*He*), observed heterozygosity (*Ho*), and genetic structures (*Fst*) were estimated using Popgene version 1.31 (Yeh & Boyle, 1997) (Table 2). The genetic variation among the 22 sesame accessions was estimated based on heterozygosity. Expected heterozygosity (*He*) ranged from 0.12 to 0.48 with a mean 0.29 and the mean observed heterozygosity (*Ho*) was 0.15 with a range of 0.00–0.30. The highest observed heterozygosity was at locus SI-ssr01 with a value of 0.30, while the lowest was 0.0 at locus SI-ssr16. The average population differentiation (*Fst*) was 0.40 suggesting that 40% of total genetic variation was found within accessions. In average, the percent polymorphism, number of alleles, *He* and *Ho* per accession were 48.9%, 5.9, 0.18 and 0.14 (Table 3).

The SSR markers have become one of the most widely used molecular markers in recent years and are preferable over many other markers. It was used to analyze the genetic diversity in sesame, and revealed high genetic variation in Korean sesame (Dixit *et al.*, 2005), and Chinese and some exotic sesame (Wei *et al.*, 2008). In the present study, 4 SSR markers found moderate level of genetic variation among 22 sesame accessions and were able to discriminate between the different sesame genotypes. Powell *et al.* (1996) reported that SSR markers give a good discrimination between closely related individuals in some cases even when only a few loci are applied. Therefore, SSR markers would be useful molecular markers in sesame genetic diversity analyses and in marker assisted breeding programs.

Table 2 Description of microsatellite loci for sesame (*Sesamum indicum* L.) accessions

Primer name	Sequence 5'-3'	Repeat motifs	*Ta* (°C)	Expected size bp	Observed size bp	*Na*	*Ho*	*He*	*Fst*
SI-ssr01	F-agcaagagacaagatgacga	(CCG)4	60	168	161-170	2	0.30	0.26	0.19
	R-tggtggatgagcaggtaata								
SI-ssr19	F-tccattgagaactaccagca	(AT)7(ATGT)6	61	350	359-401	2	0.22	0.48	0.36
	R-gccacctgaaaatctgaaaa								
SI-ssr30	F-gattgcagaaattgacacca	(CT)8	61	239	242-244	2	0.06	0.32	0.64
	R-cactaggcgaagaattcaaga								
SI-ssr16	F-cgaaactctcatctacccaag	(CT)9	60	386	388-422	2	0.00	0.12	0.36
	R-cagctcgtacttcccatgta								
SI-ssr11	F-ttcctctccctttctcctt	(CT)5(ATT)7	61	500	504	1	---	---	---
	R-cctgcattccacaatttcat								
SI-ssr21	F-ttccatacccattttcacct	(GTTT)4	60	393	390	1	---	---	---
	R-gttcgctttcttgaccattc								
SI-ssr29	F-tacaggcggagagagagatt	(TC)6	60	388	402	1	---	---	---
	R-ctccacgcacacaatagg								
Mean							0.15	0.29	0.40

Ta: annealing temperature; *Na*: number of alleles
He: expected heterozygosity; *Ho*: observed heterozygosity; *Fst*: Wright fixation index – F-statistics

Table 3 SSR markers based estimates of different genetic diversity parameters for 22 sesame accessions

ID	Acc	*%P*	*Na*	*Ho*	*He*
1	162ND	50.0	6	0.08	0.16
2	112TH	50.0	6	0.00	0.19
3	191HD	25.0	5	0.00	0.12
4	121NB	50.0	6	0.00	0.19
5	29VINH	50.0	6	0.18	0.25
6	66NL	50.0	6	0.25	0.22
7	8HN	50.0	6	0.30	0.22
8	77DC	75.0	7	0.21	0.19
9	143TB	0.0	4	0.00	0.00
10	39TN	75.0	7	0.00	0.24
11	65AG	0.0	4	0.00	0.00
12	52CT	25.0	5	0.00	0.06
13	9GL	75.0	7	0.43	0.32
14	12TN	75.0	7	0.23	0.25
15	26CT	75.0	7	0.32	0.30
16	6DT	75.0	7	0.39	0.26
17	26HUE	50.0	6	0.08	0.15
18	3KG	50.0	6	0.25	0.22
19	107CABO	25.0	5	0.08	0.10
20	54CABO	25.0	5	0.00	0.21
21	134CABO	50.0	6	0.35	0.24
22	25CABO	75.0	7	0.00	0.13
	Mean	48.9	5.9	0.14	0.18

%P: Percentage of polymorphism; *Na*: number of alleles
He: expected heterozygosity; *Ho*: observed heterozygosity

5 Conclusion, recommendation and future prospects

5.1 Conclusion

A high genetic diversity exists in the sesame germplasm collected in Vietnam and Cambodia suggesting a high potential of the crop in this region for genetic improvement. A high variation was also found among the populations growing in the different geographical regions and localities that would be useful for plant breeding and conservation purposes.

The agro-morphological and molecular markers were complementary in assessing the extent of genetic variation in sesame and the combined use of these methods helps to better understand the genetic diversity and relationship within and among sesame populations.

There was a weak association between genetic variation of sesame accessions and their ecological regions of origin and the most likely important factor affecting the genetic structure of sesame in this region is human activities. Some sesame accessions collected from different eco-geographic regions of Vietnam and Cambodia may have actually originated from the same population.

Sesame accessions collected in Vietnam and Cambodia showed significant variation in oil content, and the oil content had a strong positive correlation with oleic acid.

There was a positive correlation between capsules per plant and seed yield per plant and, therefore, these traits should be considered in the seed yield improvement programs in sesame. The present study also found that day/night temperature, day length and light intensity are among

environmental factors affecting seed yield in sesame. For example, sesame varieties from Africa and El Salvador were poor in seed yield when grown in the Mekong delta, Vietnam.

The results achieved from morphology and oil content analyses are useful as background information for sesame breeding program. Future breeding strategies to improve sesame need to exploit the results of this study to increase seed yield potential and other desirable traits.

5.2 Recommendations and future prospects

A lack of information about genetic diversity has been a barrier to improve sesame in Vietnam and Cambodia. The results of this study provided a better understanding of sesame populations in different ecological regions. Therefore, the wise use of results obtained in this study would facilitate the improvement of sesame though breeding and the *in situ* and *ex situ* conservation of sesame genetic resources in this region.

Desirable genotypes in terms of seed yield and oil content have been identified in this study. The use of these genotypes in the breeding program would lead to improved sesame varieties with a high seed yield as well as high oil content.

Several varieties from agro-morphological evaluation were found to have high seed yield potential, and the use of these varieties would lead to an increased seed production in Vietnam and Cambodia. In Vietnam, sesame is grown in the field with low agricultural inputs, which resulted in more pests, diseases, and a low seed yield in this crop. Seed yield in sesame is influenced by different factors such as planting dates, weather patterns, relative humidity, irrigation, fertilizers, and farmers' practices. Cultivation of sesame under good growing conditions would result in a relatively high seed yield. Therefore, in parallel to developing high yielding varieties the cultivation techniques need to be improved.

Vietnam is endowed with favorable ecologies for sesame cultivation. However, the seed yield of sesame is low (500 kg ha^{-1}) mainly because of the varieties used. Hence, a long term strategy of sesame cultivation in Vietnam and Cambodia should include further genetic diversity analyses and the development of permanent genomic resources. These resources can be used to identify desirable parental materials that will be used to improve the

existing varieties with various desirable traits, including seed yield and oil content. Overall, for increased and stable production of sesame in the region, the "four homes": Governments, Research Institutes/Universities, Companies and Farmers, need to be involved and work hand in hand.

References

Ahmad, M., Khan, M.A., Zafar, M. & Sultana, S. (2010). Environment-friendly Renewable Energy from Sesame Biodiesel. *Energy Sources Part a-Recovery Utilization and Environmental Effects* 32(2), 189-196.

Annussek, G. (2001). Sesame oil. In: *Gale encyclopedia of alternative medicine*. Gale Group and Looksmart.

Arriel, N.H., Di Mauro, A.O., Arriel, E.F., Uneda-Trevisoli, S.H., Costa, M.M., Barbaro, I.M. & Muniz, F.R. (2007). Genetic divergence in sesame based on morphological and agronomic traits. *Crop Breeding and Applied Biotechnology* 7(3), 253-261.

ASGA *Sesame Markets*. [online] (2011) Available from: http://www.sesamegrowers.org/usesofsesame.htm. [Accessed 27/01/2011].

Ashri, A. (1998). Sesame breeding. *Plant Breed. rev.* 16, 179-228.

Azeez, M.A. & Morakinyo, J.A. (2011). Genetic diversity of fatty acids in sesame and its relatives in Nigeria. *European Journal of Lipid Science and Technology* 113(2), 238-244.

Banerjee, P.P. & Kole, P.C. (2010). Heterosis, inbreeding depression and their relationship with genetic divergence in sesame (*Sesamum indicum* L.). *Acta Agronomica Hungarica* 58(3), 313-321.

Baydar, H., Marquard, R. & Turgut, I. (1999). Pure line selection for improved yield, oil content and different fatty acid composition of sesame, *Sesamum indicum*. *Plant Breed.* 118(5), 462-464.

Bedigian, D. (1981). Origin, diversity, exploration and collection of sesame. *In* Sesame: Status and Improvement. *FAO Plant Production and Protection Paper* 29, Pages 164-169.

Bedigian, D. (2004). History and lore of sesame in Southwest Asia. *Econ. Bot.* 58(3), 329-353.

Bedigian, D. & Harlan, J.R. (1986). Evidence for cultivation of sesame in the ancient world. *Econ. Bot.* 40(2), 137-154.

Beheshti, A.R. & Fard, B.B. (2010). Dry matter accumulation and remobilization in grain sorghum genotypes (*Sorghum bicolor* L.

Moench) under drought stress. *Australian Journal of Crop Science* 4(3), 185-189.

Bhat, K.V., Babrekar, P.P. & Lakhanpaul, S. (1999). Study of genetic diversity in Indian and exotic sesame (*Sesamum indicum* L.) germplasm using random amplified polymorphic DNA (RAPD) markers. *Euphytica* 110(1), 21-33.

Bisht, I.S., Mahajan, R.K., Loknathan, T.R. & Agrawal, R.C. (1998). Diversity in Indian sesame collection and stratification of germplasm accessions in different diversity groups. *Genet. Resour. Crop Evol.* 45(4), 325-335.

Black-Samuelsson, S., Eriksson, G., Gustafsson, L. & Gustafsson, P. (1997). RAPD and morphological analysis of the rare plant species *Vicia pisiformis* (Fabaceae). *Biological Journal of the Linnean Society* 61(3), 325-343.

Carlsson, A.S., Chanana, N.P., Gudu, S., Suh, M.C. & Were, B.A. (2008). Sesame. In: Kole, C.*, et al.* (Eds.) *Compendium of transgenic crop plant - Transgenic oilseed crops*. pp. 227-246. Texas, USA: Wiley Blackwell; 2. ISBN 978-1-405-16924-0.

Cooney, R.V., Custer, L.J., Okinaka, L. & Franke, A.A. (2001). Effects of dietary sesame seeds on plasma tocopherol levels. *Nutrition and Cancer-an International Journal* 39(1), 66-71.

Choudhary, C.S., Prasad, S.M. & Kudada, N. (2006). Evaluation of sesame genotypes for resistance against *Corynespora* blight. *Journal of Research, Birsa Agricultural University* 18(1), 149-151.

Chung, J., Lee, J. & Choe, E. (2004). Oxidative stability of soybean and sesame oil mixture during frying of flour dough. *J. Food Sci.* 69(7), 574-578.

Dey, S.S., Singh, A.K., Chandel, D. & Behera, T.K. (2006). Genetic diversity of bitter gourd *(Momordica charantia* L.) genotypes revealed by RAPD markers and agronomic traits. *Scientia Horticulturae* 109(1), 21-28.

Dixit, A., Jin, M.H., Chung, J.W., Yu, J.W., Chung, H.K., Ma, K.H., Park, Y.J. & Cho, E.G. (2005). Development of polymorphic microsatellite markers in sesame (*Sesamum indicum* L.). *Molecular Ecology Notes* 5(4), 736-738.

Durai, S.R., Saravanan, S., Pandiyan, K.S. & Sevaguperumal (2010). Investigation on hybrid vigour for yield traits in sesame (*Sesamum indicum* L.). *Research on Crops* 11(2), 476-478.

Eltinay, A.H., Khattab, A.H. & Khidir, M.O. (1976). Protein and oil compositions of sesame seed. *J. Am. Oil Chem. Soc.* 53(10), 648-653.

Ercan, A.G., Taskin, M. & Turgut, K. (2004). Analysis of genetic diversity in Turkish sesame (*Sesamum indicum* L.) populations using RAPD markers. *Genet. Resour. Crop Evol.* 51(6), 599-607.

Esquinas-Alcazar, J. (2005). Protecting crop genetic diversity for food security: political, ethical and technical challenges. *Nat Rev Genet* 6(12), 946-953.

Faircloth, B.C. (2008). Msatcommander: detection of microsatellite repeat arrays and automated, locus-specific primer design. *Molecular Ecology Resources* 8(1), 92-94.

FAOSTAT *Food and agriculture organization of the united nations*. [online] (2008) Available from: http://faostat.fao.org/site/567/default.aspx#ancor. [Accessed 27/01/2011].

Furat, S. & Uzun, B. (2010). The use of agro-morphological characters for the assessment of genetic diversity in sesame (*Sesamum indicum* L.). *Plant Omics* 3(3), 85-91.

Gaikwad, K.B., Lal, J.P. & Bhakre, R.L. (2009). Combining ability and heterosis for seed yield and related traits in sesame (*Sesamum indicum* L.). *Annals of Plant Physiology* 23(1), 57-61.

Ganesh, S.K. & Thangavelu, S. (1995). Genetic divergence in sesame (*Sesamum indicum*). *Mad. Agric. J.* 82(4), 263-265.

Geleta, M. (2007). *Genetic diversity, phylogenetics and molecular systematics of Guizotia Cass. (Asteraceae)*. Diss. Alnarp:Swedish University of Agricultural Sciences.

Ghafoorunissa Dietary fats/oils and heart diseases. In: Prasad, M.V.R. (Ed.) *Proceedings of Sustainability in oil seeds*, Hyderabad1994. pp. 486-490: Indian Society of Oil Seed Research.

Greene, S.L., Gritsenko, M. & Vandemark, G. (2004). Relating morphologic and RAPD marker variation to collection site environment in wild populations of red clover (*Trifolium pratense* L.). *Genetic Resources and Crop Evolution* 51(6), 643-653.

Gupta, V.K. & Gupta, Y.K. (1977). Variability, interrelationship and path-coefficient analysis for some quantitative characters in Sesame. *Indian Journal of Heredity* 9(1), 31-37.

Haller, H.L., McGovran, E.R., Goodhue, L.D. & Sullivan, W.N. (1942). The synergistic action of sesamin with pyrethrum insecticides. *The Journal of Organic Chemistry* 07(2), 183-184.

Hamrick, J.L. & Godt, M.J.W. (1989). Allozyme diversity in plants. In: Brown A. H. D., Clegg M. T., Kahler A. L. and Wei B. S. (eds), Plant Population Genetics, Breeding and Genetic resources. *Sinauer, Sunderland, Massachusetts*, 43-63.

Hasan, A., Begum, S., Furumoto, T. & Fukui, H. (2000). A new chlorinated red naphthoquinone from roots of *Sesamum indicum*. *Bioscience Biotechnology and Biochemistry* 64(4), 873-874.

Hibasami, H., Fujikawa, T., Takeda, H., Nishibe, S., Satoh, T., Fujisawa, T. & Nakashima, K. (2000). Induction of apoptosis by *Acanthopanax senticosus* HARMS and its component, sesamin in human stomach cancer KATO III cells. *Oncol. Rep.* 7(6), 1213-1216.

Hiltebrandt, V.M. (1932). Sesame (*Sesamum indicum* L.). *Bull. Appl. Bot. Ser. 9.* (2), 3-114.

Jellin, J.M., Gregory, P., Batz, F. & Hitchens, K. (2000). Pharmacist's letter/prescriber's letter natural medicines comprehensive database. 3rd ed. Therapeutic Research Faculty, Stockton, CA. p. 1–1527.

John, P., Tripathi, N.N. & Naveen, K. (2005). Evaluation of sesame germplasm/cultivars for resistance against charcoal rot. *Research on Crops* 6(1), 152-153.

Johns, M.A., Skroch, P.W., Nienhuis, J., Hinrichsen, P., Bascur, G. & MunozSchick, C. (1997). Gene pool classification of common bean landraces from Chile based on RAPD and morphological data. *Crop Science* 37(2), 605-613.

Kamaleldin, A., Yousif, G., Iskander, G.M. & Appelqvist, L.A. (1992). Seed lipids of *Sesamum indicum* L. and related wild-species in Sudan .1. Fatty-acids and Triacylglycerols. *Fett Wissenschaft Technologie-Fat Science Technology* 94(7), 254-259.

Kantety, R.V., Zeng, X.P., Bennetzen, J.L. & Zehr, B.E. (1995). Assessment of genetic diversity in dent and popcorn (*Zea mays* L) inbred lines using inter simple sequence repeat (ISSR) amplification. *Mol. Breed.* 1(4), 365-373.

Kang, C., Kim, S., Lee, S., Mathur, P.N., Hodgkin, T., Zhou, M. & Lee, J. (2006). Selection of a core collection of Korean sesame germplasm by a stepwise clustering method. *Breeding Science* 56(1), 85-91.

Kim, D.H., Zur, G., Danin-Poleg, Y., Lee, S.W., Shim, K.B., Kang, C.W. & Kashi, Y. (2002). Genetic relationships of sesame germplasm collection as revealed by inter-simple sequence repeats. *Plant Breed. rev.* 121(3), 259-262.

Kobayashi, T., Kinoshita, M., Hattori, S., Ogawa, T., Tsuboi, Y., Ishida, M., Ogawa, S. & Saito, H. (1990). Development of the sesame metallic fuel performance code. *Nucl. Technol.* 89(2), 183-193.

Khan, M., von Witzke-Ehbrecht, S., Maass, B. & Becker, H. (2009). Relationships among different geographical groups, agro-morphology, fatty acid composition and RAPD marker diversity in Safflower (*Carthamus tinctorius*). *Genet. Resour. Crop Evol.* 56(1), 19-30.

Langham, D.R., Riney, J., Smith, G. & Wiemers, T. (2008). *Sesame harvest guide* [online]: www.sesaco.net.

Langham, D.R., Riney, J., Smith, G., Wiemers, T., Peeper, D. & Speed, T. (2010). *Sesame producer guide* [online]: www.sesaco.net.

Langham, D.R. & Wiemers, T. (2002). Progress in mechanizing sesame in the US through breeding. In: Janick, J., *et al.* (Eds.) *Trends in new crops and new uses*. pp. 157-173 ASHS Press, Alexandria, VA.

Lattoo, S.K., Rekha, S.D., Khan, S., Bamotra, S., Bhan, M.K., Dhar, A.K. & Gupta, K.K. (2008). Comparative analysis of genetic diversity using molecular and morphometric markers in *Andrographis*

paniculata (Burm. f.) Nees. *Genetic Resources and Crop Evolution* 55(1), 33-43.

Laurentin, H.E. & Karlovsky, P. (2006). Genetic relationship and diversity in a sesame (*Sesamum indicum* L.) germplasm collection using amplified fragment length polymorphism (AFLP). *BMC Genet.* 7, 1-10.

Leduc, V., Moneret-Vautrin, D.A., Tzen, J.T.C., Morisset, M., Guerin, L. & Kanny, G. (2006). Identification of oleosins as major allergens in sesame seed allergic patients. *Allergy* 61(3), 349-356.

Linhart, Y.B. & Grant, M.C. (1996). Evolutionary significance of local genetic differentiation in plants. *Annual Review of Ecology and Systematics* 27, 237-277.

Lynch, M. & Milligan, B.G. (1994). Analysis of population genetic structure with RAPD markers. *Mol. Ecol.* 3(2), 91-99.

Majumdar, S.K., Barik, K.C., Bera, P.S. & Ghosh, D.C. (1987). Path coefficient analysis in sesame (*Sesamum indicum* L.) with varying levels of nitrogen and potassium. *Indian Agriculturist* 31(3), 165-169.

Miyahara, Y., Hibasami, H., Katsuzaki, H., Imai, K. & Komiya, T. (2001). Sesamolin from sesame seed inhibits proliferation by inducing apoptosis in human lymphoid leukemia Molt 4B cells. *Int. J. Mol. Med.* 7(4), 369-371.

Morris, J. (2009). Characterization of sesame (*Sesamum indicum* L.) germplasm regenerated in Georgia, USA. *Genet. Resour. Crop Evol.* 56(7), 925-936.

Morris, J.B. Food, industrial, nutraceutical, and pharmaceutical uses of sesame genetic resources. In: Janick, J.*, et al.* (Eds.) *Proceedings of Trends in new crops and new uses*, Atlanta, Georgia, USA, 10-13 November, 2001 2002. pp. 153-156: Proceedings of the Fifth National Symposium.

Nei, M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89(3), 583-590.

Nupur, M., Bhat, K.V. & Srivastava, P.S. (2010). Variation in fatty acid composition in indian germplasm of sesame. *Journal of the American Oil Chemists' Society: 87 (11) 1263-1269* 87(11), 1263-1269.

Nyanapah, J.O., Ayiecho, P.O. & Nyabundi, J.O. (1995). Evaluation of sesame cultivars for resistance to *Cercospora* leaf spot. *East Africa Agriculture and Forestry Jounal* 60(3), 115-121.

Ojiambo, P.S., Ayiecho, P.O. & Nyabundi, J.O. (1999). Effect of plant age on sesame infection by *Alternaria* leaf spot. *African Crop Science Journal* 7(1), 91-96.

Page, R.D.M. (1996). TreeView: An application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences* 12(4), 357-358.

Parameshwarappa, S.G. & Salimath, P.M. (2010). Studies on combining ability and heterosis for yield and yield components in sesame (*Sesamum indicum* L.). *Green Farming* 3(2), 91-94.

Pathak, H.C. & Dixit, S.K. (1992). Genetic variability and inter-relationship studies in black seeded sesame (*Sesamum indicum* L.). *Madras Agricultural Journal* 79(2), 94-100.

Pathirana, R. (1994). Natural cross-pollination in sesame (*Sesamum indicum* L.). *Plant Breed.* 112(2), 167-170.

Pavlicek, A., Hrda, S. & Flegr, J. (1999). FreeTree-freeware program for construction of phylogenetic trees on the basis of distance data and bootstrap jackknife analysis of the tree robustness. Application in the RAPD analysis of genus Frenkelia. *Folia Biologica* 45(3), 97-99.

Pejic, I., Ajmone-Marsan, P., Morgante, M., Kozumplick, V., Castiglioni, P., Taramino, G. & Motto, M. (1998). Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, RAPDs, SSRs, and AFLPs. *Theor. Appl. Genet.* 97(8), 1248-1255.

Persson, H.A. & Gustavsson, B.A. (2001). The extent of clonality and genetic diversity in lingonberry (*Vaccinium vitis-idaea* L.) revealed by RAPDs and leaf-shape analysis. *Mol. Ecol.* 10(6), 1385-1397.

Powell, W., Morgante, M., Andre, C., Hanafey, M., Vogel, J., Tingey, S. & Rafalski, A. (1996). The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Mol. Breed.* 2(3), 225-238.

Prajapati, N.N., Patel, C.G., Bhatt, A.B., Prajapati, K.P. & Patel, K.M. (2010). Heterosis in sesame (*Sesamum indicum* L.). *International Journal of Agricultural Sciences* 6(1), 91-93.

Pham, D.T., Bui, M.T., Werlemark, G., Bui, C.T., Merker, A. & Carlsson, A.S. (2009). A study of genetic diversity of sesame (*Sesamum indicum* L.) in Vietnam and Cambodia estimated by RAPD markers. *Genet. Resour. Crop. Evol.* 56(5), 679-690.

Pham, D.T., Nguyen, T.T.D., Carlsson, A.S. & Bui, M.T. (2010). Morphological evaluation of sesame (*Sesamum indicum* L.) varieties from different origins. *Australian Journal of Crop Science* 4(7), 498-504.

Rao, V.R. & Hodgkin, T. (2002). Genetic diversity and conservation and utilization of plant genetic resources. *Plant Cell Tissue and Organ Culture* 68(1), 1-19.

Reddy, C.D.R. & Haripriya, S. (1992). Genetic character association and path coefficient analysis in parents and their F1 sesame. *J. Maharashtra Agric. Univ.* 17, 55-57.

Rheenen, H.A.v. (1980). Aspects of natural cross-fertilization in sesame (*Sesamum indicum* L.). *Tropical Agriculture* 57(1), 53-59.

Rohlf, F.J. (2000). NTSYS-pc Numerical Taxonomy and Multivariate Analysis System. In: *Rohlf, F. J. NTSYS-pc: Numerical Taxonomy and*

Multivariate Analysis System. Applied Biostatistics, Exerter Publishing Ltd., New York, USA. Software. ISBN 0-925031-18-6.

Rozen, S. & Skaletsky, H. (2000). Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* 132(3), 365-386.

Russell, J.R., Fuller, J.D., Macaulay, M., Hatz, B.G., Jahoor, A., Powell, W. & Waugh, R. (1997). Direct comparison of levels of genetic variation among barley accessions detected by RFLPs, AFLPs, SSRs and RAPDs. *Theor. Appl. Genet.* 95(4), 714-722.

Salazar, B., Laurentin, H., Davila, M. & Castillo, M.A. (2006). Reliability of the RAPD technique for germplasm analysis of sesame (*Sesamum indicum* L.) from Venezuela. *Interciencia* 31(6), 456-460.

Silme, R.S. & Cagirgan, M.I. (2010). Screening for resistance to *Fusarium* wilt in induced mutants and world collection of sesame under intensive management. *Turkish Journal of Field Crops* 15(1), 89-93.

Stankiewicz, M., Gadamski, G. & Gawronski, S.W. (2001). Genetic variation and phylogenetic relationships of triazine-resistant and triazine-susceptible biotypes of *Solanum nigrum* - analysis using RAPD markers. *Weed Research* 41(4), 287-300.

Steiner, J.J. & de los Santos, G.G. (2001). Adaptive ecology of *Lotus corniculatus* L. genotypes: I. Plant morphology and RAPD maker characterizations. *Crop Science* 41(2), 552-563.

Suja, K.P., Abraham, J.T., Thamizh, S.N., Jayalekshmy, A. & Arumughan, C. (2004). Antioxidant efficacy of sesame cake extract in vegetable oil protection. *Food Chem.* 84(3), 393-400.

Talebi, R., Fayaz, F., Mardi, M., Pirsyedi, S.M. & Naji, A.M. (2008). Genetic relationships among chickpea (*Cicer arietinum*) elite lines based on RAPD and agronomic markers. *International Journal of Agriculture and Biology* 10(3), 301-305.

Tripathi, S.K., Tamrakar, P. & Chatterji, D. (1998). Evaluation of sesamum cultivars for resistance to *Cercospora* leaf spot. *Crop Research (Hisar)* 15(2-3), 265-269.

Ucan, K. & Killi, F. (2010). Effects of different irrigation programs on flower and capsule numbers and shedding percentage of sesame. *Agricultural Water Management* 98(2), 227-233.

Ucan, K., Killi, F., Gencoglan, C. & Merdun, H. (2007). Effect of irrigation frequency and amount on water use efficiency and yield of sesame (*Sesamum indicum* L.) under field conditions. *Field Crop. Res.* 101(3), 249-258.

Uzun, B. & Cagirgan, M.I. (2009). Identification of molecular markers linked to determinate growth habit in sesame. *Euphytica* 166(3), 379-384.

Uzun, B., Lee, D., Donini, P. & Cagirgan, M.I. (2003). Identification of a molecular marker linked to the closed capsule mutant trait in sesame using AFLP. *Plant Breed.* 122(1), 95-97.

Uzun, B., Ulger, S. & Cagirgan, M.I. (2002). Comparison of determinate and indeterminate types of sesame for oil content and fatty acid composition. *Turkish Journal of Agriculture and Forestry* 26(5), 269-274.

Varisai Muhammad, S. & Stephen Dorairaj, M. (1964). Correlation studies in *Sesamum indicum* L.: association between yield and certain yield components in different groups of *Sesamum* based on seed colour. *Madras Agricultural Journal* 51, 73-74.

Warwick, S.I. & Gugel, R.K. (2003). Genetic variation in the *Crambe abyssinica - C. hispanica - C. glabrata* complex. *Genet. Resour. Crop Evol.* 50(3), 291-305.

Wei, L., Zhang, H., Zheng, Y., Guo, W. & Zhang, T. (2008). Development and utilization of EST-derived microstellites in sesame (*Sesamum indicum* L.). *Acta Agronomica Sinica* 34(12), 2077-2084.

Were, B.A., Gudu, S., Onkware, A.O., Carlsson, A.S. & Welander, M. (2006a). In vitro regeneration of sesame (*Sesamum indicum* L.) from seedling cotyledon and hypocotyl explants. *Plant Cell Tissue and Organ Culture* 85(2), 235-239.

Were, B.A., Onkware, A.O., Gudu, S., Welander, M. & Carlsson, A.S. (2006b). Seed oil content and fatty acid composition in East African sesame (*Sesamum indicum* L.) accessions evaluated over 3 years. *Field Crop. Res.* 97(2-3), 254-260.

Williams, J.G.K., Kubelik, A.R., Livak, K.J., Rafalski, J.A. & Tingey, S.V. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res.* 18(22), 6531-6535.

Wolff, N., Cogan, U., Admon, A., Dalal, I., Katz, Y., Hodos, N., Karin, N. & Yannai, S. (2003). Allergy to sesame in humans is associated primarily with IgE antibody to a 14 kDa 2S albumin precursor. *Food and Chemical Toxicology* 41(8), 1165-1174.

Yadav, M., Chaudhary, D., Sainger, M. & Jaiwal, P.K. (2010). *Agrobacterium tumefaciens*-mediated genetic transformation of sesame (*Sesamum indicum* L.). *Plant Cell Tissue and Organ Culture* 103(3), 377-386.

Yeh, F.C. & Boyle, T.J.B. (1997). Population genetic analysis of co-dominant and dominant markers and quantitative traits. *Belgian Journal of Botany* 129(2), 157.

Yermamos, D.M. (1980). Sesame. In: Fehr W.R. and Hadley H.H. (eds), Hybridization of crop plants. Am. Soc . Agron., CSSA, Madison, Wisconsin, USA.

Yermanos, D.M., Saleeb, W., Hemstree.S & Huszar, C.K. (1972). Oil content and composition of seed in world collection of sesame introductions. *J. Am. Oil Chem. Soc.* 49(1), 20-23.

Acknowledgements

These studies were carried out at the Department of Plant Breeding and Biotechnology of the Swedish University of Agricultural Sciences (SLU) and Nong Lam University (NLU), Ho Chi Minh City, Vietnam, with financial support from the Swedish International Development Cooperation Agency, Department for Research Cooperation (SIDA/SAREC). I am grateful for the support from all of these institutions.

I would like to express my sincere thanks to:

Assoc. Prof. Anders Carlsson, my main supervisor for accepting me as your PhD student, for all your ideas, knowledge and experience and for introducing me to the area of research.

Prof. Arnulf Merker†, my co-supervisor for discussion in meetings, valuable support, you were a cheerful, friendly person. I miss you deeply.

Dr. Mulatu Geleta, my co-supervisor for your valuable support in the area of molecular marker studies and constructive discussion, and for reading and correcting my manuscripts.

Dr. Gun Werlemark, my co-supervisor for valuable support, and for reading and correcting my manuscripts

Assoc. Prof. Bui Cach Tuyen, for giving me permission to study and for supporting and encouraging me in my research work when you were the Rector of Nong Lam University and Director of RIBE at NLU

Dr. Bui Minh Tri, my co-supervisor in Vietnam as well as a co-author, for valuable advice, encouragement, enthusiasm and discussions throughout my study and research activities. Thank you very much for your kind support and supervision.

Prof. Sten Stymne for your kind support and help during my study in Sweden and for your inspiring lectures and discussion in the plant lipid field.

Assoc. Prof. Erland Liljeroth for your support and care during my PhD education. Thank you very much for your kind and friendly help.

Prof. Tomas Bryngelsson, head of Department of Plant Breeding and Biotechnology, SLU, Alnarp for signatures, and for accepting me as a PhD student at the Department.

Prof. Margareta Welander for supporting me during the course in Finland and for sharing a taxi between Alnarp and Malmo.

Dr. Knut Wålstedt for your friendly support and willingness to sit down and discuss the area of my research in English with me.

I would like to express my gratitude to Dr. Åsa Grimberg and Dr. Ida Lager for revising my thesis, thank you very much for your friendly help.

Dr. Aman Bonaventure for your helpful correction of my manuscripts, and discussion of the same.

Helén Lindgren and Susanne Hjerdin for kind help during my time in Sweden. You are very kind and you helped me with a lot of things whenever I needed something in the lab, Greenhouse, Biotron and in my apartment

Monica Lotfinia and Jonas Hansson for kind help and assistance in a number of administrative matters during my time in Sweden

Professors, lecturers and assistant lecturers in courses that I have attended during my study for sharing their valuable knowledge

I would like to express my sincere thanks to all that I have known in the Department of Plant Breeding and Biotechnology, Per, Li-Hua, Xueyuan, Eva, Ramune, Salla, Anna, Ann-Sofie, Annelie, Therese, Mariette, Helle, Mirela, Jenny, Maria, Svetlana, Pooja, Yohannes, Bill, Carlos, Isabel, Dickson, Birjan, Sergey, Maksat, Rae, Jing etc, and people in my house in Alnarp, as well as those whom I have played football together with. I would like to express my thanks to the good time I have had with all of you

I would like to express my gratitude to Dr. Trinh Truong Giang, present Rector of Nong Lam University for giving me permission to study and for support and encouragement

Assoc. Prof. Nguyen Thi Lang, CLRRI – Cuu Long Rice Research Institute and Mr. Duyet at Tien Giang College of Agriculture for providing the land necessary for my field trials and for their support during the set up and running of the trials.

Dr. Le Dinh Don, Dr. Dinh Trung Chanh, for their valuable advice, encouragement and readiness to give support.

My colleagues, The, Ha, Hue, Dung, Duong, Huong, Quang, Biet, Truong, Vu etc… at the RIBE of NLU, for sharing experiences in science. Mr Thanh, Mr Duong at the CLRRI, Can Tho for all the help in taking care of my field trials

My friends, Phuong, Thuy, Tu, Thuoc, Cuong, Nam, Hoan etc… for their great support, discussions, meeting and parties during my time in Sweden. Tuan, Truc, Phong, Duy, Lanh, Trung, Nghia, Lam, Long, Kiet, Dong, Hien, Chau, Nam etc… in Vietnam for their friendship, help, sharing and encouragement during the time away home.

My family, my mother, my wife's parents, my brother, sisters and niece for all their support and encouragement during these years, thank you for believing in me.

Finally I would like to express my deepest love to my wife. Thank you for your great support and understanding of my work, for always encouraging me in my studies, and for sharing happiness and sadness with me. Thank you very much and I love you forever.

Popularized summary in Vietnamese

Cây mè (*Sesamum indicum* L.) là cây thân thảo, hàng năm. Và là cây có hàm lượng dầu rất cao trong hạt (50 tới trên 60%). Cây mè rất dễ canh tác, vì thường chịu hạn khá cao, và nó không kén đất nên sinh trưởng tốt trên các loại đất, cũng như các vùng sinh thái khác nhau. Tuy nhiên hiện tại năng suất và sản lượng hạt mè của Việt Nam đang xếp vào trong nhóm những nước sản xuất thấp. Nguyên nhân do một vài yếu tố chủ quan và khách quan, ví dụ như nông dân không mặn mà với cây mè vì lợi nhuận không cao bằng các loại cây trồng khác, hoặc các yếu tố khác như các giống mè có năng suất thấp, không có giống cải tiến, giống mới, và kỹ thuật canh tác còn thấp. Mục tiêu của nghiên cứu này là nghiên cứu tính đa dạng di truyền, đa dạng quần thể của cây mè ở Việt Nam và Campuchia. Dựa vào những thông tin đa dạng di truyền, đa dạng quần thể để cải thiện các giống mè hiện có, cũng như lai tạo giống mới trong tương lai gần, từ đó đáp ứng nhu cầu giống mè cho địa phương và cho cả nước. Từ kết quả nghiên cứu này cho thấy cây mè ở Việt Nam và Campuchia có tính đa dạng khá cao, và hàm lượng dầu trong hạt cũng cao. Có mối liên hệ gần về mặt di truyền của các mẫu giống có nguồn gốc cùng vùng sinh thái nhưng không rõ ràng hay riêng biệt cho từng vùng, và cũng có sự giống nhau giữa các mẫu giống ở các vùng sinh thái khá xa nhau. Đây có thể là kết quả của quá trình trao đổi giống mè trong sản xuất của nông dân giữa các vùng. Qua các thí nghiệm đồng ruộng cho thấy có vài mẫu giống của khu vực miền Nam và Campuchia có tiềm năng năng xuất cao. Đây là thông tin hữu ích để các nhà tạo giống cây trồng có thể sử dụng để cải thiện các giống mè cho năng suất cao, sinh trưởng phù hợp với điều kiện của địa phương, và đáp ứng được nhu cầu canh tác cây mè. Tuy nhiên theo khuyến cáo của chúng tôi, cây mè ở Việt Nam có tiềm năng năng suất cao nếu được đầu tư đúng mức ở khâu canh tác. Vì hiện tại năng suất bình quân vào khoảng 500kg ha^{-1} (thống kê của FAO năm 2009), nhưng trong quá trình điều tra thu thập mẫu, cũng như qua các thí nghiệm đồng ruộng cho thấy có một vài mẫu giống cho năng suất khá cao, nếu được đầu tư tốt ở khâu canh tác thì năng suất hạt sẽ đạt cao hơn. Vì vậy bên cạnh đầu tư giống có tiền năng năng suất cao thì yếu tố canh tác cũng hết sức chú trọng đối với cây mè.