The University of Nottingham

APPLICATION OF GENOMICS AND MOLECULAR GENETICS IN DATE PALM (*PHOENIX DACTYLIFERA* L.)

AL-GHALIYA HUMAID AL-MAMARI

Thesis submitted to the

University of Nottingham for the degree

Doctor of Philosophy

April 2013

School of Biosciences

Division of Plant and Crop Sciences

The University of Nottingham, Sutton Bonington ,Campus

Loughborough, Leicestershire

LE12 5RD

UK

ACKNOWLEDGEMENTS

I would like to express my heartiest gratitude and thanks to the many people who have given their support during the process of this research.

Special thanks go to my supervisor Dr. Sean Mayes, Associate Professor, Crop Genetics, School of Biosciences, University of Nottingham, UK for his valuable guidance, keen interest, help and skilled advices. Thanks also to my second supervisor in Sultanate of Oman, His Excellency Dr. Ishaq Al-Ruqaishi, Under Secretary of Ministry of Agriculture and Fisheries, for his guidance, support, help and encouragement.

Thank to Prof. Dr. Matthew Dickinson for his role as an Internal Assessor, guidance and encouragement. Thank you to Dr. Ahmed Al-Bakri, General Director of Agriculture & Livestock Research Centre, Oman, Dr. Aladdin Hamwieh from the International Centre for Agricultural Research in the Dry Areas (ICARDA), Aleppo-Syria, Dr. Jean-Christophe Pintaud from the Institut de Recherche pour le Développement (IRD), France and Dr. Jamal Khan, Associate Professor, College of Agriculture & Marine Sciences, Sultan Qaboos University, Oman for their kind help, support and valuable guidance.

Thank you to all of the members of my research group and the department who have made working in the Department of Crop Sciences a pleasurable experience. Special thanks go to Dr. Katie Mayes, Mrs. F. Wilkinson and Mrs. L. Staniforth for their technical advices and assistance for laboratory work, Mrs Sue Golds, Mrs. Emma Hooley and Mrs. Kathlen Lawson, Sonoko Mitsui-Angwin and Diane Jones for their administrative help.

I am grateful and thankful to my colleagues and friends: Endah Sri Redjeki, Nariman Ahmed, Presidor Kendabie, Chai Hui Hui, Kidist Bogale Kibret, Dr. Upendra Kumar Devisetty, Dr. Odireleng Ozie Molosiwa and Dr. Gracia Ribas-Vargas.

Thank you also to my colleagues, both researchers, technical and administrative staff at Ministry of Agriculture and Fisheries, Sultanate of

Oman who help me during this project. Special thanks go to Marwa Al-Hinai, Abdullah Al-Jabri, Khamis Al-Rashdy and Juma Al-Siyabi.

I am grateful to the Ministry of Agriculture and Fisheries, Sultanate of Oman for awarding me this scholarship.

Lastly, I feel weak and deficient in vocabulary to find suitable words to express my feelings with tearful eyes for my loving parents whose hands are always raised for prayers which made me successful in every field of my life. Thank you to all my family, who have been exceptionally supportive and understanding throughout the entire PhD process. Thank you especially to my brothers and sisters for their continued moral support.

Thank you all once again and May Allah bless all these people with long, happy and peaceful lives (Amen).

ABSTRACT

Date palm (*Phoenix dactylifera* L.) is a diploid with 18 pairs of chromosomes and an estimated genome size of 658 Mb. It is a dioecious perennial monocot, with a long generation time (a period of 4-5 years until first flowering). Date palm is one of the major fruit crops grown in the Gulf countries and particularly in the Sultanate of Oman. Approximately 250 varieties of date palm are recorded throughout the country with evaluation and characterization based on morphological and reproductive traits (e.g. *fruit color*, *fruit shape* and *fruit weight*). Limited molecular characterization work has been undertaken for date palm germplasm in general and Omani date palm germplasm, in particular. The principal focus of this study was to; investigate the genetic diversity of Omani date palm germplasm and compare it with 'exotic' germplasm, to differentiate between female and male plants at the molecular level and to construct an initial genetic map for date palm.

Samples were taken from eight parents of the available Omani date palm controlled crosses (Khalas 4, Khalas 13 male, Um-Alsela, Khori male, Barni, Naghal, Bahlani male, and Khasab) with 90 date palms from the BC_1 and F_1 populations, from 194 Omani date palm accessions (151 female cultivars and 43 male trees), together with samples from Italy (Sanremo and Bordighera), USDA-ARS, France, Iraq, Libya, Sudan and Iran.

The *F*-statistics analysis showed that the genetic variation between female and male accessions based on random markers was only 2.1%, while within the broader group of Omani female and male accessions the molecular variation was 97%, suggesting that the Omani female and male accessions have little consistent divergence, compared to the large-scale divergence within Omani germplasm, so male palm have been derived from most genetic origins in Oman. Additionally, the Principal Coordinates Analysis (PCA) and bootstrap consensus phenetic tree showed that the Omani accessions were closely related to each other and there was no clear genetic differentiation between female and male cultivars.

A high degree of genetic variation was observed between germplasm from Oman, Italy, USDA-ARS, France, Iraq, Libya, Sudan and Iran as measured by *Fst* (19.7 %). The PCA showed that the Europe-Africa (Italy, France, Libya and Sudan) accessions are distinguished from West-Asia (Oman, Iraq and Iran) accessions and have their own autochthonous origin, a finding which was strongly validated by bootstrap consensus tree test.

A medium density genetic map in date palm was constructed using 53 individuals from BC_1 and 30 individuals from F_1 populations. The BC_1 map consisted of 270 markers (28 SSR and 242 SNP) distributed into 29 linkage groups with total genetic length of 1,486.7 cM, while the F_1 map consisted of 591 markers (21 SSR and 570 SNP) distributed into 30 linkage groups with total genetic length of 2,385.6 cM. A total of 25 combined linkage groups were possible by combining both BC_1 and F_1 maps through common markers.

A sex-link marker locus was developed and found to predict a high level of discrimination between male and female date palms among multiple varieties distributed across the wide range of cultivation, with an accuracy of 100% in the Omani crosses, 96% in the broad Omani material and 86% in the broadest date palm germplasm. This marker was also mapped in both BC_1 and F_1 at 42.8 cM and 4.9 cM in linkage groups 18 and 29, respectively and on combined group 19 at 42.8cM.

Table of contents

List of Tables

List of Figures

ABBREVIATIONS AND SYMBOLS

%	Percentage
'E	East
'N	North
°C	Degree Celsius
µg/µl	Microgram per microliter
µg/mL	Microgram per milliliter
µL	Microliter (s)
2,4-D	2,4-Dichlorophenoxyacetic acid
ABI	Applied Biosystems
AFLP	Amplified fragment length polymorphism
AMOVA	Analysis of Molecular Variance
B.C.	Before Christ's birth
BC_1	Backcrossed population
BLAST	Basic local alignment search tool
bp	Base pair
BSA	Bovine Serum Albumin
cm	Centimetre
cm^2	Square centimetre
DArT	Diversity arrays technology
DNA	Deoxyribonucleic Acid
dNTP	Deoxyribonucleoside triphosphate
E	Ethanol
e.g.	Example
EDTA	Ethylene Diamine Tetra-acetic Acid
F_1	F_1 population
FAO	Food and Agriculture Organisation
g	gravity

ICARDA	International Center for Agricultural Research in the Dry Areas
IRD	Institute de recherché pour le development
kg	Kilogram
km^2	Square kilometre
M	Molar
MAF	Ministry of Agriculture and Fisheries
Mb	Megabyte
Mbp	Megabase pair
$MgCl_2$	Magnesium chloride
min	Minute
mL	Millilitre (s)
mm	Millimetre (s)
mM	Millimolar
MoA	Ministry of Agriculture
mRNA	Messenger Ribonucleic Acid
NaCl	Sodium chloride
NaOH	Sodium hydroxide
ng μL^{-1}	Nanograms per microliter
ng	Nanogram
NGS	Next Generation Sequencing
nmol/µl	Nanomoles per microliter
PC1	First principal component
PC2	Second principal component
PCA	Principal Coordinate Analysis
PCR	Polymerase chain reaction
PIC	polymorphism information content
QTL	Quantitative Trait Loci
RAPD	Randomly amplified polymorphic DNA
RFLP	Restriction fragment length polymorphism

RNase	Ribonuclease
rpm	Revolutions per minute
SDS	Sodium Dodecyl Sulphate
SDW	Sterile Distilled Water
SLS	sample loading solution
SNPs	Single nucleotide polymorphism
SSRs	Simple sequence repeats
Taq	*Thermus aquaticus*
TBE	Tris-borate EDTA buffer
T_m	Temperature
U	Unit
USDA-ARS	United States Department of Agriculture-Agricultural Research Service
UV	ultraviolet light
w/v	weight to volume
X-gal	5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside
λ	Lambda
ΦPT	PhiPT

Chapter 1. GENERAL INTRODUCTION

1.1 The Sultanate of Oman

Oman is an Arab state in southwest Asia located in the southeastern part of the Arabian Peninsula between latitudes 16° 40'N and 26° 20'N and longitude 51° E and 59° 40'E which covers a total land area of 309,500 km^2 (Figure 1.1; MoA, 2011). The land area is composed of valleys and desert (82%), mountain ranges (15%), and the coastal plain (3%). The Sultanate's borders are flanked by the Gulf of Oman in the East, Saudi Arabia and United Arab Emirates in the West and Yemen in the South.

Natural features divide the Sultanate into five administrative divisions or regions: Al-Dakhiliya, Al-Batinah, Al-Wasta, Al-Sharkia, Al-Dhahira, and three governorates: Musandam, Dhofar and Muscat. Musandam Peninsula, an exclave of Oman with 1800 km^2 area projects into the Strait of Hormuz in the north, while in the east is Masirah Island (which is 649 km^2 in area) in the Arabian Sea (Figure 1.1).

Oman is situated at the entrance to the Gulf in the middle of the East-West trade routes, ensuring easy access to markets in the Middle East, India, Southeast Asia, Africa and Europe. Its geographical position gives it access to major shipping routes and allows it to serve as a port and commercial center.

Figure 1.1: The geographical regions of Oman (MoA, 2011)

1.2 The Agricultural land in Oman

According to the survey conducted by MoA (2009) and Worldstat (2007) (http://en.worldstat.info/Asia/Oman/Land) the total agricultural land in Oman is 17,990 km^2 (5.8% of the total area) out of which 600 km^2 is arable (3.3% of the total agricultural land). Most of the agricultural area is located in the Al-Batinah plains in the north of Oman, which represents about 3% of the area of the country (MoA, 2011). Oman has a high level of biodiversity, especially in less arid areas such as Dhofar Governorate, where seasonal rainfall induces diverse and dense vegetation. The fauna and flora in the north of the country

are similar to those in Iran and Pakistan, while semi-arid tropical African types are encountered in the south.

Oman has different indigenous crops that have been considered as strategic crops for food security as they form the basis of the Omani diet. Among these crops are wheat, barley, chickpea, dates, and lime, which are cultivated throughout the country due to their traditional food value. Seasonal fruit crops occupy about 37,082 hectares of the cultivated area, of which 31,365 hectares are used for date palm. The other crops cover 28,017 hectares land of which 10,735 hectares are used for field crops (MoA, 2011). Groundwater resources helped Al-Batinah region to become an area of vital importance to the agricultural economy of the Sultanate of Oman. The important agricultural crops (dates, vegetables, fodders, livestock and fruit trees) in the Al-Batinah area are irrigated from groundwater wells, which support about 28% of Oman's population.

1.3 Climate and water supply

Oman's climate differs from humid in coastal areas, to arid in the interior regions and tropical in the southern parts of the country. The temperature ranges from below zero in mountainous areas like Al-Jabal Al-Akhdar and Jabal Shams and reaches over 50°C during summer in desert areas. Precipitation on the coasts and on the interior plains ranges from 20 to 100 millimeters a year and falls during November – February. Rainfall in the mountains, particularly over Jebel Akhdar, is much higher and may reach 900 millimeters.

Because the plateau of Jebel Akhdar is porous limestone, rainfall seeps

quickly through it, and due to this the vegetation is usually semi-evergreen and has been classified as local center of plant endemism (Ghazanfar, 2003). However, a huge reservoir under the plateau provides springs for low-lying areas making it agriculturally productive.

Dhofar, benefiting from a southwest monsoon between June and September (*Kharif*), receives heavier rainfall (200 – 250 mm) and has constantly running streams, which make this region Oman's most fertile area (MoA, 2011).

Erskine *et al.* (2003) suggested that due to the limited rainfall and the scarcity of fresh water resources in most of the cultivated areas in Oman and in the Arab Peninsula, there is a dependence on irrigation systems from groundwater sources such as afalaj (falaj-singular), springs (oasis) and wells from the provision of small dams spread across the country. In addition, desalinized and treated wastewater form non-conventional sources of water that have been recently used in agriculture (MoA, 2011).

1.4 Date palm and breeding programs in Oman

Date palm (*Phoenix dactylifera* L.) is the major fruit crop grown in the Gulf countries and particularly in the Sultanate of Oman. Approximately 250 varieties of date palm are grown throughout the Sultanate covering an area of 31,365 hectares and this constitutes more than 84% of the total fruit crop area and about 42% of the total agricultural land of Oman (MoA, 2011; MAF, 2005; Al-Khatri, 2004).

El-Kharbotly *et al.* (2006) have provided an overview of the unique germplasm of the Omani date palm that has sustained the country for centuries. The history of date palm in Oman is closely linked to indigenous

farming techniques rooted in traditional wisdom. The Date Palm Research Station and the *ex-situ* Gene Bank in Wadi Quriate in the interior region of Oman were established in 1988 by the Ministry of Agriculture and Fisheries (MAF) (Figure 1.2). The Research Station was mandated to preserve the national heritage of date palm through the conservation of its genetic resources and to carryout research for the improvement and multiplication of this important crop. Additionally in order to help diversify the gene pool and to increase the income of date palm growers, new elite exotic cultivars were introduced from different regional date palm producers and a test breeding program was established. In this breeding program, the male date palm known as KI-96-13 was used in two crosses carried out in 1996. The KI-96-13 germplasm was selected based on its superior characteristics over other F_1 individuals and the synchronized flowering with the mother cultivar, allowing the cross to be made. The KI-96-13 was crossed with its mother Khalas-4, which is known to produce high quality date fruit and a backcross population (BC_1; 53 palms) was developed. It was also used as a male parent to produce an F_1 population (34 palms) with the Um-Assela cultivar, which is well adapted to the conditions in the coastal regions of Oman (high salinity and humidity) but bears low quality dates (El-Kharbotly *et al.*, 1998).

Figure 1.2: The Gene Bank of date palm in Wadi Quriate, Interior region, Oman. Photo by Al-Ghaliya Al-Mamari (2011)

The Sultanate is the eighth largest producer of dates, having around 4% of the total world production. The date palm production season in Oman ranges from May to November each year, the longest season among all date producing countries (FAO, 2005; Al-Yahyai and Al-Khanjari, 2008).

The total number of cultivated date palm trees in Oman has reached around 8 million, producing about 258,000 tons of dates per annum (Anonymous, 2009), of which 64% are used for fresh consumption and 36% for industrial processing (Al-Khatri, 2004). There are around 180 female and 48 male cultivated date varieties, out of which 81 produce yellow fruits, 24 produce red fruits, and the rest produce colours ranging from yellow to red (Figure 1.3; Al-Yahyai and Al-Khanjari, 2008). The most leading date varieties in Oman

based on their production (metric ton) from 2007 to 2009 are; Um Silla, Mabsali, Khasab, Naghal, Fard, Shahel, Khalas, Khinaizi, Madlooki, Barni (MoA, 2009; Table 1).

Figure 1.3: Two female trees of date palm bearing different color fruits

Table 1-1: The leading varieties in Omani date production and production figures (metric ton) from 2007-2009, MoA, (2009)

Rank	2007		2008		2009	
	Cultivar Name	Production	Cultivar Name	Production	Cultivar Name	Production
1	Um Silla	35,465	Um Silla	35,218	Um Silla	27,150
2	Mabsali	29,698	Mabsali	31,175	Khasab	21,961
3	Khasab	27,181	Khasab	27,944	Naghal	20,163
4	Naghal	25,069	Naghal	24,639	Mabsali	16,877
5	Fardh	18,956	Fardh	20,482	Shahel	16,516
6	Shahel	12,258	Khalas	12,658	Fardh	14,996
7	Khalas	12,134	Shahel	12,602	Khalas	14,166
8	Khunaizi	11,135	Khunaizi	11,264	Khunaizi	13,901
9	Madlooki	4,896	Madlooki	5,152	Jabri	6,425
10	Barni	4,852	Barni	5,056	Madlooki	4,171
Total (metric ton)		181,644		186,190		152,155

1.5 Biotic and abiotic threats to date palm

Several biotic (disease and pest) and abiotic (drought and salinity) factors have been found to be limiting date agricultural production in Oman for the last couple of decades. This is in addition to the effects caused by climatic change due to global warming. The high concentration of salt in irrigation water, pests and also diseases have had a direct impact on the cost of production, making the crop less popular with the government which could ultimately lead to a reduction in production.

Date palm, like any other crop, is affected by many pests and diseases, resulting in poor growth and yield, both quantitatively and qualitatively (Al-Khatri, 2004). The most destructive pests of date palm in Oman are the 'dubas' bug *Ommatissus lybicus* DeBergevin, the red palm weevil (RPW) *Rhynchophorus ferrugineus* Olivier and the lesser date moth (LDM) *Batrachedra amydraula*. The severity of infestation varies with cultivars, geographic location, climate, and cultural practices. Although currently of minor importance, many fungal pathogens have also been reported associated with date palm (Al-Saadi *et al.*, 2012).

Chapter 2. LITERATURE REVIEW

2.1 The genus *Phoenix*

The genus *Phoenix* is belonging to the palm family (Palmae or Arecaceae, subfamily Coryphoideae, tribe Phoeniceae) (Terral *et al.*, 2011), and comprises14 different species in which the *Phoenix dactylifera* L. (date palm) is present and is the major palm used for agriculture (Zaid, 2002; Pintaud *et al.*, 2010; Jain *et al.*, 2011; Terral *et al.*, 2011). The *Phoenix* species are distributed in the Old World subtropics and tropics from the Canary and Cape Verde islands in the Atlantic Ocean, throughout Africa, Madagascar and Asia, reaching Sumatra, Taiwan and the Philippines in the East (Pintaud *et al.*, 2010; Figure 2.1). The main centre of diversity of the genus *Phoenix* was found to be span from India to Indochina where eight species are found (Pintaud *et al.*, 2010). Some of *Phoenix* species are used as a source of sugar, e.g., *P. sylvestris* (L) Roxb and are mostly cultivated in the countries like India and Pakistan, while *P. canariensis* Chabeaud is cultivated in Cape Verde and the Canary Islands as an ornamental.

Different hypotheses have proposed the relationships between *P. dactylifera* and one or more other *Phoenix* species such as *P. canariensis* (Canary Islands), *P. atlantica* (Cape Verde), *Phoenix reclinata* Jacq. (sub-Saharan Africa and south-western Arabia) and *P. sylvestris* (India, Pakistan) (Terral *et al.*, 2011). Terral *et al.* (2011) have reported that the date palm is an inter-fertile with all these species where hybrids occur, therefore, the cultivated date palm is highly likely to be domesticated from one of these species or be the result of hybridization between two or several of them. However, recent genetic data suggested that the cultivated date palm derives from wild

populations of *P. dactylifera* (Pintaud *et al.*, 2010). In addition, the identification of wild populations could be complicated due to the possible existence of gene flow from the cultivated pool towards the wild one.

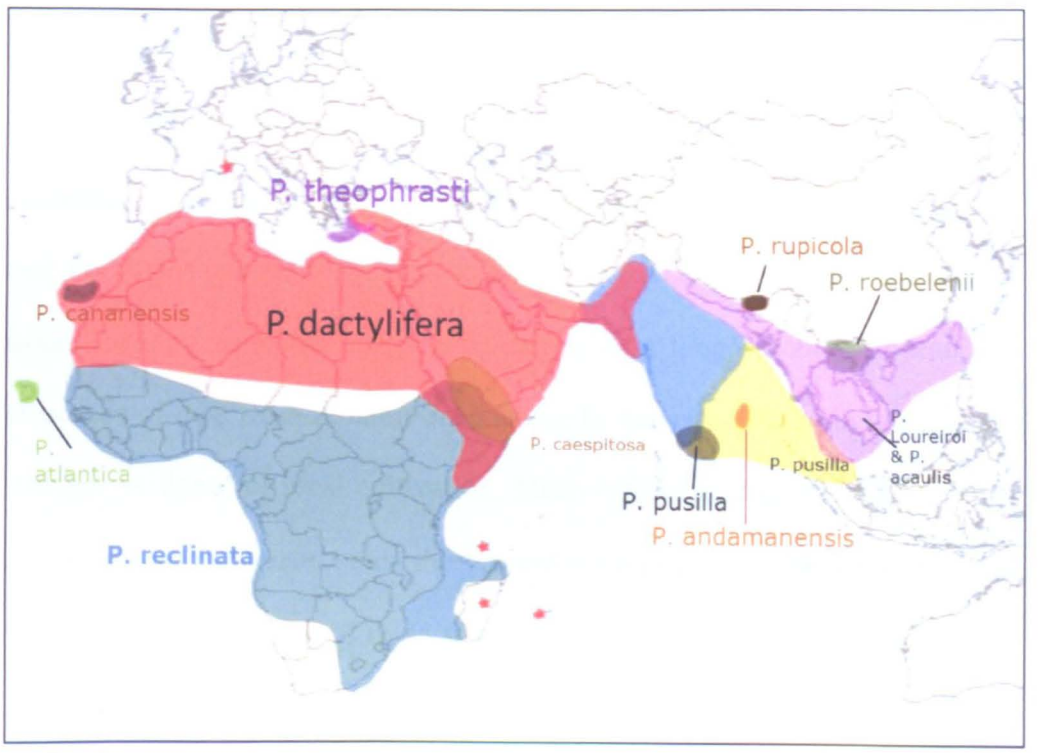


Figure 2.1: Distribution areas of *Phoenix dactylifera* L. and other *Phoenix* species, Gros-Balthazard *et al.* (In press).

2.2 The origin of date palm (*Phoenix dactylifera* L.)

The date palm is one of the most ancient and treasured trees in the Arab world. It has been the main wealth of people in past generations, the fruit serving as a source of daily nourishment, with the branches and the tree trunk proving valuable material in the construction of homes and other household materials. All Arabs including the Omani people have historical, cultural and emotional attachments with date palm. The product of the date palm has always been regarded as a luxury not only due to the challenges of growing the fruit but

also because of its delicacy and high nutritional and economic value in the world market.

The origins of domestication and history of the cultivated date palm are unclear; however, charred and mineralized seeds dating back to the 6th millennium B.C, as were recovered from some archaeological sites in the Persian Gulf, although the discovery of seeds itself does not constitute an evidence of date palm cultivation or its presence locally (Terral *et al.*, 2011). One important feature of the date as an edible fruit is that it can be dried and stored for long time, therefore can be carried over long distances. Indeed, the presences of date palm remains such as seeds, leaves and stipe fragments are stronger evidence of local cultivation. Many bioarchaeological collections of date palm are confirmed since the 5th millennium B.C. in Mesopotamia and since the 4th millennium B.C. in south-eastern Iran and the Oman Peninsula. In addition, the archaeological data suggested that a centre for date palm domestication was located in the Middle East and was supported by many authors, however, other authors proposed a north African, a tropical African, or an Indian origin (Terral *et al.*, 2011).

According to Mahmoudi *et al.* (2008) the domestication of the date palm occurred over 5000 years B.C. and it then spread from Iraq to Iran, India and Pakistan. Overall, the origin of date palm stills remains unclear and whether there was a single origin and/or domestication event (as far as the date palm is domesticated) or a number of independent origins. In terms of religious verses the date palm is mentioned over 25 times in the Holy Book of the Quran while according to Jewish beliefs it shares a holy place along with six other seed plants.

In Oman, since ancient times the date palm has been considered as a symbol of a proud heritage and culture since it represents the wealth of past generations not only for its fruit but also for its tree trunk and leaves that provided shelter from the scorching sun for dwellers and Bedouins (Al-Moharbi, 2011).

2.3 Date palm (*Phoenix dactylifera* L.)

The date palm is dioecious, having separate male and female plants. They can be easily grown from seed, but only 50% of seedlings will be female and hence fruit bearing. In addition, female plants originating from seedling usually produce late fruits of variable and generally inferior quality compared to established clonal palms (Zaid *et al.*, 2002; Chao and Krueger, 2007). Most commercial plantations thus use cuttings of heavily cropping cultivars. Plants grown from cuttings will fruit 2–3 years earlier than seedling plants. The date palm is the tallest of the *Phoenix* group and can grow up to 30 meters depending on the soil and climatic conditions. Additionally the date palm produces a single stem that ranges from 10 to 30 cm in width. Approximately 10 to 12 inflorecences are developed during the winter period in the axils of the leaf positioned immediately below growing point. The leaves of the date palm can grow up to 5 meters long with an individual leaf life span of 4 – 7 years, depending on the weather conditions, water availability or salinity of the ground water. The leaflets also have hard sharp points at their tips in order to protect the fruit from animal predation (Figure 2.2).

The date palm fruits are sweet in taste containing more than 50% sugar by weight, consisting of mainly glucose, fructose and sucrose. There are also

other edible species in the genus *Phoenix* such as *P. atlantica* A., and *P. sylvestris* Roxb. However, due to the higher concentration of sugars in *P. dactylifera* L. it remains the only cultivar with economic relevance. (Source: Naturland 2002).

The date palm fruit is a berry type. Each berry contains a single seed. Each seed has a hard endosperm which results in good preservation of the dried seeds in archaeological contexts (Terral *et al.*, 2011). The date cultivars can be classified into: soft, semi-dry or dry, based on the texture of the fruit under normal ripening conditions and also depending upon the time of harvest and associated water content (Chao and Krueger, 2007; Elshibli, 2009). Date palm cultivars are also grouped as early; mid-season or late according to the length of time needed to produce mature fruit (Jain *et al.*, 2011).

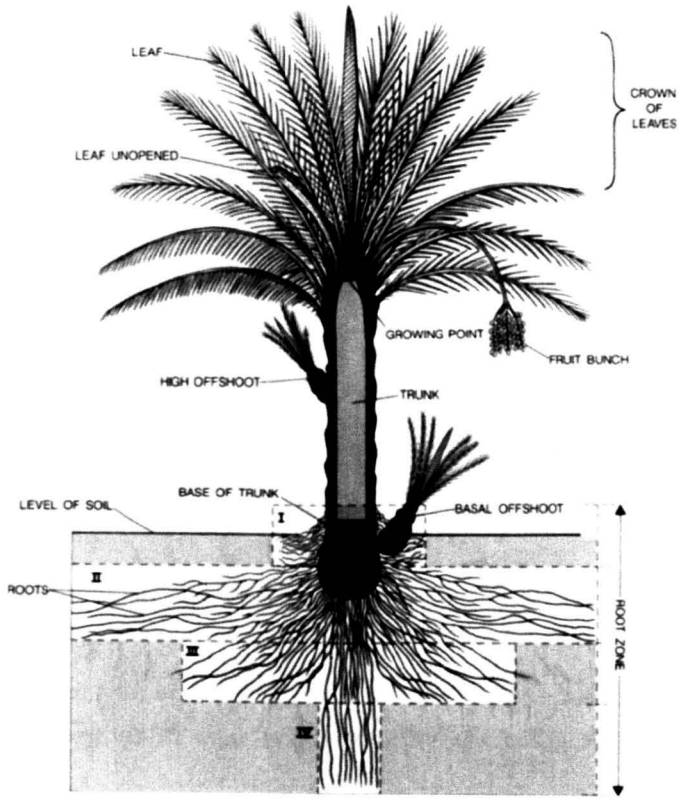


Figure 2.2: A diagram illustrating the structure of the date palm, (Chao and Krueger, 2007).

The date fruit pass through different stages during the ripening process (Figure 2.3). When the fruit are very young they are termed 'Kemri', which are also characterized by a hard texture, green colour and a high cell density. This stage is followed by 'Khala' where cell multiplication continues, the fruit become larger in size, and starch starts to accumulate. The fruit start to mature and the colour changes from green to yellow/red in the 'Beser' stage. In the 'Rutab' stage the fruit become half ripened and the colour changes steadily from yellow/red to dark brown or black as sugars accumulate. The fruit become fully ripened in last stage 'Tamar' and contains high concentrations of sugars, mainly glucose and fructose (reducing sugar) and sucrose (non-reducing sugar) (Yin *et al.*, 2012; Elshibli, 2009).

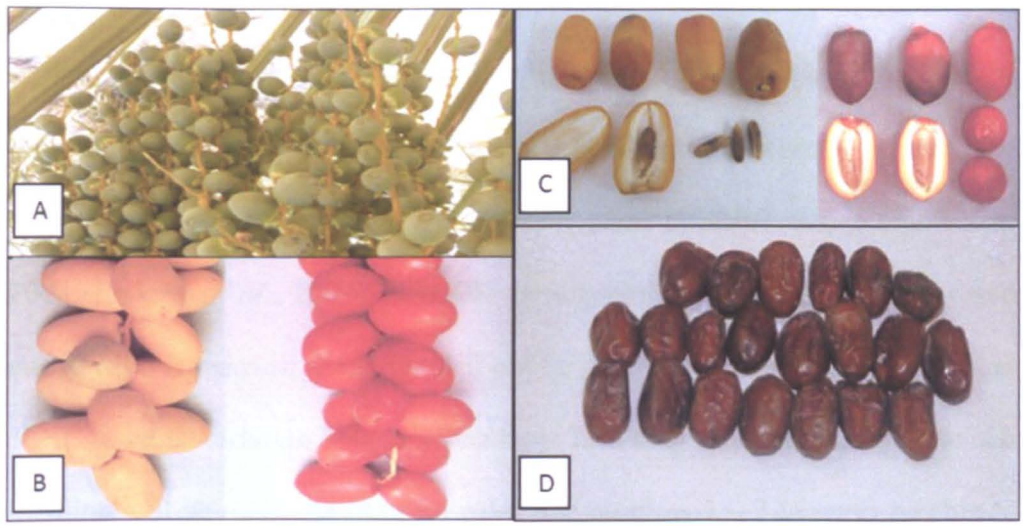


Figure 2.3: Date palm fruit at different stages of ripening: (A) 'Kemri' stage. (B) 'Beser' stage. (C) 'Rutab' stage. (D) 'Tamar' stage. Photo by Al-Ghaliya Al-Mamari.

Date palm trees can remain economically profitable for a period of 50 years, but will continue to produce fruits until the age of 100 or even older. Date palms are frequently cut down when they are about 45 feet tall because of the difficulty in climbing them to harvest the dates or when their productivity declines and they become more susceptible to pests, diseases and blow-down (Jain *et al.*, 2011; Chao and Krueger, 2007). The date palm is highly prized for its ability to adapt and endure long summers in which water is scarce. It also has the ability to thrive well in the desert, a feature that is achieved by obtaining water from underground sources. Date palm is particularly important in agriculture, as it has the ability to provide a microclimate in which other subsistence crops can be cultivated (Alhammadi and Kurup, 2012).

Date palm is propagated by seed, offshoots and tissue culture techniques. Separating offshoots is the most commonly used method. However, the offshoots are produced in limited numbers from the axillary buds on the trunk

of mother plant near the soil surface (Chao and Krueger, 2007; Al-Ruqaishi *et al.*, 2008; El-Kharbotly *et al.*, 1998; Abdulla and Gamal, 2010). Multiplication of date palms by using the offshoots will maintain the genetic identity of the date palm cultivars, therefore producing fruit which are expected to have the same quality and uniformity as the parent plant (Elshibli and Korpelainen, 2008; Erskine *et al.*, 2004). Whereas propagation by seedling is rarely used due to several reasons; seedlings will not be genetically identical to the mother plant, there is variation between seedlings in a single bunch and 50% of the seedling will be male. As the date palm tree requires 5 – 7 years to be able to identify the sex, seedlings will be field planted before the male palms can be identified and discarded. While a proportion of male palms are required in a plantation for pollination of the female palms, it is far below 50%. These palms are essentially unproductive as they do not yield dates. The main advantage of seed method that it is simple in practice and also enlarge the date palm genetic diversity. In some countries the number of date palm trees originating from natural hybrids is important like Egypt and Morocco (Jain *et al.*, 2011). Currently, tissue culture propagation is widely used for large-scale production of true-to-type plantlets from a single elite palm (Chao and Krueger, 2007; Al-Ruqaishi, 2006).

2.4 Date palm biodiversity

Biodiversity or biological diversity refers to all the variety of life that can be found on Earth (plants, animals, and micro-organisms) and includes variation at all levels of biological organization from genes to species to ecosystems. Genetic, organismal and ecological diversity are three elements of biodiversity. Each of these elements has its own components that can be

arranged in a hierarchical order that normally goes from singular living forms to species (Gaston & Spicer 2004; Frankel *et al.*, 1995; Wahid *et al.*, 2004; Elshibli, 2009).

Plant biodiversity is represented by phenotypic and genetic diversity in which these represent the most important aspects for selection, conservation and improvement of a particular species (Elhoumaize *et al.*, 2002). Genetic diversity is the variation in heritable material (whether the variation leads to a phenotypic difference or not) that is found within and between individuals or populations of plant species. While phenotypic diversity refers to the interaction between genetic and environmental variation which leads to a measurable trait.

Date palm is known to have high biodiversity, with over five thousand cultivars worldwide (Jaradat and Zaid, 2004). The large numbers of date palm cultivars has been a target of considerable research, both for phenotypic and genetic diversity. These studies help in understanding the taxonomy, origin and evolution of this tree.

2.4.1 Phenotypic diversity

There are several thousand date palm cultivars cultivated across the world, to the point that the production and industrialization of the fruit is an ever growing process that has been steadily moving forward since the 1990's (Zaid, 2002). The farmers carry out careful selection of the best cultivars, which is accompanied by the ongoing increase in the number of cultivars around the world in order to improve the quality and production of date palm (Elshibli, 2009).

Based on botanical descriptions, there are about 250 cultivars in Oman (MAF, 2005), 450 in Saudi Arabia (Al-Khalifah and Askari, 2003), 135 in the United Arab Emirates (Ghaleb, 2008), 600 in Iraq (Khierallah *et al.*, 2011a), 400 in Iran, 244 in Morocco, 250 in Tunisia and 400 in Sudan (Khanam *et al.*, 2012). A wide range of phenotypic characteristics of date palm *fruits*, *spathes*, *spadices*, *leaves*, *leaflets* and *spines* has been used to identify date palm accessions (Table 2.1, Mohamed Ahmed *et al.*, 2011; Hammadi *et al.*, 2009; Ould Mohamed Salem *et al.*, 2008). These characters are reported as part of standard descriptors in the date-palm (IPGRI, 2005).

Mohamed Ahmed *et al.* (2011) observed high levels of variability in twenty-one date palm accessions originating from different locations in the Adrar region evaluated using thirty vegetative and reproductive measurements. They also reported that characters related to *leaflets* and *spine length* and *fruit* and *seed size* accounted for a large proportion of the observed variability. In addition, they observed a typically continuous phenotypic diversity among date palm accessions and little association between cultivars with similar *fruit* characteristics from same geographic origin, such as Amsakhsi and Adaghd cultivars.

Table 2-1: Examples of measured phenotypic traits for a number of date palm cultivars, their country of origin and references.

References	Country of origin	Cultivars studied	Measured traits				
			Leaf	*Spines*	*Pinnaes*	*Fruit*	*Spadice*
Mohamed Ahmed *et al.* (2011)	Mauritania	Ahmar, Amsakhsi, Tamchkrert, Bouseker, Tiguidert, Lemdina ghailania, Tijib, Adaghd, Sembahra, Sel medina, Boudjeire, Sijoumen, Sembahmoud, Enzer, Athmenmej, Tenwazidi, Sekanni, Temazad, Tenterguel el kahla, Tadeghdit el hadi, Lemdina	*Spinelength, petiole width at the bottom, leaf length* and *width, rachis thickness between the last spine* and *the first leaflet, leaflet part length*	*Number, length and width at the middle).* *Leaflets (terminal leaflet length, leaflets number, spacing index, terminal leaflet width, leaflet width* and *length at the middle*	-	*length, width, weight, pulp & seed weight*	*length, width at the middle*
Hammadi *et al.* (2009)	Tunisia	Deglet nour, Alig, Kintichi	*length, midrib length, pinnated part length,*	*number, density, % solitary spine, spine length at the middle,*	*number, density, % of antrose pinnae,*	-	-
Ould Mohamed Salem *et al.* (2008)	Mauritania	Ahmar1, Lemdina gouchatia, Tijeb, Tiguidert, Lemdina ghailania, Ahmar2, Adaghd, Sekani, Amsakhsi, Tamchkrert, Alfa	*length, width and angle, spineted part, length of leafleted part, petiole width,*	*number, middle spine width* and *length*	-	-	-

Furthermore, three date palm cultivars (Deglet nour, Alig and Kintichi) were screened using thirty vegetative characters and the results showed the stability of 6 characters: *spine length at the middle*, *percentage of spined midrib part*, *maximal pinnae width at the top leaf*, *apical divergence angle*, *maximal spine angle* and *percentage of solitary spines*. These characters are not affected by the change of cropping conditions and could be used for cultivar identification outside the fruiting period (Hammadi *et al.*, 2009). In addition, Hammadi *et al.* (2009) found that cultivars with the same fruit consistency group together in diversity analyses. Similar findings have been reported by Ould Mohamed Salem *et al.* (2008) who studied twelve Mauritanian date palm cultivars using eighteen phenotypic traits focused on vegetative systems. Ould Mohamed Salem *et al.* (2008) have also suggested that the leaves of the date palm can provide an accurate description of the different features inherent in a specific cultivar (Figure 2.4).

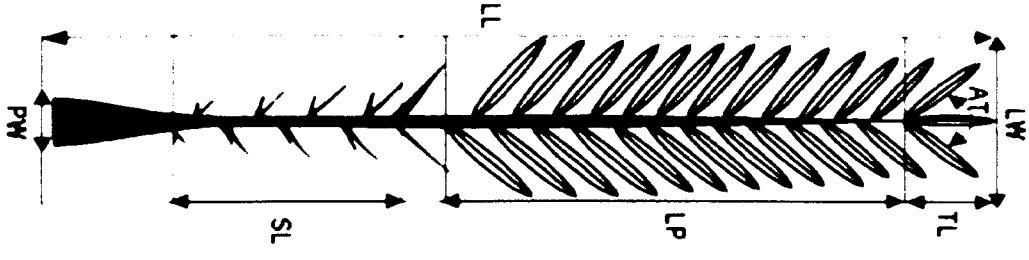


Figure 2.4: Some of traits measured in leaves of date palm (Ould Mohamed Salem *et al.*, 2008).

Elhoumaize *et al.* (2002) were able to differentiate between 26 date palm accessions from Morocco using 26 vegetative traits. The results of this study suggested that a great deal of phenotypic variability was present among the

different accessions. The study also concluded that some morphological characters were highly correlated with each other and possibly associated with resistance to Bayoud disease.

The classification of the date palm is regarded as a very important matter and due to the number of cultivars available it can sometimes be a complicated matter to identify them uniquely (Osman, 2001).

Elshibli (2009) provided an overview of the complexity of classification and reported that date palms in Sudan and Tunisia, for example, are described as 'Safra' or 'Hamra' while the same types in the north Kordofan area of Sudan tend to be described by farmers based on the color of the edible part at the Rutab stage. Elshibli & Korpelainen (2008) have provided another example in which farmers in Egypt and Sudan have kept their original cultivar names or where names have undergone only minor alterations; such is the case of Gondaila and Bitamoda in Sudan, which has been replaced by Gondila and Bertamoda in Egypt.

Several other morphological markers have been used to identify date palm cultivars. Tisserat and DeMason (1982) used the morphology of date pollen to differentiate between 4 date palm male cultivars and also different *Phoenix* species.

Yield potential is also one of the most important features that determines which cultivar to select and can be indicated by a wide range of morphological features of fruits (Elshibli, 2009).

Chemical composition of fruit, including sugars, dietary fiber, volatile matter, acidity as well as the pattern of changes that are normally present during the

different phases of ripening also can enable characterization of date palm cultivars (Elshibli, 2009).

2.4.2 Genetic diversity

Date palm is a diploid with 36 chromosomes ($2n = 2x = 36$) and the genome size is estimated to be approximately 658 Mb (Al-Dous *et al.*, 2011; Elmeer *et al.*, 2011). Three genotypes (AA, Aa, aa) and two phenotypes (A or a; assuming dominance) are to be expected for any given locus in date palm. The genotype is considered heterozygous if the two alleles are not identical, while the genotype is homozygous when the two alleles are identical, (El Hadrami *et al.*, 2011). In some cultivars, variation in chromosome number has been observed suggesting an apomictic origin (Al-Khalifah and Askari, 2006).

Different morphological features based on fruit characteristics (*colour*, *shape*, *weight*, and *texture*) plus the morphology of *leaves*, *spadice*, *spathe*, *spines* and *pinnae* have been used to describe many varieties and constitute a useful method to analyze phenotypic diversity in this important crop (Al-Khalifah and Askari, 2006; Rhouma *et al.*, 2008; Abdulla and Gamal, 2010; Mohamed Ahmed *et al.*, 2011; Khanam *et al.*, 2012; Hammadi *et al.*, 2009; Ould Mohamed Salem *et al.*, 2008; Elhoumaize *et al.*, 2002). However, many of these features may undergo changes due to environmental conditions and may not reflect the true genetic relationships. Some of these features are also time consuming to record and can only be assessed when the palm are sexually mature with the onset of fruiting takes approximately 3 to 5 years. A large set of phenotypic data are required which is difficult to collect and statistically variable, due to environmental effects (Cao and Chao, 2002; Al-Khalifah and Askari, 2003).

Date palm is a diecious plant which makes it an obligate out-crosser. This dimorphic nature leads to a highly heterogeneous genetic structure in date palm (Al-Ruqaishi, 2006). More recently, clonal propagation of date palm is preferred to seedling to preserves the genetic integrity of the cultivars, but it was recognized that even within a cultivar there has been seed propagation in its ancestry, leading to multiple genotypes within each cultivar (Khanam *et al.*, 2012; Al-Khalifah and Askari, 2007; El Hadrami *et al.*, 2011). As such, date palm cultivars are more akin to landraces than to single genotypes.

Along with the complexity of population structure, the apparent difficulties with morphological studies as well as the need to resolve cultivar identity at early stages of plant development lead to an increased interest to study date palm genetic diversity. Considering the importance of plant diversity, researchers developed different types of molecular markers that proved to be effective in assessing genetic diversity in plants and particularly in date palm species.

2.5 Molecular markers

A molecular marker is a measurable character that can identify variation in either protein or DNA sequence. Both phenotypic and genotypic traits can act as genetic markers if they identify genotypic and/or phenotypic characteristics of an individual and the inheritance of these traits can be followed through different generations. To overcome the limitations of morphological traits, other markers have been developed at the protein level (biochemical markers) and the DNA level (molecular markers) to assess the genetic variability as a

complementary strategy to more traditional approaches in plant genetic resource management (Bagali *et al.*, 2010; Farooq and Azam, 2002).

2.5.1 Biochemical markers

Biochemical markers (seed storage proteins and isozymes) are usually named 'protein markers' and were the most frequently used markers for genetic studies before the advent of DNA markers in the 1980s (Jonah, *et al.*, 2011; Farooq and Azam, 2002). These markers are generated through electrophoresis, taking advantage of the differential migrational properties of proteins and enzymes. They can be visualized by histochemical stains specific to each enzymes being assayed or through total protein stains, such as Coomassie Blue. Farooq and Azam (2002) reported that the isozymes and proteins have a neutral effect on the plant's phenotype, and are often expressed co-dominantly, resulting in discrimination between homozygotes and heterozygotes. However, protein markers are limited in number and can be affected by the environment and may be tissue/developmental stage specific; thus, the resolution of diversity can be limited and they represent only a small part of the genome (Jonah *et al.*, 2011; Mondini *et al.*, 2009; Al-Khalifah and Askari, 2006).

2.5.2 DNA markers

A number of DNA markers have been developed and have become valuable tools for detecting genetic diversity and elucidating phylogenetic relationships by identifying the differences or polymorphisms within a nucleic acid sequence between different individuals. In addition, these markers can be used for identifying markers associated with specific traits, gene introgression

through backcrossing, genetic diagnostics, germplasm characterization, the study of genome organization and the characterization of transformants (Muchugi *et al.*, 2008; Mondini *et al.*, 2009; Arif *et al.*, 2010, Karp *et al.*, 1997; Neale *et al.*, 1992; Semagn *et al.*, 2006; Jonah *et al.*, 2011, Agarwal *et al.*, 2008, Farooq and Azam, 2002; Guimaraes *et al.*, 2007; Jain *et al.*, 2002).

A genetic marker can be defined as a DNA sequence with an identifiable location on a chromosome, (Jonah *et al.*, 2011). It can be a long repeat sequence, such as minisatellites or a short one like a sequence surrounding a single base-pair change (single nucleotide polymorphism; SNP). These markers are numerous, resulting in high resolution genome sampling and they can be found in nuclear, mitochondrial and chloroplast DNA (Karp *et al.*, 1997).

Mondini *et al.* (2009) and Jonah *et al.* (2011) have stated that molecular markers can be described as differences or polymorphisms which occur naturally within a nucleic acid sequence as results of base pair deletions, insertions, translocations, mutations or duplications. Molecular markers are unlike morphological traits as they are not affected by environment and they can be applied at any stage during plant development (Jonah *et al.*, 2011).

According to Jonah *et al.* (2011) and Bagali *et al.* (2010) an ideal DNA maker should be polymorphic and co-dominant, so able to distinguish between homozygotes and heterozygotes. They should be randomly and frequently distributed throughout the genome, thereby providing a 'representative' indication of overall diversity (Muchugi *et al.*, 2008). They should also be reproducible giving the same results in different laboratories at different times.

Nowadays, a number of markers are available to detect polymorphisms in nuclear DNA and these can be classified into two categories: non-polymerase chain reaction (PCR) based markers or hybridization based markers and, polymerase chain reaction (PCR) based markers (Agarwal *et al.*, 2008, Mondini *et al.*, 2009, Jonah *et al.*, 2011).

2.5.2.1 Non - polymerase chain reaction (PCR) based markers

Restriction fragment length polymorphism

Restriction fragment length polymorphism (RFLP) was the first DNA-based molecular markers developed in early 1980s (Guimaraes *et al.*, 2007; Jonah *et al.*, 2011; Farooq and Azam, 2002). These markers are usually inherited as Mendelian characters and can detect variation in DNA sequences at the same loci in different individuals. The variation or differences in DNA sequences may arise due to simple or large-scale base pair changes as a result of translocation, inversion, deletion or transpositions. According to Jonah *et al.* (2011) these changes result in a loss or gain of recognition sites at the small scale or major alterations which in turn lead to restriction fragments of different lengths.

RFLP markers combine the use of hybridization with restriction endonucleases (Southern, 1975). Restriction endonucleases are bacterial enzymes able to cut DNA creating polynucleotidic fragments with different sizes (Mondini *et al.*, 2009).

In this method, DNA is digested with restriction enzyme such as *Eco*RI. The digested DNA fragments are separated on a gel using electrophoresis. The DNA fragments are blotted as denatured (single stranded) DNA on a hybridization membrane and probed with a labeled clone, washed and exposed

to x-ray film (Guimaraes *et al.*, 2007; Mondini *et al.*, 2009; Farooq and Azam, 2002).

The RFLP markers are co-dominantly inherited, highly polymorphic and reproducible (Agarwal *et al.*, 2008). They also allow screening of many samples at the same time (Mondini *et al.*, 2009). The use of this method is, however, restricted due to several limitations and it involves expensive, radioactive and toxic reagents. It also requires large quantity of high quality genomic DNA and is time consuming (Agarwal *et al.*, 2008).

2.5.2.2 Polymerase chain reaction (PCR) based markers

Polymerase chain reaction (PCR) was first discovered by Mullis *et al.* (1986) for DNA amplification and is considered as an important milestone in molecular biology research (Bhat *et al.*, 2010). In PCR, *Taq* DNA polymerase (or similar; a thermo-stable enzyme) makes copies of a target sequence starting from two artificial primers, which are complementary to the sequences bracketing the target. The amplification of target sequence will pass through different stages; heating to separate the double stranded DNA and cooling to allow the primers to re-anneal. Polymerase chain reaction based markers - unlike RFLPs - require less DNA and are able to process large numbers of samples quickly and efficiently (Guimaraes *et al.*, 2007).

Randomly amplified polymorphic DNA (RAPD)

Randomly amplified polymorphic DNA (RAPD) was the first PCR-based molecular marker to be used in genetic variation analyses. It detected the polymorphism in DNA by using a single, short arbitrary oligonucleotide sequence; mostly ten bases long with at least 50% GC content (Agarwal *et al.*,

2008, Mondini *et al.*, 2009; Bhat *et al.*, 2010). These oligonucleotide sequences can amplify many loci at the same time, allowing multiple markers to be assayed in a single PCR reaction. DNA segments to be amplified will be selected randomly. Furthermore, the RAPD products are separated on agarose gel in the presence of ethidium bromide and visualized under ultraviolet light (Bhat *et al.*, 2010).

RAPD markers can be applied directly to any species as no sequence data are needed for the organism being tested (Guimaraes *et al.*, 2007). The major drawback of this method is that RAPD markers are generally dominant in nature, causing a loss of information relative to markers which show co-dominance. Additionally, the RAPD method is unreliable and sensitive to a number of factors including DNA quality, reagents, PCR conditions and equipment, which can vary between two different laboratories (Jones *et al.*, 1997).

Amplified fragment length polymorphism (AFLP)

Amplified fragment length polymorphism (AFLP) is another PCR-based method derived from the selective amplification of restriction fragments and was developed by Vos *et al.* (1995). It combines restriction digestion and PCR-based technology. It is highly reproducible and equally applicable to all species (Ovesna *et al.*, 2002; Guimaraes *et al.*, 2007; Bhat *et al.*, 2010, Agarwal *et al.*, 2008; Jonah *et al.*, 2011; Arif *et al.*, 2010; Karp *et al.*, 1997).

According to Muchugi *et al.* (2008) AFLPs fragments are normally between 80 and 500 base pairs (bp) in length. This technique can generate a large number of polymorphisms (Farooq and Azam, 2002). It can also produce

fingerprints of any DNA without prior knowledge of DNA sequence (Agarwal *et al.*, 2008). However; AFLPs are dominant markers (Guimaraes *et al.*, 2007; Karp *et al.*, 1997). AFLPs are widely used for fingerprinting studies and in the screening of biodiversity as well as for the detection and evaluation of genetic variation in germplasm collections (Werner *et al.*, 2000). It is also used in plant genetic mapping, establishing linkage groups in crosses (Yin *et al.*, 1999).

The AFLP technique involves extraction of highly purified DNA, restriction endonuclease digestion of DNA (usually with two specific enzymes, one a rare cutter and the other a frequent cutter), ligation of oligonucletide adapters, pre-selective amplification, selective amplification and polyacrylamide gel analysis of amplified fragments. The AFLPs bands can be detected by silver staining or by labeling of the primers with a radioactive isotope (Ovesna *et al.*, 2002; Bhat *et al.*, 2010; Arif *et al.*, 2010). Alternatively, the bands can be also detected with a higher throughput using an automated DNA sequencer with fluorescently labeled primers (Guimaraes *et al.*, 2007).

Simple sequence repeats (SSRs) or microsatellites

Simple sequence repeats (SSRs) or short tandem repeats or microsatellites consist of tandemly repeated mono-, di-, tri- or tetra-nucleotides (e.g., $[A]_n$, $[CA]_n$, $[AGC]_n$, $[GACA]_n$), where n refers to the total number of repeats. SSR occur as interspersed repetitive elements in all eukaryotic genomes with different lengths of repeat motifs, both in coding and non-coding regions (Ijaz, 2011; Jonah *et al.*, 2011; Guimaraes *et al.*, 2007). Knowledge of the sequence

of these repeats is used for designing specific amplifying primers for regions flanking the microsatellite repeat (Ijaz, 2011).

SSRs can be found in nuclear, chloroplast and mitochondrial genomes (Soranzo *et al.*, 1999). Agarwal *et al.* (2008) and Guimaraes *et al.* (2007) reported that the number of repeats of the SSR is highly variable and this is mainly due to slipped strand mis-pairing during DNA replication causing frequent gain or loss of repeat units. Agarwal *et al.* (2008) suggested that microsatellite loci tend to be hyper-variable because slippage in replication occurs at higher frequencies than point mutations. SSRs markers often present high levels of genetic variation based on differences in the number of the tandemly repeating units of a locus (Jonah *et al.*, 2011). In general, the more repetitions of a repeat, the more likely it is to be polymorphic. For example, a $[CA]_{10}$ repeat is more likely to be polymorphic than a $[CA]_4$ repeat (Queller *et al.*, 1993; Farooq and Azam, 2002). However, longer repeats can often lead to poorer amplification in PCR or the generation of more stutter bands, due to *in vitro* slippage in PCR. SSRs markers are co-dominant markers, which can detect substantial variation within populations and between populations with the highest polymorphic content (PIC) of commonly used markers and they have high reliability/reproducibility (Ijaz, 2011; Farooq and Azam, 2002). These characteristics of microsatellites further their application in fingerprinting and different molecular studies.

SSRs markers are robust tools that can be used efficiently by different research laboratories simply by distributing primer sequences (Saghai-Maroof *et al.*, 1994) and do not require large amounts of DNA (Kloda, 2004). A possible problem associated with the use of microsatellites as molecular markers, is the

occurrence of null alleles at some loci (Callen *et al.*, 1993; Farooq and Azam, 2002). Dakin and Avise (2004) defined null alleles as any allele at a microsatellite locus that consistently fails to amplify during the PCR reaction. Microsatellites can be discovered by screening libraries of clones when prior DNA sequence is not available and usually the methodology for their development takes time, is complex and costly (Ijaz, 2011; Kloda, 2004, Guimaraes *et al.*, 2007). Kloda (2004) reported that isolation and characterization of individual loci is essential for the development of locus-specific microsatellite markers. Generally this process involves the construction and screening of a DNA library with specific probes and DNA sequencing of positive clones and subsequent PCR primer synthesis and testing. McCouch *et al.* (1997) provided a good review for microsatellite marker development. Various methods have been employed to increase the efficiency of microsatellite isolation, including enrichment techniques (Edwards *et al.*, 1996), concatenation of sequences for sequence-tagged microsatellite profiling (Hayden and Sharp, 2001), dot blot selection against high copy number sequences (Scotti *et al.*, 2002) and the application of next-generation sequencing (NGS) technologies (Zalapa *et al.*, 2012). Zalapa *et al.* (2012) have reported that NGS technologies allow efficient identification of large numbers of microsatellites at a fraction of the cost and effort of traditional approaches. In addition, NGS methods can produce large amounts of sequence data from which to isolate and develop numerous genome-wide and gene-based microsatellite loci (Zalapa *et al.*, 2012).

SSRs have already been applied in a variety of ways in several plant species and proven to be useful tools for DNA genotyping, genome mapping,

population and parentage analysis, individual identification, phylogenetic studies, conservation and the management of genetic resources (Ijaz, 2011).

According to Agarwal *et al.* (2008) PCR amplification protocols used for microsatellites employ either primer pairs with one of the primers being radiolabelled or fluorolabeled, or unlabeled primer pairs. The PCR product of unlabeled primer pairs can be visualized by either polyacrylamide or horizontal agarose gels. The PCR products of fluorescently labeled microsatellite primers can be separated by capillary electrophoresis using an automated sequencer such as Applied Biosystems' ABI PRISM or Beckman Coulter's CEQ 8000 Genetic Analysis System. However, the fluorescently-labeled microsatellite primers are costly to buy. A procedure was introduced by Schuelke (2000) in which three primers are used for the PCR amplification for each microsatellite: an SSR-specific forward primer with a M13 tail, an SSR-specific reverse primer and the universal fluorescent-labeled M13 primer. The allelic size for each SSR marker can be scored and analyzed using a range of different software.

Single nucleotide polymorphism (SNPs)

Single nucleotide variation present between the genome sequences of individuals in a population is known as a Single Nucleotide Polymorphism (SNP; Agarwal *et al.*, 2008; Mondini *et al.*, 2009). The occurrence and distribution of SNPs throughout the genome varies among species.

SNPs are the most abundant DNA markers in plant genomes and can detect changes in nucleotide sequences down to single base pairs (Chen *et al.*, 2011; Farooq and Azam, 2002). Guimaraes *et al.* (2007) have reported that SNPs

could be found very close to or within a gene of interest. SNPs can be also used to detect a known functional nucleotide polymorphism and are (potentially, depending on the detection system) co-dominant markers (Guimaraes *et al.*, 2007). The benefits of SNP assays include higher map resolution and throughput, lower error rate and the parallel assay of multiple SNP (Guimaraes *et al.*, 2007).

The SNPs are usually more common in the non-coding regions of the genome as their presence in the coding regions can generate mutation (e.g. non-synonymous mutations that result in an amino acid sequence change or synonymous mutations that do not cause any change in the amino acid sequence) (Agarwal *et al.*, 2008). Synonymous changes can result in phenotypic differences by modifying the mRNA splicing, altering active site function or protein folding, among other causes of modified function.

Detection of SNPs requires an initial DNA sequence in a reference individual plus re-sequencing in other varieties to find variable base pairs.

Diversity arrays technology (DArT)

Diversity arrays technology (DArT) has some aspects of the AFLP procedure, but using hybridization to a microarray for genome wide discovery with high throughput. The applications for DArT include genetic diversity analysis and cultivar identification, genetic map construction and quantitative trait loci (QTL) identification and genome profiling (Guimaraes *et al.*, 2007; Wittenberg *et al.*, 2005).

DArT was initially reported by Jacoud *et al*, (2001) and has been applied successfully in many crops including barley (Wenzl *et al.*, 2004), wheat

(Akbari *et al.*, 2006; Mantovani *et al.*, 2008), rice (Xie *et al.*, 2006), *Arabidopsis thaliana* (Wittenberg *et al.*, 2005), cassava (Xia *et al.*, 2005) and Musa (Risterucci *et al.*, 2009). No DNA sequence information or site specific oligonucleotides are required for a species to be studied (Wittenberg *et al.*, 2005; Semagn *et al.*, 2006; Stodart *et al.*, 2007; Amorim *et al.*, 2009). Mantovani *et al.* (2008) and Xia *et al.* (2005) reported that with proper setup and software, this particular application has the potential of processing hundreds to thousands of individual samples and producing hundreds of high-quality genomic marker based on polymorphisms between individuals that are cost and time efficient compared to the other markers. DArT is highly reproducible and the patent for this technique is essentially under a free license through its application in an open-source model (Semagn *et al.*, 2006). DArT technology passes through several steps: complexity reduction of DNA, library construction, printing and processing of microarrays onto glass slides, hybridization of fluorescently labeled amplicons onto slides, washing and scanning of slides for hybridization signal, and data extraction and analysis (Wittenberg *et al.*, 2005; Mondini *et al.*, 2009).

DArT is available with limited development costs and analysis can be performed by any experienced researcher who can prepare genomic DNA, although there is service cost of the analysis. The recent development of DArTSeq – a sequence Tag-based variant of DArT which uses Next Generation Sequencing to develop data – has great potential to integrate marker data with subsequent genome sequence, as well as generating over 10x the number of markers that an equivalent slide-based array would generate (Tinker *et al.*, 2009).

Next generation sequencing (NGS) technologies

Recently, next or second generation sequencing (NGS) technologies with high-throughput sequencing and low cost have become the first choice for researchers carrying out molecular research, avoid the handling of individual clones from shotgun libraries and produce thousands or millions of sequences in one assay (Zalapa *et al.*, 2012; Rounsley *et al.*, 2009). Imelfort and Edwards (2009) described NGS as 'platforms that can produce millions of short DNA sequence reads of length usually between 25 and 400 bp'; although these reads are shorter than the traditional Sanger sequence reads.

The three brands of NGS technologies to be commercialized are 454 Life Sciences (Roche), Solexa (Illumina) and ABI SOLID (Agencourt Biosciences). The 454 platforms can produce longer read lengths while Solexa and ABI SOLID produce very large quantities of very short reads (Rounsley *et al.*, 2009).

Zalapa *et al.* (2012) reported that pyrosequencing technology was initially developed by Pal Nyrén in the 1990s (Nyrén, 2007) and 454 Life Sciences (Roche Diagnostics, Indianapolis, Indiana, USA) were the first to optimize this method as an NGS platform. GS20 was the first successful pyrosequencing system developed and commercialized by Roche in which over 20 million base pairs was sequenced in just over 4 hours (Imelfort and Edwards, 2009). In 2007, GS20 was replaced with another model called GS FLX, with the ability to produce over 100 million base pairs of sequence in a similar amount of time (Imelfort and Edwards, 2009). More recently, Titanium chemistry was combined with this technology, increasing read-length to more than 400 Mbp of sequence and an average read-length of around 400bp.

However, another two high throughput sequencing systems (SOLiD and Solexa) now compete with GS FLX. The Solexa Genome Analyzer (GAIIx) system possesses a reversible termination property and can generate up to 50,000 million bases of data per run whereas SOLiD is based on sequential ligation with dye labeled oligonucleotides and can generate more than 20 gigabases of data per run (Imelfort and Edwards, 2009). Most of these platforms have gone through multiple rounds of improvements and upgraded specifications (Rounsley *et al.*, 2009).

2.6 Examples of molecular marker application in date palm

A wide range of molecular markers (e.g., RFLPs, RAPD, AFLPs and SSRs) have been used for a number of potential objectives in date palm, including; identification of genetic variation in date palm (Bodian *et al.*, 2012; Hamza *et al.*, 2012; Haider *et al.*, 2012; Khierallah *et al.*, 2011a,b; Johnson *et al.*, 2009; Elshibli, 2009; Al-Ruqaishi *et al.*, 2008; Rhouma *et al.*, 2008; El-Tarras *et al.*, 2007; Al-Moshileh *et al.*, 2004; Sedra *et al.*, 1998) and development of markers to distinguish between male and female trees during the early stages before inflorescences (Al-Mahmoud *et al.*, 2012; Elmeer and Mattat, 2012; Younis *et al.*, 2008, Ahmed *et al.*, 2006). These markers were also used to test somaclonal variation in regenerated plants of date palm (Ahmed *et al.*, 2009), confirmation of some cultivars as a landrace (e.g., Medjool cultivar; Elhoumaizi *et al.*, 2006), detection of genetic stability of date palm plantlets derived from *in vitro* culture (Bader *et al.*, 2007) and study of the genetic variation from offshoots and tissue culture (Gurevich *et al.*, 2005; Al-Khalifah and Askari, 2007).

RFLP markers have been used for the identification of date palm but in relatively few studies due to the high cost and large amounts of DNA required. Four date palm cultivars (Kenessy, Lulu, Nabtha Saif, and Sheshi) obtained from United Arab Emirates plantation were analyzed by RFLP (Corniquel and Mercier, 1997). The analysis was performed on offshoot leaves surrounding the shoot tips of the four cultivars. A cultivar-specific hybridization patterns with a single cDNA probe was generated with total DNA digested by *Eco*RI. Corniquel and Mercier (1997) have reported high levels of polymorphism between the four cultivars, with no variation between the individuals tested (e.g. two individuals for Kenessy, Nabtha Saif, Sheshi and four individuals for Lulu). In contrast, a complex and specific hybridization patterns was reported by Corniquel and Mercier (1994) for five date palm cultivars (Barhee, Deglet Nour, Khalas, Khadrawy and Medjool) and none of these patterns were identical to those reported earlier (Corniquel and Mercier 1997).

In other studies, RAPD markers have been used for the identification and DNA fingerprinting of date palm accessions and appear to be very effective in identifying these accessions, although the exhibited polymorphism was low in some studies (Sedra *et al.*, 1998; El-Tarras *et al.*, 2007). Sedra *et al.* (1998) used RAPD to investigate the genetic variation among 43 date palm accessions, including 37 accessions from Morocco and 6 cultivars from Iraq and Tunisia. They observed a weak association identified by cluster analysis and low levels of polymorphism, which could be related to the mode of introduction and exchange of the Moroccan date palm germplasm between plantations. However, some morphologically similar accessions were found to cluster together (Sedra *et al.*, 1998). The genetic similarity between four

female date palms (Zaghloul, Amhat, Samany and Siwi) and four unknown male trees of Egyptian date palm was studied using the RAPD technique (Soliman *et al.*, 2003). The genetic similarity between the four females ranged from 87.5% to 98.9% and the banding profiles indicated that two out of four male plants were genetically related to the four female cultivars (Soliman *et al.*, 2003). A similar study was conducted by Ahmed *et al.* (2006) to detect the genetic relationship and similarities between four known females (Sakkoty, Malkabi, Bartamoda and Dagana cultivars) and three unknown males of Egyptian date palm. The percentages of similarities ranged between 79.0% and 91.2% and the three males were found to be closely related to the tested four females (Ahmed *et al.*, 2006).

Using the RAPD technique, a number of the Saudi Arabian cultivars were also identified and fingerprinted (Al-Moshileh *et al.*, 2004; El-Tarras *et al.*, 2007). Al-Moshileh *et al.* (2004) found that the genetic similarity for five date palm cultivars (Barhi, Nabtet ali, Rothanah, Ajwa, and Sokkari) ranged between 70% and 85%, with Sokkary distantly related to the Barhi and Ajwa cultivars. These finding are in agreement with El-Tarras *et al.* (2007) who found low levels of polymorphism between another six Saudi Arabian cultivars (Sukkari, Sifri, Sullage, Khalas, Makfazi and Maktoum) indicating that most of the examined Saudi Arabian cultivars are likely to have a narrow genetic base and that RAPD is a reliable technique for the identification of Saudi Arabian date palm cultivars, within lab at least.

Furthermore, RAPD was used to test for the presence of somaclonal variation in 180 plantlets of date palm in comparison with their original mother palm (Ahmed *et al.*, 2009). The clonal plantlets for this experiment were

regenerated from juvenile leaves on regimes using 2,4-D, to induce somaclonal variation (Ahmed *et al.*, 2009). However, RAPD analysis revealed that the tested plantlets were identical with the original mother at the loci tested. Ahmed *et al.* (2009) concluded that no somaclonal variation was detected in this materials under different doses of 2,4 D (1 mg/l, 10 mg/l and 100 mg/l) using RAPD markers.

AFLP fingerprinting has been used for the assessment of the genetic diversity for different date palm cultivars (Khierallah *et al.*, 2011a; Jubrael *et al.*, 2005), for potential mapping populations (El-Kharbotly *et al.*, 1998) and for characterizing genetic variation in clones propagated from offshoots and through tissue culture (Gurevich *et al.*, 2005). Cao and Chao (2002) studied 21 date palm cultivars from California with AFLP and found that cultivars separated into two major groups, demonstrating that AFLP can be used efficiently to distinguish between date palm cultivars.

According to Adawy *et al.* (2004) who used AFLP to study fourteen date palm accessions collected from different locations in Egypt, representing six Egyptian cultivars (Sakkoty, Bertmoda, Malkaby, Gandila, Fraihy and Siwi), the levels of detected polymorphism were low. However, they observed that the genotypes of some cultivars clustered together (e.g. Fraihy and Gandila). They also found that the genotypes of the Siwi cultivar clustered together, although they exhibited some degree of variation. Sakkoty, Bertmoda, and Malkaby cultivars showed a higher degree of variation (Adawy *et al.*, 2004). Additionally, the AFLP assay separated the cultivars according to their location (e.g. Siwi and Fraihy from Aswan; Adawy *et al.*, 2004). Conversely, Jubrael *et al.* (2005) reported a high level of polymorphism among 18 Iraqi

date palm varieties indicating that Iraqi varieties are genetically distinct and that there are likely to be fewer multiple names for the same variety. The high level of polymorphism could be due to several reasons, such as the strong out-crossing mechanism in date palm, which is highly likely to increase the polymorphism or due to the AFLP technique itself and the selected primer combinations (Jubrael *et al.*, 2005). More recently, Khierallah *et al.* (2011a) have observed a large range of genetic diversity between 18 date palm varieties (11 females and 7 males) collected from the center of Iraq using six primer pairs of AFLP. The tested varieties clustered independently of their geographic origin and of their phenotypic characteristics and that all primer combinations contributed to the differentiation between varieties (Khierallah *et al.*, 2011a)

AFLP analysis was also used by Elhoumaizi *et al.* (2006) to confirm that Medjool in Morocco is not genetically uniform. In this study they used 66 Medjool accessions from Morocco, six from Egypt, and four from California plus one accession of Deglet Noor. Elhoumaizi *et al.* (2006) found that a minimum of 79% genetic similarity was shared between the 66 Medjool accessions from Morocco, supporting the idea that Medjool is not genetically uniform and it exists as a landrace in Morocco. This finding increases the possibility that other date palm cultivars may also be landraces in different regions.

Sixteen date palm specific SSRs primer pairs were initially developed by Billotte *et al.* (2004), and used in various studies (Zehdi *et al.*, 2004; Al-Ruqaishi *et al.*, 2008; Elshibli and Korpelainen, 2008, 2009; Ahmed and Al-Qaradawi, 2009; Pintaud *et al.*, 2010; Zehdi *et al.*, 2012). These 16 markers

revealed a high rate of polymorphism, supporting their efficacy for germplasm diversity studies as well as cultivar identification, pedigree analysis and genetic mapping studies. Furthermore, Pintaud *et al.* (2010) evaluated sixteen date-palm SSRs in 308 accessions of *Phoenix* representing 12 species, and revealed high levels of polymorphism and the success of their cross species application strongly indicating the transferability and utility of these SSRs between *Phoenix* species.

Zehdi *et al.* (2004) examined the genetic diversity of 49 date palm accessions from three main Oases with little geographic structure within Tunisia using the 14 SSRs primer pairs developed by Billotte *et al.* (2004). They observed high levels of polymorphism among the 49 accessions and a large number of SSR alleles (7.14 per locus). These results are comparable to other studies conducted by Hammadi *et al.* (2011) and Zehdi *et al.* (2012), which showed a high degree of polymorphism based on microsatellite markers, indicating that the Tunisian date palm collection is characterized by a high degree of genetic diversity. Zehdi *et al.* (2004) and Zehdi *et al.* (2012) also found that the genetic diversity revealed exhibited a unique structure for all accessions independent of both the geographic origin and the sex of trees. This finding is in agreement with Khierallah *et al.* (2011a) who have observed that AFLP profiles of the 18 varieties from Iraq clustered independently of their origin and phenotypic characteristics.

Over recent decades, many studies have reported the use of SSRs markers to genetically study and characterize the date palm germplasm of many countries (Oman, Bahrain, Iraq, Sudan, Morocco and Qater). Al-Ruqaishi *et al.* (2008) observed a high level of polymorphism among 21 date palm accessions

collected from Oman, Bahrain, Iraq and Morocco using the same SSRs primer pairs. The analysis also showed that the Omani accessions were genetically close to accessions from Bahrain and Iraq, while accessions from Morocco appeared distinct from all studied accessions. Furthermore, Elshibli and Korpelainen (2008) reported that the germplasm from Sudan and Morocco were also highly polymorphic, possessing a large number of alleles. A total of 343 alleles with a mean of 21.4 per locus were detected. A high level of observed heterozygosity (0.853) was observed among all accession, while the mean of observed heterozygosity values of the Sudan cultivars, Sudan males and Morocco cultivars were 0.841, 0.799 and 0.820, respectively. These differences are reflection of structure within the population being studied.

Based on F_{ST} values and genetic distances, Morocco accessions showed significant differentiation compared to the Sudanese accessions (Elshibli and Korpelainen, 2008). However, Ahmed and Al-Qaradawi (2009) reported 40 alleles with a mean of 4 alleles per locus by examining 15 Qatari date palm cultivars using the same markers.

Due to the economic importance of the date palm and the effectiveness of SSRs markers, Akkak *et al.* (2009) developed another 17 SSRs markers by constructing two microsatellite enriched libraries of date palm using $(GA)_n$ and $(GT)_n$ synthetic repeats. These SSRs have been used to evaluate 31 cultivars and clones from Algerian and Californian germplasm, which also exhibited a high level of polymorphism among the analyzed samples. The authors also showed the marker transferability in other species across the genus *Phoenix* (Akkak *et al.*, 2009). More recently, 1000 SSR primers were developed at the International Center for Agricultural Research in the Dry

Areas (ICARDA) by Hamwieh *et al.* (2010). The sequences of these markers were derived from the draft of date palm genome generated by whole genome shotgun DNA sequencing (Elmeer *et al.*, 2011). Elmeer *et al.* (2011) used 30 of these SSRs to assess the genetic diversity in 11 cultivars from different locations in Qatar. They noted that out of the thirty, only ten primers were polymorphic, possessing a total of 77 alleles with a mean of 7.7 alleles per locus and an average of genetic diversity 0.80.

Khierallah *et al.* (2011b) and Bodian *et al.* (2012) found high levels of intervarietal polymorphism, suggesting a wide genetic background among date palm accessions collected from different locations in Iraq and Morocco, respectively, using a combination of SSRs markers developed by Billotte *et al.* (2004) and Akkak *et al.* (2009).

The sequences of the nuclear and chloroplast genomes of date palm (Al-Dous *et al.*, 2011 and Yang *et al.*, 2010) have been released and soon will lead to the establishment of a genetic linkage map for date palm as well as more molecular markers being developed. These sequences will also stimulate and reactivate some of the breeding activities that had been abandoned in this crop.

2.7 Sex determination

The production of the date palm fruit mostly takes place in the arid regions of Asia, North Africa and the Middle East. The product is highly valued across the world mostly as a confectionery or fruit crop that in turn provides an important source of income and sustenance in all desert regions. Among all crops, *Phoenix dactylifera* is one of the most important species and the

plantation process is normally extremely costly and requires a considerable amount of time before a plant is able to produce the much desire fruit.

Younis *et al.* (2008) have reported that for those farmers harvesting date palms, on top of the long and costly investment that they have to bear, one of the major hurdles appears in the process of identifying the sex of seedlings in order to cultivate their orchards with enough female trees to harvest and minimize the number of male trees in their plantations.

Multiple attempts have been made to determine the sex of this species which is dioecious during the early life stages but most of them have failed, except in very isolated cases where it has been possible to differentiate markers through sex in one or two varieties but until this day there is no system that has produced markers that have worked across a wide range of cultivars (Al-Mahmoud *et al.*, 2012).

Al-Mahmoud *et al.* (2012) in order to distinguish male/female genders during the early stage of growth of the date palm applied two different approaches. They concluded that their results should be helpful in discriminating between male and female date palm, as well as saving time.

2.8 Resistance in date palm

According to Al-Khatri (2004) date palm cultivars are affected by a wide range of different pests. In some cases pests attack the fruit while in others the fronds and sometimes the trunk. Al-Khatri (2004) has suggested that there are more than 24 different species of arthropods associated with date palm cultivars with the most detrimental pests being: Dubas bug *Ommatissus*

lybicus DeBergevin, Red Palm weevil (RPW), and Lesser date moth (LDM) all of which affect the date palm quantitatively or qualitatively.

It is very hard to ignore the progress made by the application of DNA-based markers in regards to quality assurance. To date, no genetic engineering and no markers have been identified to improve the resistance of date palm cultivars. Nevertheless, Jain *et al.* (2011) have suggested that such applications will be most likely used in the near future. It was not until very recently that studies have been conducted at the genome level of date palms. The different technologies previously described (RFLP, RAPD, AFLP and SSRs) could be used for molecular detection purposes.

To overcome the threat posed by most of the diseases that can affect date palm, the most suitable strategy to follow is to conduct an integrated management approach. To successfully conduct this approach it is necessary to combine different techniques aiming to sanitize, prevent, exclude when necessary, all those palms in need of proper care, preventing propagation to other cultivars of any disease or pest so minimizing the losses either in quantity or quality of a cultivar.

2.9 Salinity tolerance of the date palm

Even though the date palm is known for thriving and growing without many problems in arid regions, there are also multiple environmental conditions in which the date palm manages to survive, including different levels of water deficit, salt and other stress tolerances. According to Pavez *et al.* (2007) date palm is considered to be salt tolerant as it grows under different levels of salinity, however there is no systematic approach to characterize such

genotypes and identify the genes involved (Alhammadi and Kurup, 2012). Erskine *et al.* (2003) have suggested that certain varieties of date palm can survive 22,000 ppm of salt, but their growth and productivity levels are lower than those grown in lower conditions of salinity.

Increasing of salinity is considered an important problem in date palm production in areas where the negative effects are visible (Dakheel, 2005). Supporting such views, Al-Yahyai & Al-Khanjari (2008), have reported that since 2001 there has been a reduction in the population of date palms within the Sultanate of Oman, despite several stress factors being the focus for improvement. The increased levels of salinity in the major growing areas of date palms within the country are believed to be the main cause for the reduction in palm numbers, although the Red Palm Weevil and the Dubas bug are considered the most dangerous pests responsible for the decimation of particular cultivars.

2.10 Genetic mapping

Paterson (1996) defined a linkage map as a 'road map' of chromosomes derived from two different parents and used to measure the relative genetic distances between markers along chromosomes as well as to locate their position based on Mendelian principles of segregation and recombination. A large number of linkage maps based on different marker types have been constructed for many plant species, such as rice, maize, wheat, barley and other cultivated plants (Mohan *et al.*, 1997).

Mapping and sequencing of plant genomes will help in identifying chromosomal locations containing genes and QTLs associated with traits of

interest, gene tagging, evolutionary studies, as well as improve selection activities (Collard *et al.*, 2005). Different molecular markers, such as RFLPs, RAPD, AFLP, and SSR have been used to construct linkage maps of various plant species (Kaga *et al.*, 1996; Peng *et al.*, 2000; Roger *et al.*, 2000). In order for researchers to develop a successful genetic map, an appropriate mapping population and sufficient numbers of markers are needed to perform linkage analysis (Collard *et al.*, 2005).

2.10.1 Mapping population

The choice of mapping population is the most critical decision in constructing a linkage map. The selection of parents to cross is very important for a population segregating for both traits of interest and genetic markers. In particular, high levels of polymorphism should be detected between the parents, so that they can be crossed to obtain segregating offspring for genome mapping (Young, 1994). In general, cross-pollinating species possess higher level of DNA polymorphism compared to inbreeding species (Collard *et al.*, 2005).

The size of mapping population is also important for constructing a reliable map. According to Mohan *et al.* (1997) for preliminary genetic mapping studies the population size ranges from 50 to 250 individuals, however far larger populations are required for high-resolution mapping if the intention is to positionally clone genes.

2.10.1.1 Types of mapping populations

Various types of mapping populations are often used in linkage mapping including: F_2 population, Backcrosses, Recombinant Inbred Lines (RILs),

Near-isogenic Lines (NILs) and Doubled haploids (DHs). Generally, highly inbred plant varieties are crossed to generate the F_1 generation. In this cross type, F_1 individuals are identical to each other and are often largely heterozygous. These can be either crossed to themselves (self-pollinated) or crossed to one of the parental inbreds (backcross, BC) (Figure 2.5; Grant and Shoemaker, 2001). The F_2 population produces an expected Mendelian allele segregation ratio of 3:1 for dominant markers and of 1:2:1 for co-dominant markers. The backcross population shows a 1:1 segregation ratio of alleles at each locus. Selfing of F_2 will produce F_3, such selections can continue for six to eight generations (e.g F_31, F_32, F_33, etc.). According to Mendel's laws, that selfing of generation will reduce the heterozygosity of the progeny by half and repeated selfing of generations eventually results in new inbred lines, sometimes referred to as recombinant inbred lines or RILs. Recombinant Inbred Lines (RILs), Near-isogenic Lines (NILs) and Doubled haploids (DHs) shows expected 1:1 ratios, irrespective of whether the genetic markers are dominant or co-dominant as no heterozygotes exist within these population types and the presence of a single band implies a homozygote (Semagn *et al.*, 2006; Grant and Shoemaker, 2001).

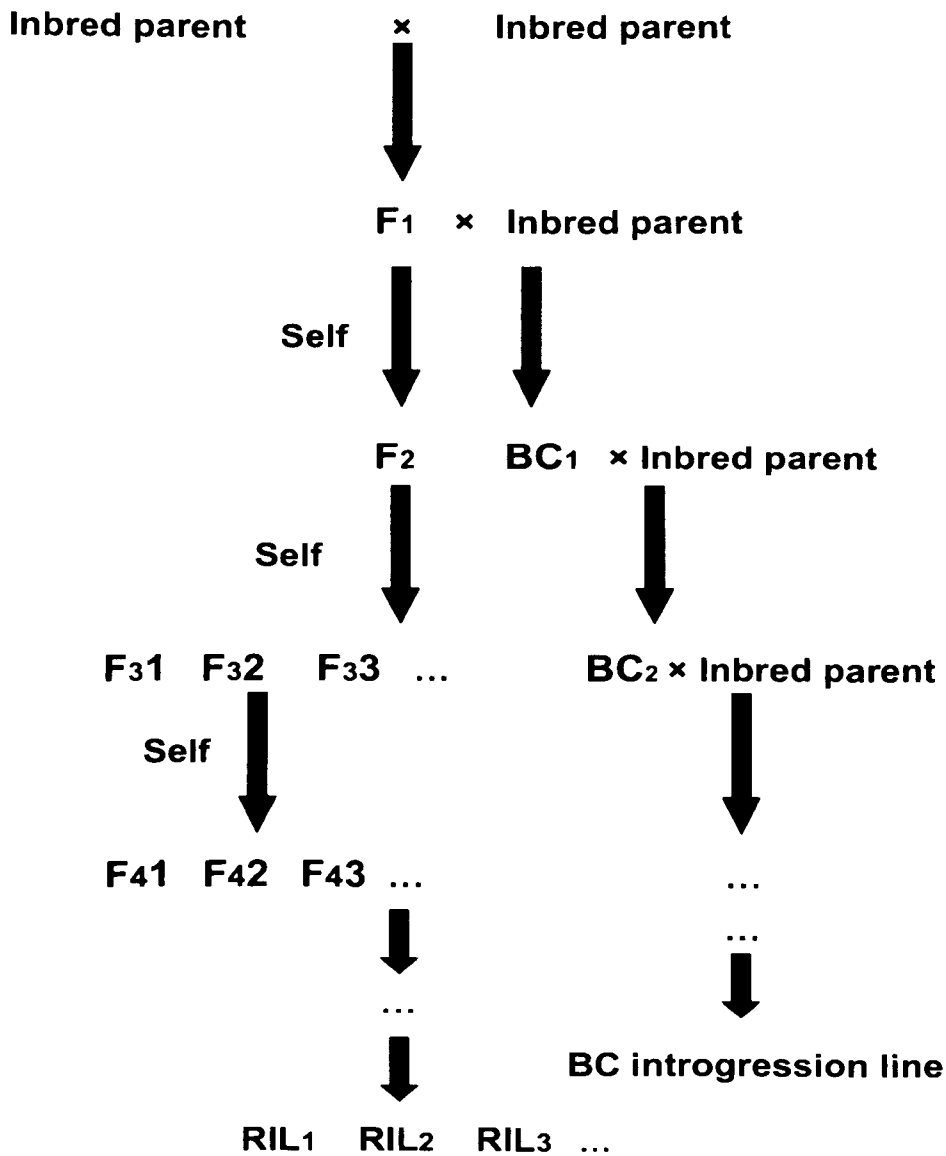


Figure 2.5: Generalized scheme for developing some of the most common population types used in genetic mapping (Grant and Shoemaker, 2001).

2.10.2 Selection of molecular markers

The selection of accurate molecular marker techniques for genome mapping is important and depends on several factors including: the breeding habit and genome size of the plant or organism to be mapped, information already known on the genome organization and the facilities available. Young (1994) have reported that organisms with smaller genomes may require less DNA per

sample than species with very large genomes using such markers for mapping (e.g. RFLPs). Problems may arise with RFLP mapping if too little DNA is used. Generally, RFLPs represent a single copy sequences, the amount of any target sequence in a genomic DNA sample can be vanishingly small. It may be impossible to see a signal after hybridization; if too little DNA is loaded onto the gel for blotting.

Co-dominant markers like RFLP and SSRs are more informative than dominant markers like RAPD and AFLP when mapping in an outbred population or early generations of an inbreeding population after cross-pollination. RAPD and AFLP markers do not require any previously cloned DNA fragments or the DNA sequence information of the genome to be known. According to Powell *et al.* (1996) the SSR marker is a critical technique for genome mapping due to the high information content, high discrimination power, very high reproducibility and ease of scoring. DArT has recently been used but the dominant inheritance is still a limitation for mapping although as this is hybridization based, markers are often likely to be detecting the same locus in the genome in different crosses (Semagn *et al.*, 2006). The more recently developed DArT Seq overcomes a number of these limitations, producing a mixture of classical DArT (presence/absence) markers and SNP (sequence variants) with both being based on 64bp sequence taqs.

Single nucleotide polymorphism (SNPs) markers are an important marker type which can detect changes in nucleotide sequences down to single base pairs and can occur both in coding and non-coding parts of the genome (Chen *et al.*, 2011). In addition, SNPs have been reported that were found very close to or within a gene of interest (Guimaraes *et al.*, 2007). However, SNPs required a

considerable amount of funding for their development in terms of sequence. eSNPs could be one approach, based on transcriptomes generated from expressed genes and their variants in RNASeq.

2.10.3 Genotyping of the mapping population

Once polymorphic markers have been identified, 'genotyping' of the population is essential to distinguish the segregation patterns of particular markers across individuals in the entire mapping population, including the parents if possible. The expected segregation ratios for co-dominant and dominant markers are summarized in Table 2.2. Markers should segregate with Mendelian expectations although distorted segregation ratios may be encountered (Collard *et al.*, 2005).

Table 2-2: Expected segregation ratios for co-dominant and dominant markers in different population types

Population type	Co-dominant markers	Dominant markers
F_2	1: 2:1	3:1
Backcross (BC)	1:1	1:1
Recombinant Inbred Lines (RIL	1:1	1:1
Doubled haploids (DHs)	1:1	1:1

2.10.4 Linkage analysis and map construction

Various computer packages are presently available to create a genetic linkage map and the most widely used is JoinMap which accepts data with different expected segregation ratios and can integrate data from different populations

(Stam, 1993a). The LINKAGE software based on a chi-square analysis and allows only the evaluation of pairwise recombination values (Suiter *et al.*, 1983), while MAPMAKER/EXP performs multipoint analysis using maximum likelihood in F_2 and backcross generations (Lander *et al.*, 1987). GMENDEL (Echt *et al.*, 1992) and Map Manager QTX (Manly *et al.*, 2001) are also programs used to create genetic maps. All programs are freely available from the internet except JoinMap which is a commercial program. In addition, all programs have the same basic principles for map construction, and the major steps in the following linkage analysis are described using JoinMap as an example.

2.10.4.1 Segregation distortion

Significant deviations from expected ratios for each segregating marker can be analyzed using chi-square tests. A deviation of the observed genotypic frequencies from Mendelian expectations in a given genotypic class within a segregating population is called segregation distortion (Semagn *et al.*, 2006; Lu *et al.*, 2002). There are several reasons for segregation distortion, including: sampling/selection during population development, small population size, genotyping score errors, the consequence of missing data, gametophytic competition and sterility factors (Millan *et al.*, 2010). Abortion of the male or female gametes or zygotes and the action of transposable element and environmental agents would also be counted among the factors involved in segregation of markers (Yamagishi *et al.*, 2010; Knox and Ellis, 2002).

Segregation distortion was first reported in maize by Mangelsdorf and Jones (1926), and later was reported in many other crops such as wheat, tomato, rice,

coffee, sorghum, tobacco and barley (Kumar *et al.*, 2007; Loegering and Sears, 1963; Paterson *et al.*, 1988; Zhang *et al.*, 2010; Ky *et al.*, 2000; Pereira *et al.*, 1994; Cameron and Moav, 1957; Goloenko *et al.*, 2002). It occurs in wide crosses as a normal phenomenon, therefore it is important that individual marker locus should be tested for segregation distortion and if necessary, the marker showing high degree of segregation distortion be removed from further calculation. According to Lu *et al.* (2002); Matsushita *et al.* (2003) and Sibov *et al.* (2003) it is better to study the distorted loci after calculating the map as these markers may be distorted towards the same parental alleles or clustered in a small chromosome region. These may be genuine genetic effects and may be important to understanding the biology of the system.

2.10.4.2 Create linkage groups

A linkage group is a group of mutually linked loci which may correspond to positions on the same chromosome. Statistically, it is referred to as a group of loci inherited together according to certain statistic criteria (Ma, 2003). Stam (1993a) have reported that markers are assigned to linkage groups using a logarithm of odds (LOD) value or LOD score, which refers to the ratio of the probability that two loci are linked in a given recombination value over a probability that the two are not linked. If the LOD score is above a critical 'linklod', the marker pairs are provisionally considered to be linked, while if the LOD score is less than 'linklod', they are provisionally considered to be unlinked (Semagn *et al.*, 2006). A LOD of 3 as the minimum threshold value has been used in several studies in order to decide whether or not loci were linked, and this value indicates that linkage is 1000 times more likely than no

linkage (Stam, 1993a). Small LOD threshold values will tend to generate few linkage groups with large number of markers per group, whereas higher LOD threshold values will create fragmented linkage groups, each with smaller number of markers. Generally, if two markers or more are not linked, they will be placed in distinct linkage groups (Semagn *et al.*, 2006). Ideally, a number of linkage groups that is the same as the chromosomes numbers of the species under study are obtained. However, determining number of linkage groups is not straight forward because loci on different chromosomes might appear to be linked by chance or more than one linkage group might be obtained for a single chromosome, which results in a high number of linkage groups compared to the chromosomes number. Where a species has been mapped extensively with co-dominant markers (such as wheat; Somers *et al.*, 2004), then composite maps of placed markers can be generated which allows the researcher to choose markers from known locations and also use these to infer chromosome identity for the obtained linkage groups. This is not a guarantee for individual markers, but the consistent presence of markers previously mapped to the same group is good evidence that they are detecting the same chromosome as previously reported.

2.10.4.3 Estimate map distance and locus order

Several parameters need to be considered for calculating map distances and determining locus order including: recombination threshold value, minimum 'maplod', jump threshold value, and mapping function. Map distances are calculated using only information for marker pairs with a LOD score above 'maplod'. The selection of 'maplod' values is ranged between 0.01 (low

value) and 3.0 (high value). Sometime, the value of 'maplod' should be set between 0.5 and 1.0 to make sure that no extra information is used from distant markers. Constructing a map is a process of adding loci one by one, starting with loci pair having most linkage information (Stam, 1993b; regression mapping). The best position for each added locus is searched for by the program and a goodness-of-fit measure is calculated. If the overall goodness-of-fit for a locus is reduced too sharply (too large a 'jump'), or negative distances are observed then the locus should be eliminated. Several rounds could be run until all loci have been handled once. All loci removed in the first round can be retested in the second round. Markers remaining unmapped in the second round can be forced (by relaxing the 'jump' and negative distance rules) into the final order.

Additionally, one of the mapping functions (Kosambi or Haldane) should be selected to construct a genetic map. Kosambi's mapping function (Kosambi, 1944) assumes a certain degree of interference between crossovers in meiosis, while Haldane's mapping function (Haldane, 1931) assumes absence of interference. These mapping functions are necessary to translate recombination frequencies into linear and additive map units' centimorgans (cM). This term is derived from Thomas Hunt Morgan, who proposed that one map unit is equal to one percent of recombinant phenotypes, or one centimorgans. This is also known as the 'Direct' mapping function, as it translates from Recombination Fraction (Rf) directly to linear distances. In practice, this holds true for small Rf (<0.05) where double recombination events are unlikely to occur. The Haldane mapping functions gives greater map distance than Kosambi, if the recombination frequencies are above 10%,

as Kosambi postulates interference between close recombination events, while Haldane does not. The total map length will be greater for the Haldane than Kosambi mapping functions.

Locus ordering is performed using one of the three locus ordering criteria: weighted least squares, maximum likelihood and minimum sum of adjacent recombination fractions. A study was performed to compare the performance of these three criteria in the presence of missing values, typing errors and distorted segregation ratios (Hackett and Broadfoot, 2003). The study concluded that map inflation was high with the maximum likelihood criterion. While using weighted least-squares, the distances between markers are calculated from the map distances between all pairs of markers on a chromosome, as a result the impact of typing errors on the distance between adjacent markers is less severe.

2.10.5 Quantitative trait locus (QTL) mapping

Many economically important traits in plants such as yield, quality, heat tolerance, drought tolerance and some forms of disease resistance are controlled by quantitative trait loci (QTL) (Collard *et al.*, 2005). A quantitative trait locus (QTL) is a region of any genome which is responsible for variation in the quantitative trait of interest (Kearsey, 1998). Therefore, studying and locating the genes controlling these traits is very important. For this purpose QTL maps have been developed to identify the genomic regions associated with traits of interest (Collard *et al.*, 2005). To carry out a QTL analysis, it is necessary first to obtain a progeny segregating for the character of interest. The segregating population is then scored for its trait values in each individual using an appropriate design.

2.10.6 Construct a genetic map for date palm

In date palm, molecular breeding is still in its infancy and the inheritance patterns of traits in this crop are not fully understood due to the non-availability of segregating populations derived from controlled crosses. Therefore, no physical or linkage maps have yet been constructed for date palm (Jain *et al.*, 2011).

Screening for a desirable palm cultivar could be possible at the seedling stage by using a marker-based selection strategy. The same method can be used to distinguish between male and female palms before flowering. The prerequisite for marker-based selection is the identification of a marker tightly linked to a trait of agronomic interest. This objective can be achieved through the initial construction of a complete linkage map (Gebhardt and Salamini, 1992). The cosegregation of the molecular marker and the trait of interest, in a progeny segregating for the trait, is an indication of linkage between them. This is helpful in shortening the breeding program especially in date palms, which requires many years before flowering. Marker assisted breeding, developing molecular markers to segregate sex before flowering, mapping and sequencing of chromosomal locations containing genes and QTLs associated with traits of interest in date palm is needed in a country like Oman where molecular assisted date palm improvement is at the developmental stage. Considering the importance of the date palm to Omani people, the construction of a genetic map for date palm has been carried out and the results are presented in the relevant section.

2.11 Aims of the present study

The conservation of date palm genetic resources is an important issue for the development of the date palm industry and for food security in Oman and many other countries. It is now a matter of urgency that new date cultivars should be bred with higher and sustainable yield potentials, superior quality, and multiple resistances to diseases and pests. It is also important that various valuable date palm germplasm sources should be identified and utilized to improve the socio-economic conditions of the grower and for future food security.

Furthermore, constructing a date palm genetic map and the determination of genetic variability (allowing proper cultivar identification) would be of major importance. In particular, date palm breeding programmes are very immature due to the diccious nature of date palm as fruit bearing palms are only identified many years after planting. A marker to sex determination in date palm would make breeding programmes more viable.

The aims of this study can be summarised in the following goals:

1. **To screen and develop new microsatellite markers (SSR) for date palm (*Phoenix dactylifera* L.) (Chapter 4).**
2. **To investigate the genetic diversity of Omani germplasm and compare accessions with other countries' germplasm (Chapter 5).**
3. **To begin the genetic mapping of date palm and to facilitate the identification of markers linked to traits of interest, such as sex determination gene or resistance genes for disease and salinity (Chapter 6).**

4. **To develop new markers for gender discrimination in date palm that can be used in breeding programmes (Chapter 7).**

Chapter 3. GENERAL MATERIALS AND METHODS

3.1 Introduction

The following section will outline all materials and methods, which have been used throughout the whole study. Specific data analysis for individual experiments will be described in each chapter.

3.2 Experiments

Most of the molecular analyses, such as DNA extraction from date palm accessions, PCR and fragment analysis were performed at the Tissue Culture and Biotechnology Research Laboratory, Directorate General of Agriculture and Livestock Research, Ministry of Agriculture and Fisheries, Sultanate of Oman. Developing and screening of new microsatellite for date palm was carried out at the South Lab, Plant and Crop Sciences, School of Biosciences, The University of Nottingham, UK.

3.3 Plant material

The majority of date palm (*Phoenix dactylifera* L.) leaf materials used in this study was collected by the author from Oman. In addition, DNA samples from Italy (Sanremo, Bordighera), USDA-ARS (United States Department of Agriculture-Agricultural Research Service) and France were kindly provided by Dr. Jean-Christophe Pintaud from IRD (Institute de recherché pour le developpement) in Montpellier. Samples from different origins were donated by people from Iraq, Libya, Sudan and Iran and are also used in this study.

3.3.1 Samples for screening and testing of microsatellite primers (SSRs)

Eight parents of the available Omani date palm controlled crosses (Khalas 4, Khalas 13 male, Um-Alsela, Khori male, Barni, Naghal, Bahlani male, and Khasab) (Chapter 4) were used to screen 171 microsatellite primers, as produced by Billotte *et al.* (2004), Akkak *et al.* (2009), ourselves and Hamwieh *et al.* (2010), all primer sequences are listed in Appendix 1.

3.3.2 Samples for genetic diversity

One hundred ninety-four accessions (151 female cultivars and 43 male trees) of Omani date palm, summarized in Chapter 5, were used to study the genetic structure of Omani date palm. These were collected from the National Germplasm Collection at Wadi Qurayat Research Station, Bahla, Sultanate of Oman. Each female accession was represented by a single tree, but a number of the male trees were represented by 2-3 replicates. The study also included samples from Italy (Sanremo, Bordighera), (USDA-ARS), France, Iraq, Libya, Sudan and Iran for comparison of Omani germplasm with germplasm from other countries and these are all listed in Chapter 5.

3.3.3 Samples for genetic mapping

A total of 83 palms were used for the purpose of genetic linkage mapping analysis, along with three samples of available parents. Fifty-three palms are from a BC_1 population and the other 30 palms are a F_1 population. The two populations were developed using the same male (Kl-96-13) and two different females Khalas 4 and Um-Alsela (El Kharbotly *et al.*, 2006) as listed in Chapter 6.

3.3.4 Samples for testing new microsatellite primers (SSRs) for gender discrimination

One hundred and ninety-four Omani accessions and 96 samples from BC_1 and F_1 populations were used to screen and test new microsatellite primers for gender discrimination. Samples from Italy, USDA-ARS, France, Iraq, Libya, Sudan and Iran, giving a total of 96 samples, were also used to screen and test the new microsatellites. All samples are listed in Tables 7.1 and 7.2; Chapter 7.

3.4 DNA extraction

Mature leaves were collected and stored in a cool box until returned to the laboratory and frozen at -80 °C until DNA was extracted. Total genomic DNA was extracted from samples using the DNeasy plant Maxi kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions.

3.4.1 DNA extraction using DNeasy Plant Maxi Kit

The sample was pulverized under liquid nitrogen using a pre-chilled mortar and pestle. One gram of the powdered tissue was transferred to a 15 mL tube, 5 mL of preheated Buffer AP1 and 10 µl of RNase were added with vigorous mixing. The samples were incubated for about 30 min at 65°C with frequent swirling. This was followed by disruption of the cell membranes to release the DNA into the extraction buffer achieved by adding 1.8 mL of Buffer AP2 and incubating for 10 min on ice. After incubation, the lysate was spun at 3000–5000 x *g* for 5 min at room temperature. The supernatant was carefully decanted and transferred to a QIAshredder Maxi Spin Column and placed in a 50 mL collection tube and spun at 3000–5000 x *g* for 5 min at room temperature. The flow-through was then transferred to a new 50 mL tube. 1.5

volumes of Buffer AP3/E were added. After mixing, the samples were applied to the DNeasy Maxi Spin Column and centrifuged at 3000–5000 x *g* for 5 min. The flow-through was discarded and the precipitated nucleic acids were washed with 12 mL of Buffer AW on the DNeasy Maxi Spin Column. The Spin Column was centrifuged for about 10 min at 3000–5000 x *g* to dry the membrane and ensure that no residual ethanol was left. The dried membranes were transferred to a new 50 mL tube, 0.75–1 mL of Buffer AE was added and incubated for 5 min at room temperature (15–25°C). The DNA was then eluted by centrifuging the membrane for 5 min at 3000–5000 x *g*.

3.4.2 Agarose gel electrophoresis

Agarose gel electrophoresis is the most common and easy way for analyzing and separating DNA PCR products. The concentration of agarose varies depending on the separation required. Generally 1.0% was used for DNA quantitation while 2.0% was used for resolving small PCR products.

After extraction and purification, the concentration of the DNA was estimated on a 1.0 % (w/v) agarose gel. An electrophoresis-grade 1% agarose gel was prepared by melting 1.8 g of agarose in 180 mL of 0.5x TBE buffer (Tris Borate EDTA, Appendix 2) in a microwave for approximately 2-3 min. The solution was then allowed to cool at room temperature for a couple of minutes, 1 µg/mL final concentration of ethidium bromide was added and swirled to mix. The ethidium bromide absorbs UV light and re-emits it as visible light when bound to DNA molecules by intercalating between bases (Muchugi *et al.*, 2008). The molten agarose was then poured into the supplied tray. Combs were inserted to make wells and the gel allowed to set for a minimum of 20-30

min at room temperature on a flat surface. After the gel had set, the tray was placed in a gel tank filled with 0.5x TBE buffer. The combs were removed and 5 µL of the DNA samples or PCR product with 2 µL of 6x loading Buffer (Appendix 2) added were loaded into the wells.

Electrophoresis was carried out at a constant 80 volts. The gel was then exposed to UV light using an image capture system (Bio-Rad documentation system, Bio-Rad Laboratories, Hercules, CA) and images were saved.

3.4.3 Quantitation of genomic DNA

The DNA samples were quantified by loading 5 µL of the DNA samples mixed with 2 µL of 6x loading Buffer into the gel wells. Standard samples of 10, 5, and 2.5 µL of Lambda (λ) (Promega) were also loaded alongside the samples and used as a reference to estimate the concentration of unknown DNA samples. The Lambda DNA samples had a total loading of 500 ng μL^{-1}, 250 ng μL^{-1} and 125 ng μL^{-1}.

DNA ladder 2-log (Appendix 3) was also used to check the integrity of the DNA. The samples were then separated on a 1.0 % (w/v) agarose gel and visualized with UV light (Section 3.4.2).

The genomic DNA concentration was estimated by comparing fluorescence between the test samples and the lambda control. After quantitation, the samples were then diluted to approximately 10 ng μL^{-1}. For further accuracy, 5 µL of the diluted samples was loaded onto a second 1 % (w/v) agarose gel electrophoresis and ran under the same conditions to allow concentration of the diluted PCR template to be confirmed. DNA solutions from IRD and other countries were also re-quantified and used directly or after dilution at 10 ng

μL^{-1} when needed. The genomic DNA was used for PCR amplification or stored at -20°C.

3.5 Microsatellite markers

3.5.1 Development of new microsatellite primers for genetic diversity

3.5.1.1 Microsatellite library construction

A new microsatellite-enriched library of date palm (*Phoenix dactylifera* L.) was constructed essentially as described by Kloda (2004) with some modifications based on the method of Edwards *et al.* (1996). The amplified microsatellite-enriched amplicons were sequenced by Roche 454 Pyrosequencing and two hundred simple sequence repeats were identified for microsatellite construction.

Preparation of filters

Small filters with a notch in the top right hand corner were made by cutting Hybond™ N+ Hybridisation Transfer Membrane into 1 cm^2 squares to make it easier to identify the DNA side of the membrane.

Ten microliters of the following oligonucleotides (MWG Biotech) at 1 µg/µl [$[GC]_{17}$, $[AC]_{17}$, $[GA]_{17}$, $[GT]_{17}$, $[CAA]_{10}$, $[GCC]_{10}$, $[CTG]_{10}$, $[CAG]_{10}$] were added to a 2 ml tube. The volume then made up to 1 ml with 3x SSC.

A total of 80 µl of the oligonucleotide mix was pipette onto the filters on Whatman 3M paper in 20 µl aliquots. The filters were then air dried for one hour. After that, they were covered with another sheet of 3M paper, wrapped in foil and dried for two hours at 80°C. Stratalinker™ 2400 (Stratagene) UV

Cross-linker was used to fix the oligonucleotides to the hybridisation membrane.

Four filters were then transferred to new 50 mL tubes and 30 mL of hybridisation buffer was added with gentle shaking and incubated for two days at 45°C. The filters were further washed with new hybridisation buffer for another two days.

Filters from the library were placed in 60 ml *0.1x SSC, 0.1%* SDS (w/v) to remove any unbound oligonucleotides and heated for 5 minutes at 95°C. Then, they were washed with *1x* SSC, wrapped with sheets of Saran film (Dow Chemical Company) and stored at –20°C.

Dilution and phosphorylation of adapters

Oligonucleotide sequences called MICRO ADAPT.1 (P) and MICRO ADAPT.2 were manufactured by MWG Biotech. They were diluted to 1 nmol/µl. MICRO ADAPT.1 (P) was also diluted to 1 µg/µl to be used as a blocking agent, and diluted to 400 ng/µl to be used as a primer.

MICRO ADAPT.1 (P): 5' – CTC TTG CTT ACG CGT GGA CTA – 3' 21-mer

MICRO ADAPT.2 : 5' –TAG TCC ACG CGT AAG CAA GAG CAC A – 3' 25-mer

A phosphorylation reaction for MICRO ADAPT.2 was performed in a total reaction volume of 20 µl containing 10 µl MICRO ADAPT.2 at 1 nmol/µl, 4

µl 5x exchange buffer (Invitrogen), 2 µl 100 mM rATP (Sigma), 1 µl T4 Polynucleotide Kinase (Invitrogen), 3 µl SDW and incubated for 1 hour at 37°C. The reaction was stopped by heating the tube for 30 minutes at 70°C. The phosphorylated MICRO ADAPT.2 was mixed with 10 µl of MICRO ADAPT.1 (P) and 10 µl SDW to produce a double-stranded DNA adapter mix, which were then heated to 70°C for 10 minutes and allowed to cool to room temperature for 15 minutes.

Digestion of genomic DNA

Genomic DNA of *Phoenix dactylifera* L. was digested using a four-base pair blunt cut-site restriction endonuclease (*Rsa*I) in a total reaction volume of 50 µl containing 5 µl 10x *Rsa*I buffer (Amersham Pharmacia), 3 µl *Rsa*I (10 units/µl, Amersham Pharmacia), 5 µl DNA (200 ng in 5 µl) and 37 µl SDW and incubated for 3 hours at 37°C to digest completely the DNA.

Ligation of adapters to digested DNA

The digested DNA was ligated with adapters in a reaction containing the following: 4.5 µl SDW, 40 µl *Rsa*I digested DNA, 1 µl 100mM rATP (Sigma), 1 µl prepared adapter mix and 1 µl T4 DNA ligase (Promega). The reaction was then incubated for two hours at 37°C. Ligation of the adapters to the *Rsa*I restriction fragments continued with the digestion.

PCR amplification of ligated fragments

PCR was performed to amplify the ligated fragments using the MICRO ADAPT.1 (P) primer, which is homologous to the adapters.

Three replicates of PCR reactions were set up in a 50 µl final volume containing; 1 µl ligated DNA, 1.5 mM $MgCl_2$, 5 µl 10X PCR buffer (Invitrogen), 0.2 mM dNTP (Invitrogen), 0.8 µg/µl MICROADAPT.1(P), 0.4 µl *Taq* DNA Polymerase (Invitrogen) and SDW . The PCR programme was 35 cycles of 40s at 95 °C, 60s at 60 °C, and 180s at 72 °C with a final elongation step of 10 min at 72 °C.

The amplification products were visualized by loading 10 µl of products on a 2% (w/v) agarose gel for electrophoresis. The PCR products were further pooled, purified with phenol: chloroform: isoamyl alcohol extraction and resuspended in 300 µl of SDW.

Enrichment for microsatellites

The purified amplicon DNA was denatured by heating 50 µl to 95°C for five minutes and immediately added to a new 2 ml tube containing one pre-washed filter, 2 µg of each adapter primer was added as blocking agent and 1 ml *hybridisation buffer*. The tube was incubated overnight at 50°C.

After hybridisation, filters were first washed with buffer I six times for five minutes followed by four washes with buffer II for five minutes. All washes should be at 60°C.

Elution of enriched DNA fragments

The enriched DNA was eluted after washing. Each filter was placed in a new tube 2 ml containing 500 µl SDW and was boiled for five minutes. The filters were then removed from the tube. The DNA was precipitated by addition of 12 µl of 5M NaCl and 1.5 ml absolute ethanol at -20°C and incubating on ice

for one hour. The tubes were then centrifuged for 30 minutes at 13,000 rpm using a Biofuge centrifuge (Heraeus) and the supernatant was discarded. The pellet was air-dried and then suspended in 25 µl SDW.

After elution, the enriched DNA was amplified by PCR using three replicates for each library. The reaction mix and PCR programme was the same as for the pre-enrichment amplification using 1 µl of eluted DNA as template. The amplification product of the three replicates was then visualized by loading 10 µl of products onto an agarose gel for electrophoresis. The products were pooled, cleaned with phenol: chloroform: isoamyl alcohol extraction method and resuspended in 35 µl SDW. After separation on an agarose gel, the fragments between 500 and 100 bp were excised and purified using QIAquick Gel Extraction Kit (QIAGEN) according to the manufacturers' instructions. The amplified insert was sent for pyrosequencing (Roche 454; MWG service; $1/16^{th}$ plate, non-Titanium reagents). Sequences were received as FASTA files and were screened for the presence of potential microsatellite repeats using the MISA.pl Perl script (http://pgrc.ipk-gatersleben.de/misa/misa.html) using the default settings.

Manufacture and screening of microsatellite primers

Out of the two hundred simple sequence repeats identified, forty-one pairs of SSR primers were synthesised; forward and reverse primers were designed flanking the motif regions of the microsatellite using the PRIMER3 software (Rozen and Skaletsky 2000; http://frodo.wi.mit.edu/Primer3/). The expected product size was limited to 200 bp, the melting temperature is fixed to be around 60 °C, and the length of the primers varied between 18 and 23 bases.

All other parameters in the PRIMER3 software were kept as default values. The primers were synthesized by Eurofins MWG Operon, UK.

Of the forty-one microsatellite loci identified as potentially suitable for development as molecular markers, only 13 amplified while 28 failed to work or amplify clear bands of expected size. The M13-extension tail CACGACGTTGTAAAACGAC was added at the 5' end of the forward primer for the thirteen microsatellites identified from the library and synthesized again by Eurofins MWG Operon, UK (Chapter 4). The sequences of all designed SSR primer pairs from the library are summarized in Appendix 1.

3.5.2 Screening and testing 30 published microsatellite (SSRs) primer pairs

Thirty SSRs primer pairs developed by Billotte *et al.* (2004) and by Akkak *et al.* (2009) with reportedly high levels of allelic diversity and clear and stable amplification were screened on DNA samples from eight parents of Omani date palm (Chapter 4).

3.5.3 Screening and testing 100 new microsatellite (SSRs) primer pairs

One hundred pairs of new SSRs primers were screened and tested with DNA samples from eight parents of Omani date palm. The sequences of these untested primers were kindly provided by the International Centre for Agricultural Research in the Dry Areas, Aleppo-Syria (ICARDA). These primers were generated from an assembly of the draft genome of date palm and published by Hamwieh *et al.* (2010) (Chapter 4).

3.6 Genetic relationships between Omani date palm and other germplasm using SSRs markers

A total of 12 microsatellite (SSRs) primer pairs were selected and used to study genetic relationships among one hundred ninety-four accessions from Omani germplasm and forty-eight accessions from different countries. These are listed in Chapter 5.

3.7 Construction of the genetic linkage map

A genetic linkage map of date palm was developed using the JoinMap4.1 software (Van Ooijen, 2006) by combining data from SSRs and SNPs markers. Two different populations (BC_1 and F_1) along with their parents were used as described in Section 3.3.3. Most palms have reached flowering stage and the gender is known for most of them (Chapter 6).

Microsatellite (SSRs) markers

Seventy-three selected microsatellite primer pairs were used for genome analysis of both populations, BC_1 and F_1.

SNPs markers

SNPs marker assays for the BC_1 and F_1 populations were performed by DArT Pty. Ltd (Yarralumla, Australia; *www.diversityarrays.com)* (Wenzl *et al.* 2004; Akbari *et al.* 2006; Semagn *et al.* 2006).

3.8 Development of new microsatellite primers (SSRs) for gender discrimination in date palm

Five new microsatellite primers were designed using the sequences of the three scaffolds that showed SNPs segregating in association with the sex determination locus in date palm as recently published by Al-Dous *et al.* (2011). The sequences for the three scaffolds were obtained from the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/) by using the following scaffold IDs (PDK_30s1150131, PDK_30s1038101, and PDK_30s1038231). The PCR primer pairs for microsatellite amplification were designed using the WebSat software (Martins *et al.*, 2009). They were designed to amplifying fragments ranging in size between 150 and 350 bp. Each sequence was subjected to a single PCR primer design except the sequence scaffold PDK_30s1038101, which was subjected to three PCR primer designs. An M13-extension tail 5'-CACGACGTTGTAAAACGAC-3' was added at the 5' end of each forward primer. All primers were then synthesized by Eurofins MWG Operon, UK and are described in Chapter 7.

3.9 Optimizing the annealing temperature for the microsatellite primers

An equimolar mixture of eight parental DNA samples (Khalas 4, Khalas 13 male, Um-Alsela, Khori male, Barni, Naghal, Bahlani male, and Khasab) was used to determine the best annealing temperature for each primer pair. A PCR reaction mix was prepared as follow:

A 100 µl of 10X standard reaction buffer (including Mg^{2+}), 8 µl of dNTPs mix (25mM), 5 µl of 0.5 U *Taq* DNA polymerase (Invitrogen), 1 µl of 200 µM

M13-Blue extension primer (CACGACGTTGTAAAACGAC), 100 µl of DNA template (mixture of eight DNA samples) and 686 µl of Sigma water to give a total volume of 900 µl. After mixing, the 900 µl of the PCR mix was aliquoted into four 1.5 ml tubes (225 µl per tube). For each tube, 25 µl of reverse primer (10X) and 2.5 µl of forward primer (10X) were added. The final mix was aliquoted into a 96-well PCR plate, 20 µl per well. The PCR amplifications were performed in a Thermal Cycler (GeneAmp PCR system 9700, Applied Biosystems, Singapore) with the following programme; an initial denaturation of 3 min at 94 °C, 35 cycles of 1 min at 94 °C, 50-60°C for 1 min (gradient 15°C) and 72 °C for 2 min, followed by a final elongation step of 72 °C for 10 min. The PCR products were resolved on 2% (w/v) agarose gels and visualized under UV light.

3.10 Polymorphic marker screening

PCR reactions were performed in a total reaction mixture of 20 µl containing: 10 ng of total cellular DNA (2 µl) as a template, 2 µl of 10X standard reaction buffer (including Mg^{2+}), 0.2 µl of dNTPs mix (25mM), 0.2 µl of 0.5 U *Taq* DNA polymerase (Invitrogen), 0.02 µl of 200 µl M13-Blue and/or M13-Green extension (CACGACGTTGTAAAACGAC), 0.2 µl of forward primer with M13 tail (10X), 2 µl of reverse primer (10X) and 13.38 µl of Sigma water. The amplifications were performed in a Thermal Cycler (GeneAmp PCR system 9700, Applied Biosystems, Singapore) with the following conditions; an initial denaturation of 5 min at 95 °C, 35 cycles of 30s at 95 °C, the specific annealing temperature for each SSR primer for 1 min, extension 72 °C for 1

min, followed by a final elongation step of 72 °C for 7 min. The amplification product was visualized under UV light using a 2% (w/v) agarose gel.

3.11 Fragment analysis using the CEQ 8000

Fragment analysis was carried out using the Beckman Coulter CEQ 8000 Genetic Analysis System (Beckman-Coulter, Fullerton, CA). 25 µl of sample loading solution (SLS) and 0.25 µl of the CEQ 400 size standard were aliquoted into each well in the CEQ sample plate. One µl of PCR product was added per well, layered with a drop of mineral oil and spun for 1 min prior to running on the CEQ system.

3.12 Data collection and analysis

The allelic size range for each marker was scored across all available samples and recorded as presence/absence. Data analysis was carried out using various software packages and each will be explained in more detailed in the relevant Chapters.

Chapter 4. SCREENING AND DEVELOPING NEW MICROSATELLITE MARKERS (SSRs) FOR DATE PALM (*PHOENIX DACTYLIFERA* L.)

4.1 Introduction

Date palm has preferably been propagated through clones, with morphological markers being used to try to ensure proper consistency of those individual cultivars. This approach is often unreliable and hard to evaluate due to trait-environment interactions (Elhoumaizi *et al.*, 2002; Bodian *et al.*, 2012). As a result the application of DNA technology has become a crucial aspect of genetic improvement of date palm for accurately identifying the cultivars and analyzing their genetic diversity and phylogenic relationships.

Within date palms, various molecular markers including isozymes (Benaceur *et al.*, 1991; Bendiab *et al.*, 1993 & 1997), RFLP (Corniquel and Mercier, 1994 & 1997), RAPD (Sedra *et al.*, 1998; El-Tarras *et al.*, 2007), AFLP (Elhoumaizi *et al.*, 2006), ISSR (Zehdi *et al.*, 2004; Karim *et al.*, 2010; Hamza *et al.*, 2012 and SSR (Al-Ruqaishi *et al.*, 2008; Ahmed and Al-Qaradawi, 2009, Elmeer *et al.*, 2011; Khierallah *et al.*, 2011b; Bodian *et al.*, 2012) have been used to assess the genetic diversity of many cultivars from different germplasm belonging to Morocco, Tunisia, Sudan, Iraq, Oman, Saudi Arabia, and California.

Microsatellite or simple sequence repeat (SSRs) molecular markers have been developed and used to study the genetic diversity of date palm and they have proven to be very powerful due to their locus-specificity, co-dominance, high reproducibility as well as revealing highly levels of polymorphism (Khanam *et*

al., 2012; Zehdi *et al.*, 2012). The first set of 16 date palm specific primer pairs for microsatellite amplification were developed from a $(GA)_n$ microsatellite-enriched library by Billotte *et al.* (2004), followed by another 17 SSRs markers developed by construction of two microsatellite enriched libraries of date palm, using $(GA)_n$ and $(GT)_n$ repeat sequence oligonucleotides (Akkak *et al.*, 2009).

More recently, Hamwieh *et al.* (2010) have reported the design of 1000 SSRs primer pairs. Thirty of these SSRs were used to investigate the genetic diversity of eleven date palm genotypes from Qatar (Elmeer *et al.*, 2011). Out of the thirty SSRs used, only ten were able to produce distinct polymorphic amplified SSR bands. Allele sizes ranged from 108 bp to 274 bp (Elmeer *et al.*, 2011). However, this number of SSRs markers is still insufficient to cover the entire genome and to give a comprehensive measurement of the genetic diversity of date palm.

In this study, we have developed and screened a set of high-quality microsatellite markers suitable for differentiation between and within date palm genotypes. We have also screened these SSRs for their ability to show polymorphism between the parents of available mapping crosses.

4.2 Plant material

Samples from mature leaves of eight Omani date palm representing the parents of the available controlled crosses (Table 4.1) were selected to screen 171 microsatellite primers, (13 SSRs primer pairs; Billotte *et al.*, 2004, 17 SSRs primer pairs; Akkak *et al.*, 2009,41 SSRs primer pairs; ourselves, 100 SSRs primer pairs; Hamwieh *et al.*, 2010).

Table 4-1: The eight Omani parents selected to screen 171 microsatellite primer pairs.

Sample No.	Accession Name	Gender
1	Khalas 4	Female
2	Khalas 13	Male
3	Um-Alsela	Female
4	Khori	Male
5	Barni	Female
6	Naghal	Female
7	Bahlani	Male
8	Khasab	Female

4.3 Data analysis

Molecular data were analyzed using the PowerMarker software (Version 3.25) (Liu and Muse, 2005). The allele number, gene diversity, polymorphism information content (PIC) and expected heterozygosity were calculated. The genetic distances between the cultivars were calculated according to Nei (1973).

Phenetic analyses were performed between the eight parents to obtain a clearer picture of the genetic relationships among them by means of the Neighbor-Joining clustering algorithm using PowerMarker Ver. 3.25 (Liu and Muse, 2005). Bootstrap analysis (1000 replicates) was performed to estimate stability and support for the inferred clades (Felsenstein, 1987). Bootstrap values were calculated using PowerMarker and visualized by MEGA software v.4 (Kumar *et al.*, 2004) with confidence limits placed at the major nodes.

4.4 Results

4.4.1 Quantitation of genomic DNA for the eight parents

The DNA concentration of the eight parents was estimated (Figure 4.1) and the required dilutions (10 ng μL^{-1}) made; Table 4.2. The dilution was further confirmed by running a 5 µL aliquot of each sample on a new gel (Figure 4.2).

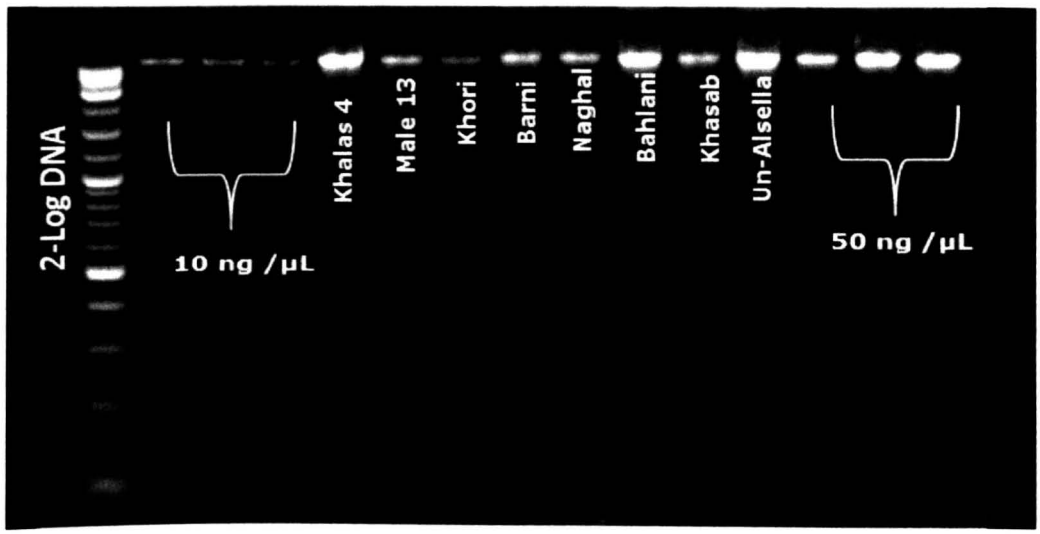


Figure 4.1: Quantitative estimation of genomic DNA concentration of the eight parents on agarose gel electrophoresis using Lambda DNA (λ) as a comparator and 2-log DNA ladder.

Table 4-2: Dilution factor for the eight parents of date palm estimated from photographed DNA using Lambda DNA (λ) 10 ng/µl and 50 ng/µl as the comparator .

Sample No.	Genotypes Name	DNA Conc. (ng/µl)	Dilution (µl)	Final Volume (µl)
1	Khalas 4	100	5 µl DNA + 45 µl H_2O	50
2	Male 13	25	25 µl DNA + 25 µl H_2O	50
3	Khori	10	50 µl DNA	50
4	Barni	25	25 µl DNA + 25 µl H_2O	50
5	Naghal	25	25 µl DNA + 25 µl H_2O	50
6	Bahlani	100	5 µl DNA + 45 µl H_2O	50
7	Khasab	25	25 µl DNA + 25 µl H_2O	50
8	Um-Alsella	100	5 µl DNA + 45 µl H_2O	50
Final DNA Concentration = 10ng/ µl				

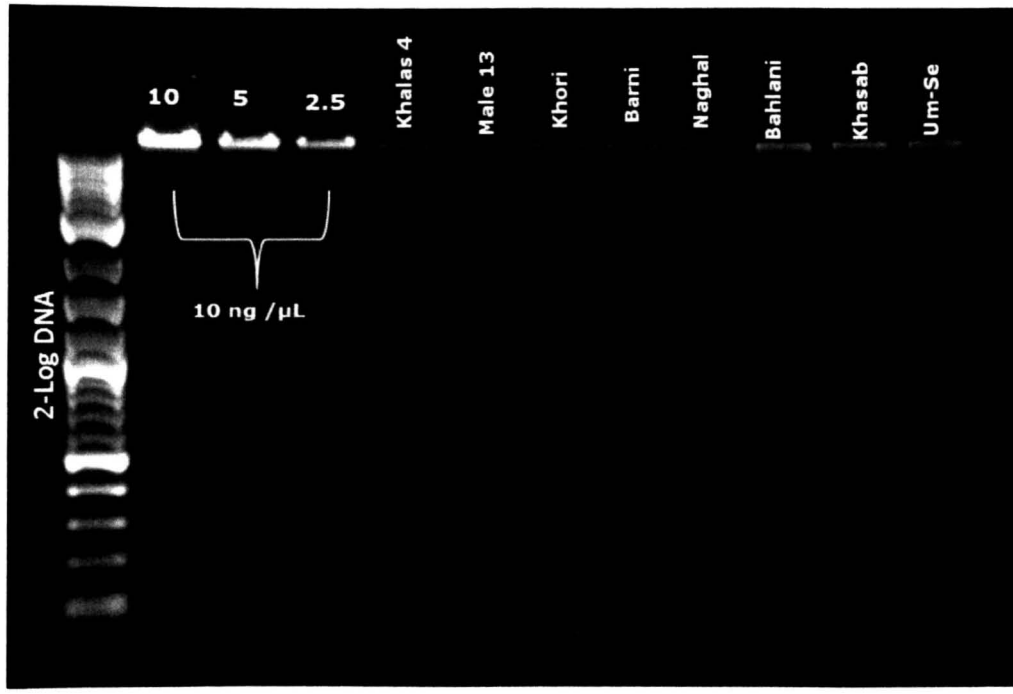


Figure 4.2: Final DNA concentration (10ng/µl) for the eight parents as presented in Table 4.2.

4.4.2 Development of new microsatellite primers

A genomic enriched-library for date palm (*Phoenix dactylifera* L.) was constructed as described in Chapter 3 (Section 3.5.1.1) and sequence data generated using Roche 454 Pyrosequencing (non-Titanium reagents). Sequences containing microsatellite repeat motifs were identified using the MISA.pl script and forward and reverse primers were designed around sequence data flanking the motif using the PRIMER 3 software (Rozen and Skaletsky 2000; http://frodo.wi.mit.edu/Primer3/). Out of the two hundred simple sequence repeats identified, the best 41 sequences were selected. The SSRs primer length varied between 18 and 27 bases with an optimal melting temperature around 60 °C.

The 41 primers were initially tested to determine if they could amplify date palm genomic DNA. A total of thirteen (31.7%) of the primer pairs amplified clear bands; whereas 28 SSRs did not amplify the date palm DNA to give clear single bands of around the expected size.

The PCR product of the amplified DNA was further analyzed using the CEQ 8000 Genetic Analyzer in order to determine monomorphic and polymorphic primers. The results indicate that eleven of the thirteen primers generated polymorphic fragments when amplified from the available eight parental samples (Figure 4.3). These primers are: DateS1, DateS8, DateS9, DateS12, DateS16, DateS17, DateS103, DateS110, DateS111, DateS130, and DateS131. However two primers; DateS21 and DateS138 produced monomorphic fragments only. The observed sizes for these primers and working temperature are summarized in Table 4.3.

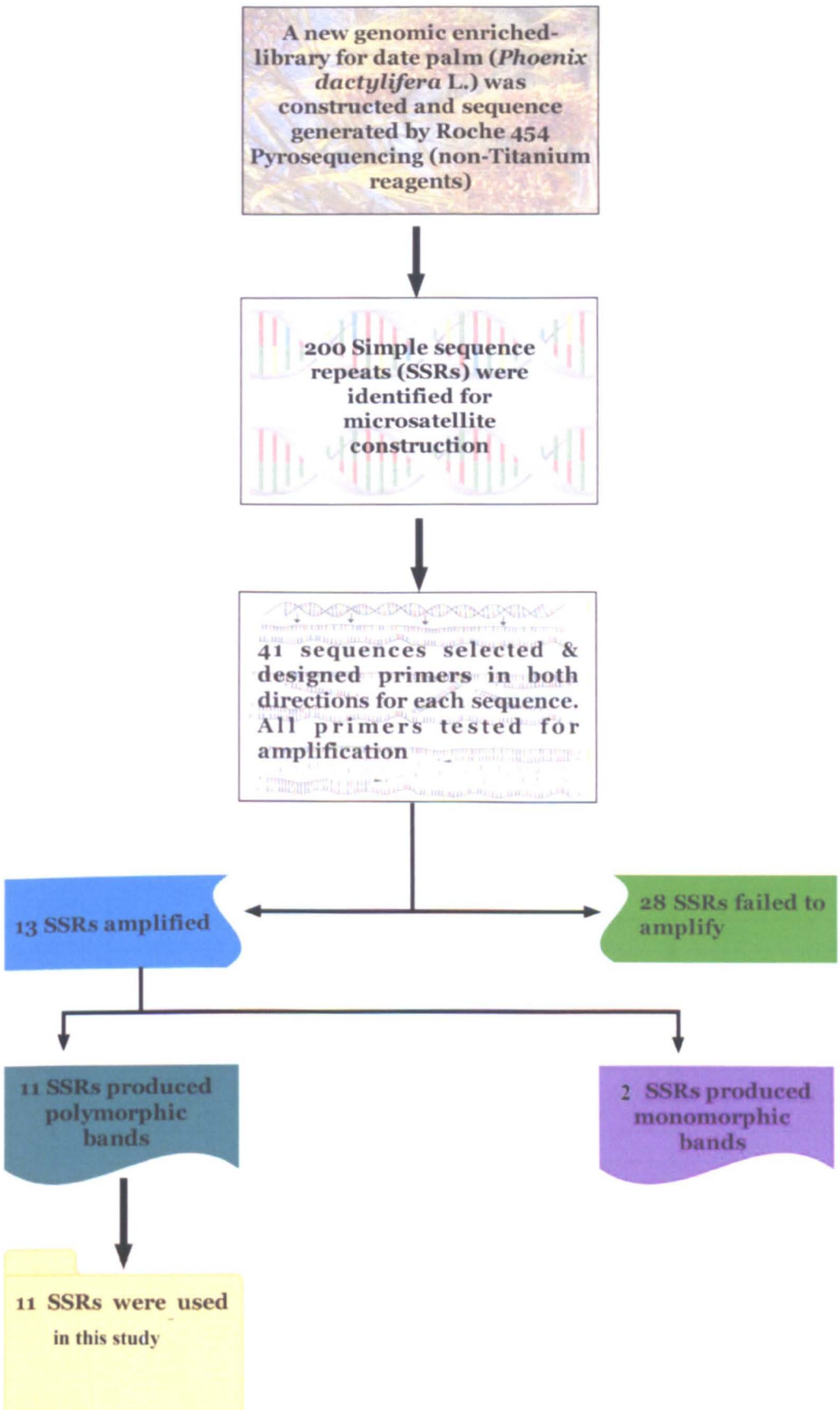


Figure 4.3: A flow chart illustrating the process for developing and selection of the eleven new polymorphic SSR.

Table 4-3: Lists the 41 microsatellite primers designed for date palm, marker name, annealing temperature Tm (°), motif repeat, observed allelic size range (bp) and status of amplification.

Marker name	Annealing T_m (°)	Motif repeat	Observed size (bp)	Status of amplification
DateS1	50	$(TC)_{12}$	137-151	++
DateS2	50	$(CTC)_{16}$	-	-
DateS3	52	$(GA)_8$	-	-
DateS4	50	$(CT)_{18}$	-	-
DateS8	55	$(GA)_{10}$	195-199	++
DateS9	55	$(CT)_{10}$	188-224	++
DateS10	55	$(CCG)_5$	-	-
DateS11	55	$(CT)_{19}$	-	-
DateS12	55	$(GA)_{22}$	158-173	++
DateS13	55	$(GT)_{17}$	-	-
DateS14	55	$(GA)_{29}$	-	-
DateS15	55	$(GAA)_6$	-	-
DateS16	55	$(GCC)_8$	126-139	++
DateS17	55	$(CA)_{11}$	167-183	++
DateS18	55	$(CTC)_{13}$	-	-
DateS19	50	$(GA)_{28}$	-	-
DateS21	52	$(CTC)_{16}$	204-214	+
DateS22	50	$(GA)_{12}$	-	-
DateS23	52	$(GA)_{17}$	-	-
DateS24	52	(TCC)4	-	-
DateS25	52	(GA)19	-	-
DateS26	52	$(CT)_{22}$	-	-
DateS41	50	$(GA)_6$	-	-
DateS57	50	$(TC)_7$	-	-
DateS60	50	$(CA)_5$	-	-
DateS78	55	$(CA)_{12}$	-	-
DateS84	55	$(GA)_{18}$	-	-
DateS88	60	$(GA)_8$	-	-
DateS90	60	$(CA)_8$	-	-
DateS100	55	$(CGG)_5$	-	-
DateS103	55	$(GA)_{15}$	203-206	++
DateS110	52	$(CT)_{10}$ $(GT)_9$	200-207	++
DateS111	52	$(GT)_{11}$	154-177	++
DateS116	52	$(ATT)_3$	-	-
DateS120	55	$(TTC)_5$	-	-
DateS130	52	$(CT)_{11}$	183-198	++
DateS131	52	$(GA)_{13}$	195-200	++
DateS137	52	$(CGC)_6$	-	-
DateS138	52	$(GA)_{10}$	170-172	+
DateS176	55	$(CAT)_5$	-	-
DateS185	55	$(CT)_7$	-	-

++; amplification of polymorphic band, +; amplification of monomorphic band, -; no amplification

The gene diversity, or expected heterozygosity, is a frequently used measure of genetic variation applied in diverse areas of population genetics (DeGiorgio and Rosenberg 2009), and was ranged from 0.33 to 0.75 with an average of 0.50 using the eleven new SSRs markers amplified with the eight cultivars. The eleven primers were able to detect a total of 42 alleles with a mean of 3.82 alleles per locus. The number of alleles varied from 2 alleles for primers DateS1 and DateS130 and 6 alleles in primer DateS111. In addition, the mean major allele frequency was 0.64 and ranged from 0.38 to 0.81. The heterozygosity for the eight cultivars was found to be 0.38 and the polymorphism information contents PIC value for the markers ranged from 0.30 to 0.71 for DateS1 and DateS110, respectively with an average of 0.45 (Table 4.4).

The average genetic distance among the eight cultivars varied from 0.09 to 0.50 (Table 4.5). The highest genetic distances value (0.50) was recorded between Khalas4 and Barni, while lowest value (0.09) was between Khori male and Naghal, followed by Khalas4 and Khalas13 with a genetic distance value of 0.14.

Table 4-4: Marker name, major allele frequency, number of genotypes identified, number of alleles generated, gene diversity, heterozygosity and polymorphism information content (PIC) using the eight parents of Omani date palm amplified with eleven new SSR markers.

Marker	Major allele frequency	Genotype no.	Allele no.	Gene diversity	Heterozygosity	PIC
DateS1	0.75	3	2	0.38	0.25	0.30
DateS8	0.75	4	3	0.40	0.25	0.35
DateS9	0.81	3	4	0.33	0.25	0.31
DateS12	0.75	4	4	0.41	0.25	0.39
DateS16	0.81	4	4	0.33	0.38	0.31
DateS17	0.63	5	4	0.54	0.38	0.48
DateS103	0.50	3	3	0.59	0.25	0.51
DateS110	0.38	6	5	0.75	0.88	0.71
DateS111	0.38	5	6	0.73	0.25	0.69
DateS130	0.69	2	2	0.43	0.63	0.34
DateS131	0.63	5	5	0.57	0.38	0.54
Mean	0.64	4	3.82	0.50	0.38	0.45

Table 4-5: The average genetic distance based on Nei (1973) for eight Omani cultivars (Khs4: Khalas4, Kh13: Khalas13, Ums: Um-Alsela, KhrM: Khori male, Barni, Nag: Naghal, Bah: Bahlani male, Khb: khasab) using eleven new SSR markers.

Parents ID	Khs4	Kh13	Ums	KhrM	Barni	Nag	Bah	Khb
Khs4	0.00							
Kh13	0.14	0.00						
Ums	0.32	0.16	0.00					
KhrM	0.48	0.43	0.45	0.00				
Barni	0.50	0.34	0.36	0.30	0.00			
Naghal	0.45	0.36	0.39	0.09	0.18	0.00		
Bah	0.39	0.39	0.41	0.43	0.41	0.32	0.00	
Khb	0.50	0.39	0.39	0.45	0.30	0.25	0.32	0.00

4.4.3 Screening and testing of 30 published microsatellite (SSRs) primer pairs

Thirty microsatellite primer pairs reported by Billotte *et al.* (2004) and Akkak *et al.* (2009) were tested on the eight parents of Omani date palm (Table 4.6). All SSRs amplified successfully except for markers mPdCIR044 and mPdCIR048. Nineteen markers (63.3%) showed polymorphic fragments among the eight cultivars, whereas nine generated monomorphic fragments. The polymorphic primers as presented in Table 4.6 and are: mPdCIR010, mPdCIR015, mPdCIR016, mPdCIR025, mPdCIR050, mPdCIR057, mPdCIR078, mPdCIR085, mPdCIR093, PDCAT2, PDCAT5, PDCAT10, PDCAT11, PDCAT12, PDCAT14, PDCAT17, PDCAT18, PDCAT20 and PDCAT21.

Table 4-6: Lists of 30 microsatellite primers designed for date palm by Billotte *et al.* (2004) and Akkak *et al.* (2009), marker name, annealing temperature Tm (°), motif repeat, observed allelic size range (bp) and status of amplification

Marker name	Annealing T_m (°)	Motif repeat	Observed size (bp)	Status of amplification
mPdCIR010	52°C	$(GA)_{22}$	141-181	++
mPdCIR015	52°C	$(GA)_{15}$	143-156	++
mPdCIR016	52°C	$(GA)_{14}$	151-157	++
mPdCIR025	52°C	$(GA)_{22}$	224-236	++
mPdCIR044	52°C	$(GA)_{19}$	-	-
mPdCIR048	52°C	$(GA)_{32}$	-	-
mPdCIR050	52°C	$(GA)_{21}$	191-224	++
mPdCIR057	52°C	$(GA)_{20}$	273-283	++
mPdCIR070	52°C	$(GA)_{17}$	208	+
mPdCIR078	52°C	$(GA)_{13}$	146-165	++
mPdCIR085	52°C	$(GA)_{29}$	174-200	++
mPdCIR090	52°C	$(GA)_{26}$	177-190	+
mPdCIR093	52°C	$(GA)_{16}$	160-181	++
PDCAT1	55°C	$(TC)_{21}$	164-167	+

PDCAT2	55°C	CTCGCTG$(TC)_3(TC)_3$T $(TC)_3T(TC)_3T(TC)$4TTCT GTCCCG$(TC)_{16}T(TC)$	152-179	++
PDCAT3	55°C	$(CA)_8$ - $(GT)_3(CA)_4$	192-130	+
PDCAT4	55°C	$(CA)_8TT(CA)_4(GA)_{20}$	249-261	+
PDCAT5	55°C	$(AG)_{16}$	155-183	++
PDCAT6	55°C	$(CA)_{14}(GA)_{23}$	104-106	+
PDCAT8	55°C	$(TC)_{16}$	220	+
PDCAT10	55°C	$(TC)_{16}$	243-247	++
PDCAT11	55°C	$(TC)_7(TC)_{20}$	184-215	++
PDCAT12	55°C	$(CT)_{19}$	105-124	++
PDCAT13	55°C	$(GA)_{21}$GCA(GGA)GA $(GGA)_3$	181-195	+
PDCAT14	55°C	$(TC)_{19}(TC)_{16}$	135-168	++
PDCAT15	55°C	$(GA)_{13}$-$(GA)_8(GA)_6$	144-148	+
PDCAT17	55°C	$(GA)_{21}$	144-158	++
PDCAT18	55°C	$(CT)_{13}G(CT)_8CG(CT)_3$ $CG(CT)_3$	126-169	++
PDCAT20	55°C	$(GA)_{29}$	343-362	++
PDCAT21	55°C	$(GA)_5T(GA)_2TA(GA)_2$ $GC(GA)_5(GT)7$	163-170	++

++, amplification of polymorphic band, +, amplification of monomorphic band, - no amplification

Nineteen polymorphic markers generated a total of 137 alleles, ranging from 2 to 10 alleles per locus, with an average of 7.21 alleles per locus. High levels of heterozygosity at 0.60 were detected among the eight cultivars with marker average polymorphism information content (PIC) of 0.75, ranging from 0.36 to 0.87 at locus PDCAT10 and PDCAT11, respectively. The major allele frequency at each locus ranged from 0.19 to 0.63, with an average of 0.34. The gene diversity or expected heterozygosity varied from 0.47 at locus PDCAT10 to 0.88 at loci PDCAT11 and PDCAT18, with a mean value of 0.78 (Table 4.7).

Table 4-7: Marker name, major allele frequency, number of genotypes identified, number of alleles generated, gene diversity, heterozygosity and polymorphism information content (PIC) using the eight parents of Omani date palm amplified with the nineteen SSR markers developed by Billotte *et al.* (2004) and Akkak *et al.* (2009).

Marker	Major allele frequency	Genotype no.	Allele no.	Gene diversity	Heterozy-gosity	PIC
mPdCIR010	0.31	7	9	0.83	0.75	0.81
mPdCIR015	0.38	6	7	0.78	1.00	0.75
mPdCIR016	0.31	5	6	0.79	0.88	0.76
mPdCIR025	0.44	5	5	0.70	0.50	0.66
mPdCIR050	0.44	7	7	0.73	0.63	0.69
mPdCIR057	0.50	4	5	0.68	0.50	0.64
mPdCIR078	0.25	7	9	0.84	0.63	0.83
mPdCIR085	0.19	8	9	0.87	0.50	0.85
mPdCIR093	0.31	6	6	0.78	0.25	0.75
PDCAT2	0.38	7	10	0.81	0.75	0.80
PDCAT5	0.25	7	8	0.83	0.38	0.81
PDCAT10	0.63	2	2	0.47	0.00	0.36
PDCAT11	0.19	7	10	0.88	1.00	0.87
PDCAT12	0.50	5	5	0.66	0.25	0.62
PDCAT14	0.19	8	9	0.87	0.75	0.85
PDCAT17	0.31	7	7	0.80	0.63	0.78
PDCAT18	0.19	7	10	0.88	0.75	0.86
PDCAT20F	0.38	7	7	0.78	0.63	0.75
PDCAT21F	0.25	6	6	0.80	0.63	0.77
Mean	0.34	6	7.21	0.78	0.60	0.75

The average genetic distance was computed based on Nei (1973) and ranged from 0.24 to 0.80 between the eight cultivars using these 19 SSR markers (Table 4.8). The smallest genetic distance value (0.24) was calculated between Khalas 4 and Khalas 13, followed by 0.33 Khalas 13 and Um-Alsela (Table 4.8). Cultivar Naghal was found to have the greatest genetic distance (0.80) compared with Khalas 13, followed by 0.78 between Naghal and Khalas 4 and 0.76 between Naghal and Um-Alsela. However, Naghal seemed to be genetically close to the Khori male giving the lowest value of 0.30 as compared to those of other cultivars tested.

Table 4-8: The average genetic distance based on Nei (1973) for eight Omani cultivars (Khs4: Khalas4, Kh13: Khalas13, Ums: Um-Alsela, KhrM: Khori male, Barni, Nag: Naghal, Bah: Bahlani male, Khb: khasab) using nineteen SSR markers developed by Billotte *et al.* (2004) and Akkak *et al.* (2009).

Parents ID	Khs4	Kh13	Ums	KhrM	Barni	Nag	Bah	Khb
Khs4	0.00							
Kh13	0.24	0.00						
Ums	0.34	0.33	0.00					
KhrM	0.68	0.67	0.64	0.00				
Barni	0.58	0.59	0.57	0.46	0.00			
Nag	0.78	0.80	0.76	0.30	0.50	0.00		
Bah	0.64	0.64	0.62	0.47	0.36	0.55	0.00	
Khb	0.62	0.62	0.62	0.50	0.46	0.49	0.47	0.00

4.4.4 Screening and testing 100 new microsatellite (SSRs) primer pairs

One-hundred new SSRs primer pairs sequences were obtained from the International Centre for Agricultural Research in the Dry Areas (ICARDA), synthesised and screened against the eight parents of Omani date palm crosses. Seventy-nine percent of primers amplified successfully while 21% primers failed to amplify genomic DNA of date palm (Table 4.9) at annealing temperatures which ranged from 55°C to 61°C. The DNA bands of date palm amplified by the 79 primers were analysed on the Beckmann CEQ 8000. From this analysis 42 were polymorphic and 37 monomorphic when amplified from the eight parental cultivars. The successful markers produced clear amplified SSR bands ranging in size from 110 bp to 356 bp.

Table 4-9: List of 100 new microsatellite primers designed for date palm, marker name, annealing temperature Tm (°), motif repeat, observed allelic size range (bp) and status of amplification

Marker name	Annealing T_m (°)	Motif repeat	Observed size (bp)	Status of amplification
DPALM301	55°C	(TAAA)5	202	+
DPALM302	55°C	(ATTT)5	225-230	++
DPALM303	55°C	(TATG)5	183-193	++
DPALM305	55°C	(AAAG)5	224-229	++
DPALM306	55°C	(AGAT)5	222-225	+
DPALM307	58°C	(ATTT)11	193-204	++
DPALM308	60°C	(TTA)20	-	-
DPALM309	58°C	(TATC)6	189-213	++
DPALM310	59°C	(GA)24	-	-
DPALM311	50°C	(TACA)6	199-204	++
DPALM312	58°C	(GAA)8	205-212	++
DPALM315	58°C	(ATG)8	266-278	++
DPALM316	57°C	(CTTG)5	222-225	+
DPALM317	60°C	(CAAA)5	-	-
DPALM318	55°C	(TAA)8	217	+
DPALM319	55°C	(TG)30	167-182	++
DPALM320	55°C	(TTAT)9	207	+
DPALM321	55°C	(CTT)8	229	+
DPALM322	57°C	(AGG)9	248	+
DPALM323R	55°C	(TACA)6	189	+
DPALM324	55°C	(GA)24	198-209	+
DPALM325	55°C	(AT)23	188-194	++
DPALM326	55°C	(TTC)8	219	+
DPALM327	55°C	(ATCT)6	238-255	++
DPALM328	55°C	(TG)31	157-204	++
DPALM329	60°C	(TC)27	-	-
DPALM331	55°C	(ATGT)6	230-234	+
DPALM332	55°C	(ACAG)6	168-178	++
DPALM333	55°C	(ACAT)5	274-290	++
DPALM335	60°C	(ATGT)5	-	-
DPALM336	55°C	(GA)22	178-195	++
DPALM337	50°C	(TTA)12	160	+
DPALM338	59°C	(TC)27	-	-
DPALM339	55°C	(ATGT)5	273	+
DPALM340	55°C	(GAA)8	219-221	++
DPALM341	55°C	(CT)29	148-178	++
DPALM342	55°C	(GCT)8	195-224	++
DPALM343	58°C	(AAT)9	155-194	++
DPALM344	58°C	(TATG)5	227-231	++
DPALM345	59°C	(TTA)13	-	-
DPALM346	58°C	(TTA)8	227-228	+
DPALM347	57°C	(GAG)8	208	+
DPALM348	55°C	(ATT)8	209-223	++
DPALM349	55°C	(CA)34	181-196	++
DPALM350	61°C	(CT)47	171-178	++
DPALM351	58°C	(TGCA)10	215-217	+
DPALM352	55°C	(GAAG)6	220-224	++
DPALM353	55°C	(AAC)8	198-200	+
DPALM354	55°C	(TTTC)5	199-200	+

DPALM355	55°C	(TATG)5	268-272	+
DPALM356	60°C	(TC)36	-	-
DPALM357	57°C	(CTTT)5	272-280	++
DPALM358	55°C	(AAAT)6	213-215	+
DPALM359	60°C	(AG)27	-	-
DPALM360	58°C	(CCT)9	187	+
DPALM361	55°C	(AAT)9	138-169	++
DPALM362	57°C	(TC)29	326-356	++
DPALM363	58°C	(AAT)9	122-156	++
DPALM364	55°C	(TACA)5	212	+
DPALM365	55°C	(CAAA)5	210-212	+
DPALM366	57°C	(ATA)9	227-234	++
DPALM367	60°C	(TTA)13	-	-
DPALM368	55°C	(TGGT)6	214	+
DPALM369	55°C	(ATGG)6	209-227	++
DPALM370	55°C	(GTTT)6	224	+
DPALM371	57°C	(TTTA)14	179-192	+
DPALM372	55°C	(GTTT)6	210-219	+
DPALM373	60°C	(AG)27	-	-
DPALM374	55°C	(CA)39	192-196	++
DPALM375	55°C	(AGA)10	173	+
DPALM376	55°C	(TTGC)5	226	+
DPALM377	61°C	(AGG)9	246-251	++
DPALM378	55°C	(CAT)8	227-234	++
DPALM379	55°C	(TGTT)5	198-201	++
DPALM380	55°C	(TATC)6	211-215	++
DPALM381	55°C	(TC)54	110-111	+
DPALM383	60°C	(TG)23	-	-
DPALM385	60°C	(TGCT)7	-	-
DPALM386	60°C	(AG)30	-	-
DPALM387	60°C	(ATGT)5	-	-
DPALM388	55°C	(TTA)15	240-253	++
DPALM389	55°C	(AAG)8	227	+
DPALM390	57°C	(ATT)13	208-210	+
DPALM391	55°C	(GATA)7	113-139	+
DPALM392	55°C	(TCC)8	187	+
DPALM393	60°C	(TC)44	-	-
DPALM394	61°C	(ACC)8	230	+
DPALM395	60°C	(GA)22	-	-
DPALM397	60°C	(CT)29	-	-
DPALM398	50°C	(CTTT)5	210-215	++
DPALM399	59°C	(CA)39	-	-
DPALM400	57°C	(CT)29	224	+
DPALM401	55°C	(ATT)8	210-211	+
DPALM402	55°C	(GTAC)6	197-202	++
DPALM403	60°C	(GA)36	-	-
DPALM404	55°C	(TCA)9	151-155	++
DPALM405	55°C	(TTTC)5	223-224	++
DPALM407	60°C	(TTA)20	-	-
DPALM408	55°C	(ATGC)5	229-233	++
DPALM410	55°C	(AAG)14	205-209	++

++, amplification of polymorphic band, +, amplification of monomorphic band, - no amplification

The allele sizes for the forty-two polymorphic SSRs markers were further analysed using the PowerMarker software (Version 3.25). A total of 190 alleles with an average of 4.52 alleles per locus were observed, with the number of alleles varying from 2 to 8 per locus. The major allele frequencies ranged from 0.25 to 0.88 with a mean of 0.47. The 42 primers detected an average heterozygosity of 0.42 and an average PIC value of 0.59, ranging from 0.19 to 0.81 for loci DPALM344 and DPALM333, respectively. Gene diversity among the eight cultivars averaged 0.64 and ranged from 0.22 in locus DPALM344 to 0.83 in locus DPALM333 (Table 4.10).

Table 4-10: Marker name, major allele frequency, number of genotypes identified, number of alleles generated, gene diversity, heterozygosity and polymorphism information content (PIC) using the eight parents of Omani date palm amplified with 42 new SSR markers designed by Hamwieh *et al.* (2010).

Marker	Major allele frequency	Genotype no.	Allele no.	Gene diversity	Heterozy-gosity	PIC
DPALM302	0.63	3	3	0.53	0.00	0.47
DPALM303	0.38	6	5	0.74	0.75	0.70
DPALM305	0.56	3	3	0.54	0.13	0.45
DPALM307	0.63	3	3	0.53	0.00	0.47
DPALM309	0.69	3	3	0.46	0.13	0.40
DPALM311	0.25	5	5	0.78	0.75	0.75
DPALM312	0.50	5	4	0.66	0.13	0.62
DPALM315	0.63	4	3	0.53	0.25	0.47
DPALM319	0.63	4	4	0.55	0.38	0.51
DPALM325	0.50	5	5	0.66	0.25	0.62
DPALM327	0.38	6	6	0.75	0.88	0.71
DPALM328	0.44	7	8	0.75	0.88	0.73
DPALM332	0.50	4	4	0.66	0.50	0.60
DPALM333	0.25	8	7	0.83	0.63	0.81
DPALM336	0.31	5	6	0.78	0.88	0.75
DPALM340	0.63	3	3	0.53	0.50	0.47
DPALM341	0.31	7	8	0.80	0.88	0.78
DPALM342	0.81	2	2	0.30	0.38	0.26
DPALM343	0.50	3	3	0.59	0.25	0.51
DPALM344	0.88	2	2	0.22	0.25	0.19
DPALM348	0.56	5	3	0.59	0.38	0.52
DPALM349	0.50	5	5	0.68	0.50	0.64
DPALM350	0.31	7	7	0.81	0.75	0.79
DPALM352	0.50	3	2	0.50	0.75	0.38

DPALM357	0.50	4	4	0.66	0.25	0.60
DPALM361	0.31	7	7	0.77	0.75	0.74
DPALM362	0.25	6	7	0.82	0.38	0.80
DPALM363	0.38	4	4	0.73	0.63	0.68
DPALM366	0.31	6	5	0.75	0.25	0.71
DPALM369	0.44	7	8	0.76	0.75	0.74
DPALM374	0.50	4	3	0.55	0.75	0.46
DPALM377	0.25	5	5	0.78	0.00	0.75
DPALM378	0.31	7	6	0.78	0.50	0.75
DPALM379	0.38	5	5	0.73	0.25	0.68
DPALM380	0.69	4	4	0.49	0.13	0.46
DPALM388	0.38	4	3	0.66	0.38	0.59
DPALM398	0.50	4	5	0.64	0.25	0.58
DPALM402	0.38	6	5	0.74	0.50	0.70
DPALM404	0.50	4	4	0.60	0.25	0.53
DPALM405	0.63	3	3	0.53	0.25	0.47
DPALM408	0.56	3	3	0.54	0.13	0.45
DPALM410	0.38	5	5	0.73	0.13	0.68
Mean	0.47	4.67	4.52	0.64	0.42	0.59

The average genetic distance among the eight parents of Omani date palm was calculated to be between 0.08 and 0.64. The highest value was detected between Khalas 4 and Bahlani as well as Um-Alsela and Khasab. However, Khalas 4 and Khalas 13 showed the lowest genetic distance (0.08) followed by Khalas13 and Um-Alsela (0.21) (Table 4.11).

Table 4-11: The average genetic distance based on Nei (1973) for eight Omani cultivars (Khs4: Khalas4, Kh13: Khalas13, Ums: Um-Alsela, KhrM: Khori male, Barni, Nag: Naghal, Bah: Bahlani male, Khb: khasab) using the 42 new SSR markers developed from primer data supplied by Hamwieh *et al.* (2010).

Parents ID	Khs4	Kh13	Ums	KhrM	Barni	Nag	Bah	Khb
Khs4	0.00							
Kh13	0.08	0.00						
Ums	0.26	0.21	0.00					
KhrM	0.58	0.52	0.54	0.00				
Barni	0.57	0.54	0.58	0.42	0.00			
Nag	0.54	0.52	0.49	0.38	0.46	0.00		
Bah	0.64	0.63	0.62	0.52	0.48	0.43	0.00	
Khb	0.60	0.58	0.64	0.61	0.56	0.49	0.41	0.00

4.4.5 Analysis of eight parents of Omani date palm using the combined 72 polymorphic SSR primer pairs

The data from the different primer sources was combined and the 72 microsatellite primer pairs used to analyse genetic variation in the eight parental cultivars of Omani date palm. This resulted in a total of 369 alleles with an average of 5.13 alleles per locus. The number of alleles per locus ranged from 2 for locus PDCAT10 to 10 for loci PDCAT2, PDCAT11 and PDCAT18. The average heterozygosity of the eight cultivars was 0.46. The PIC varied from 0.19 to 0.87 at loci DPALM344 and PDCAT11, respectively, with a mean of 0.61. In addition, the major allele frequency ranged from 0.19 to 0.88 with an average of 0.46. The gene diversity or expected hetrozygosity was 0.66 varying from 0.22 for locus DPALM344 to 0.88 for loci PDCAT11 and PDCAT18, (Appendix 4).

The average genetic distances were also estimated for the eight cultivars using the combined data from 72 SSR markers and are summarized in Table 4.12. The analysis revealed that Khalas4 is highly divergent from Bahlani male with a genetic distance value of 0.60, but closely related to Khalas13 at 0.13. In addition, both Khalas4 and Khalas13 show low value of genetic distances with Um-Alsela 0.29 and 0.24, respectively.

Table 4-12: The average genetic distance based on Nei (1973) for eight Omani cultivars (Khs4: Khalas4, Kh13: Khalas13, Ums: Um-Alsela, KhrM: Khori male, Barni, Nag: Naghal, Bah: Bahlani male, Khb: khasab) using the combined data from 72 SSR markers produced by Billotte *et al.* (2004), Akkak *et al.* (2009), ourselves and from primer sequences supplied by Hamwieh *et al.* (2010).

Parents ID	Khs4	Kh13	Ums	KhrM	Barni	Nag	Bah	Khb
Khs4	0.00							
Kh13	0.13	0.00						
Ums	0.29	0.24	0.00					
KhrM	0.59	0.55	0.56	0.00				
Barni	0.56	0.52	0.55	0.41	0.00			
Nag	0.59	0.57	0.55	0.32	0.43	0.00		
Bah	0.60	0.59	0.59	0.49	0.44	0.45	0.00	
Khb	0.59	0.56	0.59	0.56	0.49	0.45	0.41	0.00

A phenetic analysis reflecting the simplest relationship, revealed significant divergence between the studied eight Omani date palm parents using different sub-sets of SSRs markers and four phenetic trees were generated (Figure 4.4). The reliability of phenetic trees was estimated using bootstrapping. Bootstrap values are shown at the appropriate nodes (Fig. 4.4). Bootstrapping can be used estimate the confidence in each clade of NJ (Neighbour-Joining) tree (Kloda, 2004).

Phenetic tree 'A' revealed Khalas4, Khals13 and Um-Alsela clustered together in one group, whereas Naghal, Khori male and Barni were in a second group. Bahlani male and Khasab appeared to be two outgroups with respect to the other samples (Fig. 4.4 A). The clustering of Naghal and Khori male with Barni was significantly supported with a bootstrap value of 99 %, although other bootstrap values were non-significant. Khalas4 and Khalas13 showed a bootstrap value of 79% with Um-Alsela. Tree 'B' showed two main cluster groups based on data from 19 SSRs markers. The first group contained

Bahlani, Barni, Khori male, Nagal and Khasab, while the second group included Khalas 13, Khalas 4 and Um_Alsela (Fig. 4.4 B). This tree had very high bootstrap values compare to tree 'A'. The clustering of both Khalas 4 and Khalas13 with Um-Alsela was significantly supported with a bootstrap value of 86 %. The clustering of Naghal and Khori male was also significantly supported with a bootstrap value of 98 %, similar to tree 'A'.

The phenetic analysis performed with 42 primer pairs reported earlier (Hamwieh *et al.*, 2010) yielded tree 'C' with two distinct groups similar to tree 'B' with little difference in clustering patterns. Cultivar Khasab formed a separate branch outside the first and second groups. The confidences within this cluster were high based on bootstrap values (Fig. 4.4 C). The phenetic results obtained with combined 72 SSR markers (Fig. 4.4 D) were similar to those with 42 SSR markers with a very high bootstrap values (Fig. 4.4 C; Hamwieh *et al.*, 2010). The branching order is the same for both trees 'C' and 'D'.

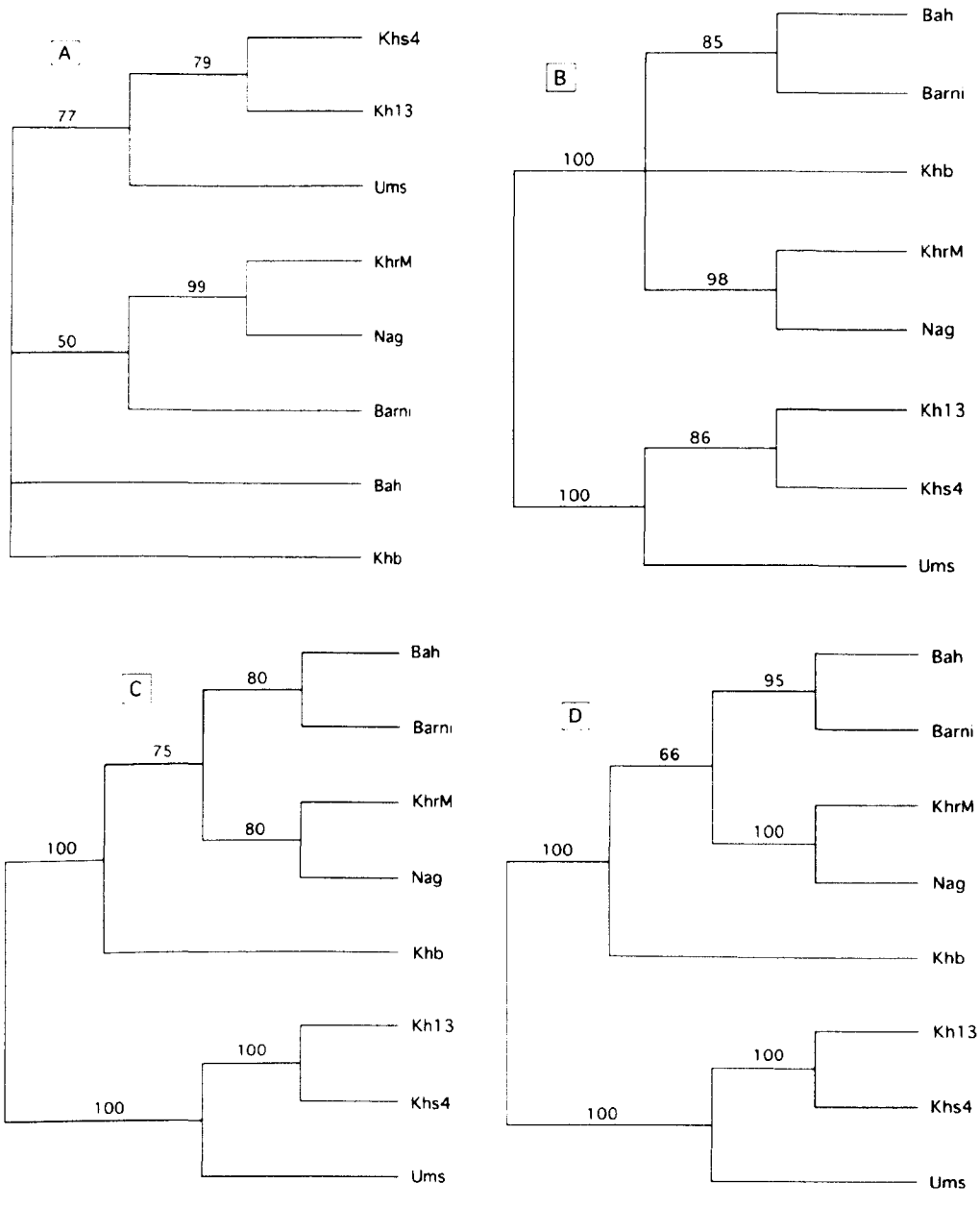


Figure 4.4: A-D Phenetic relationships between the eight parents of Omani date palm using (A) 11 new SSRs markers produced by ourselves; (B) 19 SSRs markers developed by (Billotte *et al.*, 2004 and Akkak *et al.*, 2009); (C) 42 new SSRs primer pairs designed by Hamwieh *et al.* (2010); and (D) a total of 72 combined SSRs markers. All four phenetic trees were generated by Neighbour-Joining analysis based on Nei's genetic distance. The numbers shown on each branch represent the bootstrap percentages obtained from 1000 replications. Codes correspond to the names of accessions in Table 4.

4.5 Discussion

4.5.1 Development of new microsatellite primers

Microsatellite markers reveal high levels of polymorphism among date palm genotypes (*Phoenix dactylifera* L.) giving them utility for germplasm diversity studies as well as for cultivar identification (Billotte *et al.*, 2004; Akkak *et al.*, 2009). Nevertheless, there is still a need and necessity to develop more markers to enhance the genotyping accuracy as well as for the construction of genetic maps for date palm.

In this study, we have developed 13 new SSRs markers, which were evaluated on eight Omani parents of date palm, with 11 markers producing polymorphic bands while two SSRs produced monomorphic bands. Although the number of SSR primers designed by us was low (13) the overall recovery rate from design was 30% (polymorphic). Compared to present results, Akkak *et al.* (2009) found 17 polymorphic SSRs primers (41%) out of total 41 primers tested.

The eleven SSR markers resulted in a total of 42 alleles amplified from 8 Omani date palm parents with an average of 3.82 alleles per locus. These results are similar to those reported by Ahmed and Al-Qaradawi (2009) who reported 40 alleles from 15 Qatari date palm cultivars with a mean of 4 alleles per locus, however it is lower than 77 alleles detected by Elmeer *et al.* (2011) in 11 Qatari cultivars. The observed heterozygosity with 8 Omani date palm parents was 0.38, which is significantly lower than the 0.841 and 0.820 observed by Elshibli and Korpelainen (2008) representing 45 female

(including eight females from Morocco) and 23 male palms from Sudan, respectively. The relatively low heterozygosity observed in Omani date palm could possibly be attributed to the breeding population structure, usually with relative inbreeding smaller founder populations in the population appeared with less heterozygosity. The results also revealed a lower level of polymorphism information content PIC compared to the 0.77 attained by Elmeer *et al.*, (2011). Additionally, the eight Omani cultivars showed a lower level of gene diversity (0.50) compared to 0.8 among eleven Qatari date palms (Elmeer *et al.*, 2011). A Nei's genetic distance based phenetic analysis revealed that Omani date palm cultivars, Khalas4, Khalas13 and Um-Alsela clustered together irrespective of different SSR primers used. This was also true for Naghal and Khori (Fig. 4.4).

4.5.2 Screening and testing 30 published microsatellite (SSRs) primer pairs

The results of using 30 SSRs markers developed for date palm (Billotte *et al.*, 2004; Akkak *et al.*, 2009) gave successful amplification across the eight Omani cultivars excluding locus mPdCIR044 and locus mPdCIR048 which produced multiple bands with an unclear major product size. The mPdCIR044 and mPdCIR048 markers were previously reported as yielding erratic amplification (Billotte *et al.*, 2004; Pintaud *et al.*, 2010). According to Pintaud *et al.* (2010) mPdCIR044 performed erratically problem which is probably due to "a geographical distribution of a mutation in the annealing site of one of the primers", making the amplification less stable across the available germplasm.

The SSRs markers employed in this study revealed large numbers of alleles (137 with a mean of 7.21 alleles per locus). The number of alleles detected in this study is higher than the 40 alleles scored by Ahmed and Al-Qaradawi (2009) among fifteen Qatari cultivars. However, it is lower than the previously identified 188 different alleles, varying from 3 to 21 alleles per locus from 30 Iraqi date palm samples, which was little over than three times more than our sample size (Khierallah *et al.*, 2011b). High levels of PIC (average 0.75) and heterozygosity (0.60) were detected between the eight Omani parents, which is slightly higher than those of 0.67 (PIC) and 0.50 (heterozygosity) reported for the 30 Iraqi cultivars (Khierallah *et al.*, 2011b). Dizkirici *et al.*, (2008) suggesting that the specific SSRs used, the number of markers and the choice of genotypes all influence the results obtained. The level of gene diversity (0.78) between the eight Omani cultivars is higher than the 0.695 recorded for 30 Iraqi cultivars using the same 22 SSRs markers in the present study. The results based on Nei (1973) genetic distances revealed an interesting fact that Khalas4, Khalas13 and Um-Alsela as well as Naghal and Khori are the genetically closest genotypes among the eight Omani cultivars tested.

4.5.3 Screening and testing of 100 new microsatellite (SSRs) primer pairs

In the present study, we have screened 100 new SSRs from primers previously designed on the eight Omani date palm cultivars. The microsatellites tested were highly polymorphic (42%) producing 190 alleles with a mean of 4.52 alleles per locus and a major allele frequency of 0.47. These SSRs primers were more efficient in revealing polymorphism than those reported earlier as

might be expected as they are derived from recent transcriptome and genome sequencing efforts. Elmeer *et al.* (2011) have reported that 10 primers out of 30 showed polymorphic banding while the remaining were either monomorphic or failed to amplify. They also observed 77 alleles with a mean of 7.7 alleles per locus. The higher numbers of alleles per locus are likely to be related to the number of studied genotypes and the geographical distribution of samples (Elmeer *et al.*, 2011).

Furthermore, the mean PIC of 0.59 and the gene diversity of 0.64 in this study were low as compared to Qatari cultivars which had an average PIC 0.77 and gene diversity of 0.80 (Elmeer *et al.*, 2011). Interestingly, the heterozygosity (0.42) in Omani cultivars was similar to those of Qatari cultivars (Elmeer *et al.*, 2011). It is also notable from the average genetic distance analysis that Khalas4, Khalas13 and Um-Alsela were the closest among the eight parents, showing the lowest dissimilarity.

4.5.4 Analysis of eight parents of Omani date palm using the combined 72 polymorphic SSRs primer pairs

In this section, we analyzed eight cultivars using data from the combined 72 SSRs markers derived from ourselves, Billotte *et al.* (2004); Akkak *et al.* (2009), and primer pairs provided by Hamwieh *et al.* (2010). The analysis showed 369 alleles (Appendix 4). Elshibli and Korpelainen (2008) detected 343 alleles among 68 accessions from different geographic locations using sixteen SSRs markers. The differences in numbers of alleles between different reports could be explained mainly by the number of SSR markers used and the size and genetic diversity of the populations under study.

In this study the PIC values of the 72 SSRs loci ranged from 0.19 (DPALM344) to 0.87 (PDCAT11) with an average of 0.61 (Appendix 4). Similarly, the eight Omani cultivars showed high levels of gene diversity (0.66). However, this level was lower than that observed within the eleven Qatari cultivars, which might be due to the Qatari cultivars being more divergent than the Omani cultivars.

Overall, these studies have shown that the highest gene diversity (0.75) among the eight cultivars was observed using the 19 SSRs markers from Billotte *et al.* (2004) and Akkak *et al.* (2009). However, the highest number of alleles was attained using all 72 SSRs markers, as would be expected (Table 4.13). No significant differences were noticed in the phenetic analysis of the data generated by 19 SSR, 42 SSR and 72 SSR markers (Fig. 4.4 B-D). However, the 11 new genomic SSR markers produced different patterns by clustering cultivar Barni with Khori male and Nagla and creating separate branches for Bahlani and Khasab. Interestingly cultivar Khasab formed a separate branch in all four trees. It is not clear why Khasab was separated from the other genotypes, even though it originated from the same geographic region.

Table 4-13: A summary of genetic diversity information for eight Omani cultivars using different sets of SSRs markers; 11 new SSRs markers produced by ourselves, 19 SSRs markers developed by Billotte *et al.* (2004) and Akkak *et al.* (2009), 42 new SSRs primer pairs designed by Hamwieh *et al.* (2010), and the combined 72 SSRs markers.

	Total no. of alleles	Major allele frequency	Average of allele per locus	Gene diversity	Heterozygosity	PIC
11 new SSRs	42	0.64	3.8	0.5	0.38	0.45
19 SSRs	137	0.34	7.2	0.78	0.60	0.75

42 new SSRs	190	0.47	4.52	0.64	0.42	0.59
72 SSRs	369	0.46	5.13	0.66	0.46	0.61

4.6 Conclusion

Simple sequence repeat DNA markers (SSR or microsatellite markers) are a powerful tool to provide information on the relatedness of various genotypes that could be difficult to distinguish morphologically, however the available SSRs markers for date palm have been very limited. Six main conclusions can be drawn from this study:

- This study adds a new set of SSRs markers which would be of major value in date palm improvement programs and germplasm characterization; 11 derived from a genomic library and a further 42 derived from untested primer sequences.
- This study suggests that the 19 SSR markers from Billotte *et al.* (2004) and Akkak *et al.* (2009) are highly informative for the analysis of date palm genetic diversity and are a useful resource for genetic mapping.
- Using different sets of SSR markers revealed similar grouping for (Khalas4, Khalas13, and Um-Alsela) and (Naghal and Khori) which strongly supports their closer relationship.
- The phenetic analysis obtained with 72 SSR markers were similar to those using 42 SSR markers (Hamwieh *et al.*, 2010) indicating that SSR markers are reflecting similar genetic relationships in the eight Omani parents.
- No significant differences were observed in relation to sex between the eight parents, which is unsurprising, especially when we consider

Oman as one unit regarding cultural practices and exchange of plant materials. Farmers in Oman depend on a few selected males for the pollination of females trees. Male tree are selected for the fruit quality they confer to female palm on pollination and they are exchanged between farmers on this basis.

- Determination of genetic variability among the eight parents is useful for the selection of the correct parents for the two populations, which will be used for genetic mapping.

Chapter 5. GENETIC STRUCTURE OF OMANI DATE PALM (*PHOENIX DACTYLIFERA* L.) GERMPLASM AND ITS RELATIONSHIPS WITH 'EXOTIC' GERMPLASM USING SSR MARKERS

5.1 Introduction

Different studies have been conducted on date palm cultivars either to identify the relationship between cultivars or to solve related problems, such as identifying potential sources of pest and salinity resistance. Most studies identified and compared cultivars by using a large set of morphological characters, including; tree height, number of pinnae and spines, fruit weight, flesh weight, fruit and seed size, color, and tested sugars. These morphological characters can be highly influenced by environmental conditions, such as soil and weather, and thus may not reflect the true genetic relationships (Elhoumaizi *et al.*, 2002; Zehdi *et al.*, 2004; Elshibli and Korpelainen, 2009; Al-Ruqaishi *et al.*, 2008). Furthermore, the identification of trees by their fruits is usually not possible until the onset of fruiting, which takes 3 to 5 years or sometimes even longer. Often it has been observed that one date palm cultivar is known by more than a single name in different regions. Alternatively, two different cultivars may be given the same name due to indistinguishable morphological characters. Both these issues hinder the identification of the cultivars (Al-Khalifah and Askari 2006; Cao and Chao, 2002; Elhoumaizi *et al.*, 2002).

Along with morphological characters, several other methods have been used to identify date palm cultivars, particularly biochemical markers- isozymes - and proteins. These have proven to be effectual in varietal identification of date

palm, however, they are an indirect approach for detecting genomic variation, provide limited information and can be tissue- or developmental stage-specific (Al-Khalifah and Askari, 2006; Bader *et al.*, 2007).

Recently, DNA-based markers have been used for investigating genetic diversity with a progression of markers from restriction fragment length polymorphisms (RFLPs) to a set of PCR-based technologies such as RAPD, AFLP, SSR and SNPs (Johnson *et al.*, 2009) being applied.

Simple Sequence Repeats or SSR markers (also known as microsatellites) are one of the most relevant molecular markers used for plant diversity analysis in general and date palm germplasm, in particular. SSRs are based on the incorporation of naturally occurring short repetitive oligonucleotide stretches into PCR amplified fragments. SSRs are co-dominant, so allow the unambiguous detection of heterozygotes. Palliyarakkal *et al.* (2011) have reported that the application of SSR markers in date palm has become extremely valuable and are increasingly popular, particularly due to their reproducibility and transferability between palm species. Akkak *et al.* (2009) reported the use of SSRs to evaluate genetic diversity data in date palm germplasm in the last few years with a number of different microsatellite markers developed from *P. dactylifera* L., allowing genetic analysis of Qatari, Tunisian, Omani and Sudanese date palm (Ahmed & Al-Qaradawi, 2009; Elmeer *et al.*, 2011; Zehdi *et al.*, 2004; Al-Ruqaishi *et al.*, 2008; Elshibli and Korpelainen, 2010).

The aim of the present work is to investigate the genetic diversity of Omani date palm germplasm using SSR markers in order to obtain an accurate

description and understanding of these genetic resources and to compare them with 'exotic' germplasm.

5.2 Data generation and analysis

DNA from one hundred and ninety-four date palm accessions from Oman (151 female cultivars and 43 male trees) and forty-eight accessions from Italy (Sanremo, Bordighera), USDA-ARS, France, Iraq, Libya, Sudan and Iran were amplified with 12 SSR primer pairs (Table 5.1). These microsatellites were selected on the basis of their known performance and their possession of high levels of polymorphism during screening. All date palm accessions used in this study are summarized in Tables 5.2, 5.3 and 5.4.

The PCR products were analysed on a Beckman Coulter CEQ 8000 Genetic Analysis System (Beckman-Coulter, Fullerton, CA) as described in Chapter 3; Section 3.10 and the fragment's size (bp) and profiles were recorded.

All molecular data were analysed using the GenAlex program (Version 6.4) (Peakall & Smouse, 2006). The total number of alleles, percentage of polymorphic loci and allelic frequencies were determined. Observed and expected heterozygosity (*He, Ho*) and fixation index for the F-statistics of Wright (Fis, Fst, and Fit) was also calculated using GenAlex (Wright, 1965). *F* = Fixation Index = (*He* - *Ho*) / *He* = 1 - (*Ho* / *He*) and it varies theoretically from −1 to 1. *Fit* and *Fis* were defined as genetic deviation from Hardy–Weinberg expectations within and between populations, respectively. If *Fit* and *Fis* are 0, populations are at Hardy–Weinberg equilibrium. *Fst* is an evaluation of gene differentiation between populations and varies from 0 to 1.

If there is no genetic differentiation among population the value of *Fst* is 0 (Qi-Lun *et al.*, 2008).

The genetic distances (PhiPT) between groups were tested by Analysis of Molecular Variance (AMOVA) based on 999 permutations. The AMOVA provides an estimate for the partitioning of the genetic variation that exists among populations and it takes into account correlation among loci which are not measured while computing the locus *Fst* statistics. The AMOVA does not take into account the expected heterozygosity, but takes individuals as haploid for each locus, and therefore the AMOVA Φ estimate would be expected to be higher than the Wright's *Fst* estimate (Sousa-Correia, 2008).

In addition, Principal Coordinate Analysis (PCA) was used to visualize differences between the selected populations by computing a matrix based on the Nei genetic distance. A phenetic analysis was performed with PowerMarker Ver. 3.25 software (Liu and Muse, 2005). Uninformative characters were excluded from the analysis. A phenetic tree was constructed by using the Neighbor-Joining clustering algorithm. Bootstrap analysis (1000 iterations) was performed to estimate stability and support for the inferred clades (Felsenstein, 1985) and was visualized by using the MEGA software Ver. 4 (Kumar *et al.*, 2004). Unrooted UPGMA (Unweighted Pair-Group Method with Arithmetical Averages) trees displaying the distribution of accessions from different germplasm were also generated using the DARwin 5.0 software based on the UnWeighted Neighbor-Joining method (Perrier and Jacquemoud-Collect, 2006).

Table 5-1: List of twelve SSRs markers used to study the genetic diversity of 194 Omani date palm accessions and 48 accessions from Italy (Sanremo, Bordighera), USDA-ARS, France, Iraq, Libya, Sudan and Iran. Marker name, motif repeat, annealing temperature, expected product size and original source are given.

Marker no.	Marker Name	Motif repeat	Annealing T_m (°)	Expected size (bp)	Source
1	mPdCIR010	$(GA)_{22}$	52°C	118-161	Billotte *et al.* (2004)
2	mPdCIR015	$(GA)_{21}$	52°C	120-156	Billotte *et al.* (2004)
3	mPdCIR016	$(GA)_{14}$	52°C	130-138	Billotte *et al.* (2004)
4	mPdCIR025	$(GA)_{22}$	52°C	199-231	Billotte *et al.* (2004)
5	mPdCIR050	$(GA)_{15}$	52°C	154-208	Billotte *et al.* (2004)
6	mPdCIR057	$(GA)_{20}$	52°C	251-278	Billotte *et al.* (2004)
7	mPdCIR093	$(GA)_{17}$	52°C	153-184	Billotte *et al.* (2004)
8	PDCAT2	$TCGCTG(TC)_3(TC)_3T(TC)_3T$ $(TC)_3T(TC)_4TTCTGTCCCG(TC)_{16}T(TC)$	55°C	166-194	Akkak *et al.* (2009)
9	PDCAT11	$(TC)_7(TC)_{20}$	55°C	133-155	Akkak *et al.* (2009)
10	PDCAT12	$(CT)_{19}$	55°C	145-167	Akkak *et al.* (2009)
11	PDCAT14	$(TC)_{19}(TC)_{16}$	55°C	114-145	Akkak *et al.* (2009)
12	PDCAT20	$(GA)_{29}$	55°C	290-353	Akkak *et al.* (2009)

Table 5-2: List of 151 female cultivars collected from the National Germplasm Collection at Wadi Qurayat Research Station, Bahla, Sultanate of Oman, and their laboratory code.

Lab Code	Accession Name	Lab Code	Accession Name
1a	Qash Bunaringa	77a	Abu Al Audooq
2a	Miznag Ahmar	78a	Qash Shafer
3a	Qash Tabaq	79a	Alak
4a	Fard	80a	Snah
5a	Khsab	81a	Qash Hatami
6a	Qash Hagr	82a	Qash Katerh
7a	Qash Beladsait	83a	Qash Al Rabab
8a	Khinaizi Halaw	84a	Qash Sba
9a	Hilali Hassa	85a	Qash Hiyshmi
10a	Zabad	86a	Qash Al Teebi
11a	Qash Na'im	87a	Qash Rsheed
12a	Qash Halaw	88a	Nashu Al khzma
13a	Qash Bu Rashid	89a	Berzeban
14a	Bentaami	90a	Nashu Al Wakhrh
15a	Farid	91a	Naghlt Khalas
16a	Battash	92a	Khalas Oman
17a	Khinaizi Arabi	93a	Qash Gha'roof
18a	Qash Qaroot	94a	Qash Ghafan
19a	Kalbi	95a	Qash Nas'rah
20a	Umm Alssila	96a	Qash Gheniyah
21a	Mazrm	97a	Qash Nwaihi
22a	Minaz	98a	Qash Al Masbt
23a	Mahlabbi	99a	Qasht Naghal
24a	Khamri	100a	Qash Ali
25a	Shabroot	101a	Qash Hareb
26a	Mekhildi	102a	Qash Nasir
27a	Nashu Ba'oodh	103a	Qash Safiyh
28a	Mawaz	104a	Qash Fakhrh
29a	Nashu Ghoson	105a	Qash Suwaid
30a	Qash Suwaih	106a	Qash Ba'Omar
31a	Qash Qataari	107a	Nashu Shamiss
32a	Qash Ain Al Bakar	108a	Shahl
33a	Qash Al Teeb	109a	Ramli
34a	Qash Hareer	110a	Shiham
35a	Hilali Ahmer	111a	Seedi
36a	Qash Zamel	112a	Khashkar
37a	Ghrabo	113a	Nashu Saleh
38a	Shaeeri	114a	Nashu Maneh
39a	Qash Abu Keebal	115a	Lulu
40a	Bayadh	116a	Rabai
41a	Qash Mushrab	117a	Qash Suwailim
42a	Qash Ghssan	118a	Bata
43a	Qash Hmdan	119a	Barny
44a	Qash Habeeb	120a	Nashu Fahood
45a	Qash Abu Saif	121a	Muttrahi
46a	Qash Habisha	122a	Bidaa
47a	Hessas	123a	Medairki
48a	Qash Manzef	124a	Kibkab
49a	Zaad	125a	Huzaifah
50a	Naghal	126a	Nashu Al Khashiyah
51a	Qash Qantara	127a	Hawam

52a	Ma`an	128a	Qadmi
53a	Bunaringa	129a	Qash Gammah
54a	Jebri	130a	Qash Saima
55a	Hilali Makran	131a	Medlooki
56a	Qash Humaid bin Ghareb	132a	Damoos
57a	Menhi	133a	Qash Hareez
58a	Mebseli	134a	Qash Al Looz
59a	Bershi	135a	Qash Al Semnah
60a	Rees	136a	Qash Hamreiyah
61a	Qash Al Hareem	137a	Qash Baloobiya
62a	Nashu Ewan	138a	Qash Abu Al Sohoon
63a	Malkt Deeni	139a	Qash Mishah
64a	Mayasi	140a	Qash Al Dahiyah
65a	Hadaqi	141a	Qash Al Ramliyah
66a	Abu Qareen	142a	Qash Al Wali
67a	Selahni	143a	Medgahdel
68a	Qash Sahrh	144a	Bel'aq
69a	Naboot Saif	145a	Jebreen
70a	Qash Al Yamam	146a	Qash Bussemen
71a	Qash Humaid	147a	Hilali Omani
72a	Naghl Lulu	148a	Qash Ghinuwi
73a	Qash Al Rabeea	149a	Qash A'Saba Al Aruz
74a	Mazni	150a	Serna
75a	Qash Al wa'b	151a	Khalas Al Zahra
76a	Qash Al Saghiay		

Table 5-3: List of 43 male trees collected from the National Germplasm Collection at Wadi Qurayat Research Station, Bahla, Sultanate of Oman, and their laboratory code.

Lab Code	Accession Name	Lab Code	Accession Name
152	Khori 1	174	Bu'Sab'ah 3
153	Khori 2	175	Rghad 1
154	Khori 3	176	Rghad 2
155	Khori 4	177	Rghad 3
156	Naghayli 1	178	A'reesh 1
157	Naghayli 2	179	A'reesh 2
158	Naghayli 3	180	An'bati 1
159	Medgahdel	181	An'bati 2
160	Bahlani 1	182	An'bati 3
161	Bahlani 2	183	Al Maquidha 1
162	Bahlani 3	184	Al Maquidha 2
163	Bahlani 4	185	Soo'qum 1
164	Ghareef 1	186	Soo'qum 2
165	Ghareef 2	187	Khzini 1
166	Ghareef 4	188	Khzini 2
167	Al Fahel Al dhakm 1	189	Khzini 3
168	Al Fahel Al dhakm 2	190	Do'wairah 1
169	Unknown Male 1	191	Do'wairah 2
170	Unknown Male 2	192	Al Lasah 1
171	Unknown Male 3	193	Al Lasah 2
172	Bu'Sab'ah 1	194	Al Lasah 3
173	Bu'Sab'ah 2		

Table 5-4: List of 48 female accessions included in this study from Italy (Sanremo, Bordighera), USDA-ARS, France, Iraq, Libya, Sudan and Iran, their laboratory code and country of origin. (-) indicates accession name was not available.

Lab Code	Accession Name	Country of Origin	Lab Code	Accession Name	Country of Origin
414	-	Italy/Sanremo	Mkm-Iq	Maktoom	Iraq
433	-	Italy/Sanremo	Bdm-Iq	Bdmalki	Iraq
434	-	Italy/Sanremo	Ben-Iq	Benosh	Iraq
439	-	Italy/Sanremo	Ash-Iq	Ashrasi	Iraq
441	-	Italy/Sanremo	Khs-Iq	Khastawi	Iraq
443	-	Italy/Sanremo	Say-Iq	Saylani	Iraq
444	-	Italy/Sanremo	Bhm-Iq	Bahram	Iraq
447	-	Italy/Sanremo	Aw-Ly	Awreeq	Libya
500	-	Italy/Bordighera	Kh-Ly	Khmag	Libya
501	-	Italy/Bordighera	Ta-Ly	Taghiyat	Libya
523	-	Italy/Bordighera	Am-Ly	Amreer	Libya
529	-	Italy/Bordighera	Tal-Ly	Talees	Libya
541	-	Italy/Bordighera	Sa-Ly	Saidi	Libya
Khalas	Khalas	Arabia; USDA	Aq-Ly	Aqudool	Libya
Thory	Thory	Algeria; USDA	Med-Sdn	Medina	Sudan
Hilali	Hilali	Oman; USDA	Gnd-Sdn	Gondaila	Sudan
Barhee	Barhee	Iraq; USDA	Bar-Sdn	Barakawi	Sudan
Medjool	Medjool	Morocco; USDA	Bit-Sdn	Bitamoda	Sudan
Fran1	-	France	Do-Sdn	Dogna	Sudan
Fran5	-	France	Iran3	Bentossbae	Iran
DA-Iq	Daml Asfer	Iraq	Iran9	Gentaar	Iran
B-Iq	Badmi	Iraq	Iran13	Zahedi	Iran
Sar-Iq	Sarmadti	Iraq	Iran22	Soweidance	Iran
Khd-Iq	Khadrawy	Iraq	Iran40	Halilehei	Iran

5.3 Results

5.3.1 Diversity analysis of Omani date palm cultivars

5.3.1.1 Allele number, percentage of polymorphic loci and allele frequency of microsatellites used

The analysis of one hundred and ninety-four individual accessions of date palm from Oman (151 cultivars and 43 male trees) using 12 SSR markers resulted in a total of 188 alleles with an average value of 15.7 alleles per locus.

The number of alleles per locus varied between 4 in locus mPdCIR057 and 21 in locus PDCAT2 (Table 5.5). All loci detected polymorphism in both female and male accessions. Loci mPdCIR010, mPdCIR016, mPdCIR093, PDCAT2, PDCAT11, and PDCAT14 identified the highest allele numbers in male accessions, while loci mPdCIR015, mPdCIR025, mPdCIR050, mPdCIR057, PDCAT12, and PDCAT20 revealed the highest allele numbers in female accessions (Figure 5.1).

Table 5-5: Allele number per locus for 195 accessions of Omani date palm

Locus name	Allele number	Locus name	Allele number
mPdCIR010	20	mPdCIR093	11
mPdCIR015	17	PDCAT2	21
mPdCIR016	14	PDCAT11	17
mPdCIR025	15	PDCAT12	16
mPdCIR050	19	PDCAT14	20
mPdCIR057	4	PDCAT20	14
Total number = 188			

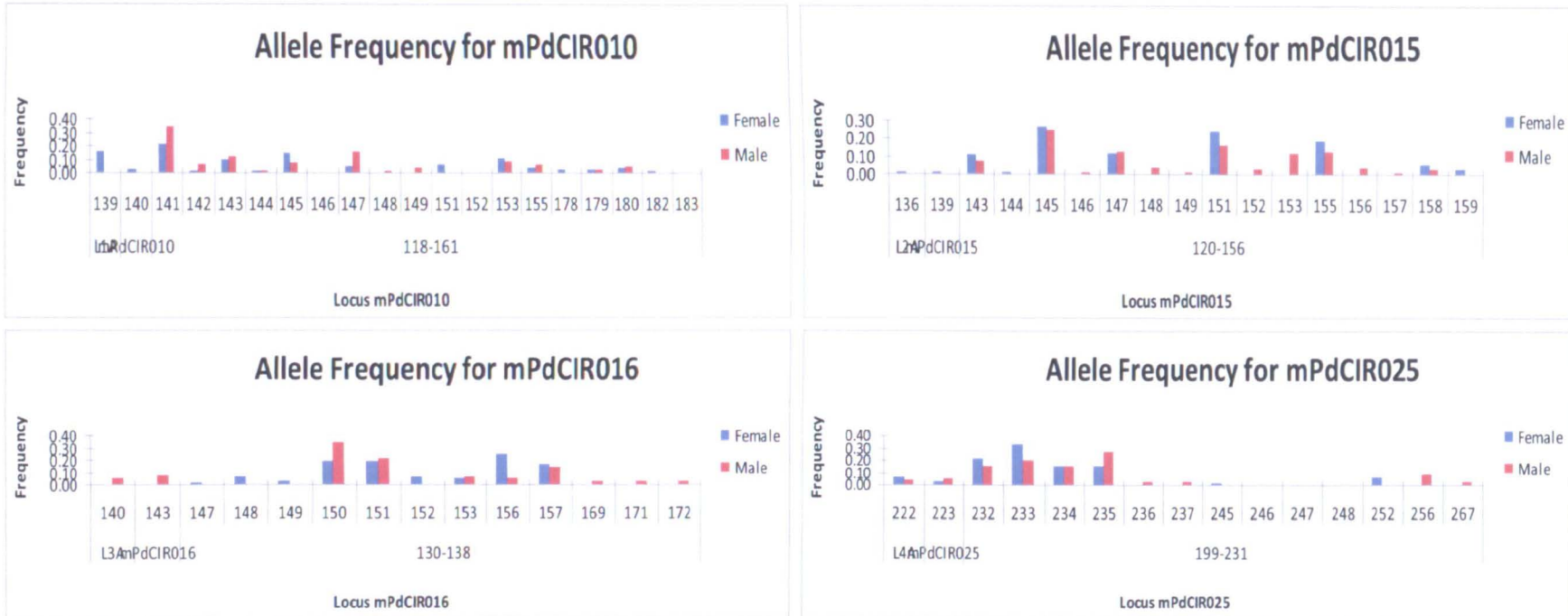


Figure 5.1: Histograms illustrating the microsatellite allelic frequency distributions in 152 female and 43 male accessions from Oman amplified by 12 pre-selected SSR markers.

Figure 5.1 (Continued)

Figure 5.1 (Continued)

5.3.1.2 Heterozygosity and fixation index

Heterozygosity detected by the 12 SSR primer pairs for Omani female and male accessions was high with the exception of one locus (mPdCIR057) and ranged from 0.241 to 0.870 (Table 5.6). The mean of expected heterozygosity (*mHe*) varied from 0.240 in locus mPdCIR057 to 0.854 in mPdCIR010, while the mean of observed heterozygosity (*mHo*) ranged between 0.260 in mPdCIR057 and 0.809 in mPdCIR015. Overall, the mean of observed heterozygosity was lower than the mean of expected heterozygosity between and within male and female accessions, except for locus mPdCIR057 which has a higher mean of observed heterozygosity (Table 5.6), potentially indicative of population structure within Omani palms.

The *Fis*, *Fit*, and *Fst* were also estimated to analyse the genetic structure of Omani date palm accessions (Table 5.6). According to Wright (1965) a system for describing the properties of subdivided natural populations was developed and three parameters were proposed in terms of individuals (I), subdivisions (S), and total population (T). The average of *Fis* (fixation index of individuals compare to subpopulations) was 0.173, varying from -0.082 (mPdCIR057) to 0.324 (PDCAT14), and the average *Fit* (fixation index of individuals relative to the total population) was 0.190 ranging from -0.078 to 0.513 at corresponding loci. The average *Fst* (fixation index of subpopulation compared to the total population) was 0.021.

Table 5-6: Data on heterozygosity and fixation index calculated with GenAlex 6.4 for 152 female and 43 male Omani date palm accessions based on 12 SSR markers.

Locus name	*Ht*	*mHe*	*mHo*	*Fis*	*Fit*	*Fst*
mPdCIR010	0.870	0.854	0.792	0.073	0.090	0.019
mPdCIR015	0.844	0.837	0.809	0.033	0.041	0.008
mPdCIR016	0.837	0.816	0.587	0.280	0.298	0.025
mPdCIR025	0.825	0.814	0.642	0.211	0.222	0.014
mPdCIR050	0.784	0.757	0.612	0.191	0.219	0.035
mPdCIR057	0.241	0.240	0.260	-0.082	-0.078	0.004
mPdCIR093	0.718	0.711	0.658	0.075	0.083	0.009
PDCAT2	0.821	0.811	0.806	0.007	0.018	0.012
PDCAT11	0.864	0.826	0.611	0.261	0.293	0.044
PDCAT12	0.704	0.685	0.343	0.500	0.513	0.026
PDCAT14	0.840	0.810	0.547	0.324	0.348	0.035
PDCAT20	0.789	0.770	0.609	0.209	0.229	0.025
Mean				0.173	0.190	0.021

Key:

Ht = Total expected heterozygosity = 1 - Sum (tpi^2) where tpi is the frequency of the ith allele for the total & Sum tpi^2 is the sum of the squared total allele frequencies

m = mean

m*He* = mean of expected heterozygosity, m*Ho* = mean of observed heterozygosity

Fis = (Mean *He* - Mean *Ho*) / Mean *He*, *Fit* = (*Ht* - Mean *Ho*) / *Ht*, *Fst* = (*Ht* - Mean *He*) / *Ht*

5.3.1.3 Genetic similarity

A high level of genetic similarity (0.866) was observed between 151 female and 43 male Omani accessions.

5.3.1.4 Analysis of molecular variance of Omani cultivars

Pairwise population comparisons were conducted using an Analysis of Molecular Variance (AMOVA). The AMOVA results indicated that most (94%) of the molecular variation in Omani date palm exists between individuals, within populations, with lower amounts between female and male populations (6%) as presented in Figure 5.2. Permutation tests (based on 999 permutations) suggest that the overall ΦPT (ΦPT = 0.057, P = 0.001) is higher than the *Fst* = 0.021, but still denotes a small, but significant, level of differentiation (Table 5.7).

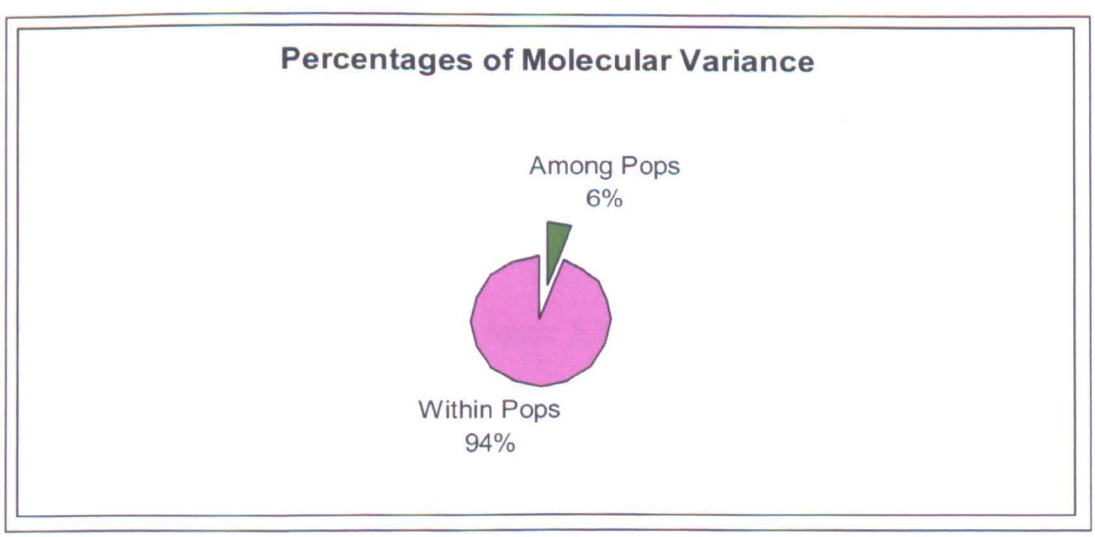


Figure 5.2: Distribution of the molecular variance within and between female and male date palm accessions in Oman obtained using 12 SSR primer pairs.

Table 5-7: Analysis of molecular variance for Omani date palm accessions (151 female and 43 male) obtained using 12 SSR primer pairs.

Source	d.f.	Sum of Squares	MS	Est. Var.	Percentage of variation	*P* value
Among Pops	1	56.416	56.416	0.676	6%	0.001
Within Pops	193	2147.625	11.128	11.128	94%	
Total	194	2204.041		11.803	100%	

Stat	Value	P(rand >=data)
ΦPT	0.057	0.001

5.3.1.5 Associations between female and male accessions of Omani date palm

Associations among 194 accessions of Omani date palm (151 female and 43 male) were investigated using Principal Coordinates Analysis (PCA). The location of accessions was defined by the first principal component (PC1) and second principal component (PC2) which were displayed graphically. PC1 and PC2 explained 22.4% and 21.9%, respectively, of the total molecular variation present (Figure 5.3). The male palms are more constrained on axis 2 than the female palms. No clear differentiation between female and male accessions from Oman could be observed, despite nearly 45% of the molecular variation being represented in the PCA presented in figure 5.3. This is consistent with the results of the AMOVA which suggested only 6% of the molecular variation could be explained by differences between male and female palms.

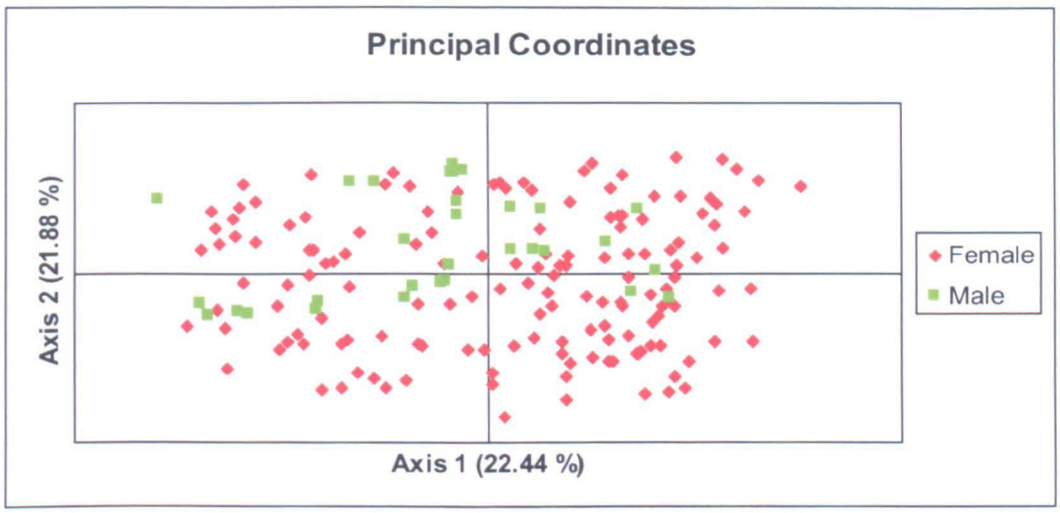


Figure 5.3: Principal Coordinates Analysis (PCA) of female and male date palm accessions from Oman. Axis 1 accounts for 22.44% and Axis 2 accounts for 21.88% of the total variation.

5.3.1.6 Cluster analysis

Bootstrap consensus phenetic trees was generated using 1000 replications based on the genetic distance index Nei (1973) to obtain a clearer picture of the genetic relationships among 194 Omani date palm accessions (151 female and 43 male). The consensus tree classified the 194 studied accessions into three major clusters (Figure 5.4). Bootstrap values were calculated based on the >50% majority rule and confidence limits were placed at the major nodes.

The tree constructed exhibited close relationships among the Omani accessions (male and female) and the confidence was confirmed by a high bootstrap values. Replicates for some accessions were used and showed strong clustering relationship with the corresponding accessions, but not identity. In general, clustering of accessions examined revealed that they were clustered independently of the accession sex.

THIS PAGE HAS BEEN LEFT BLANK

THIS PAGE HAS BEEN LEFT BLANK

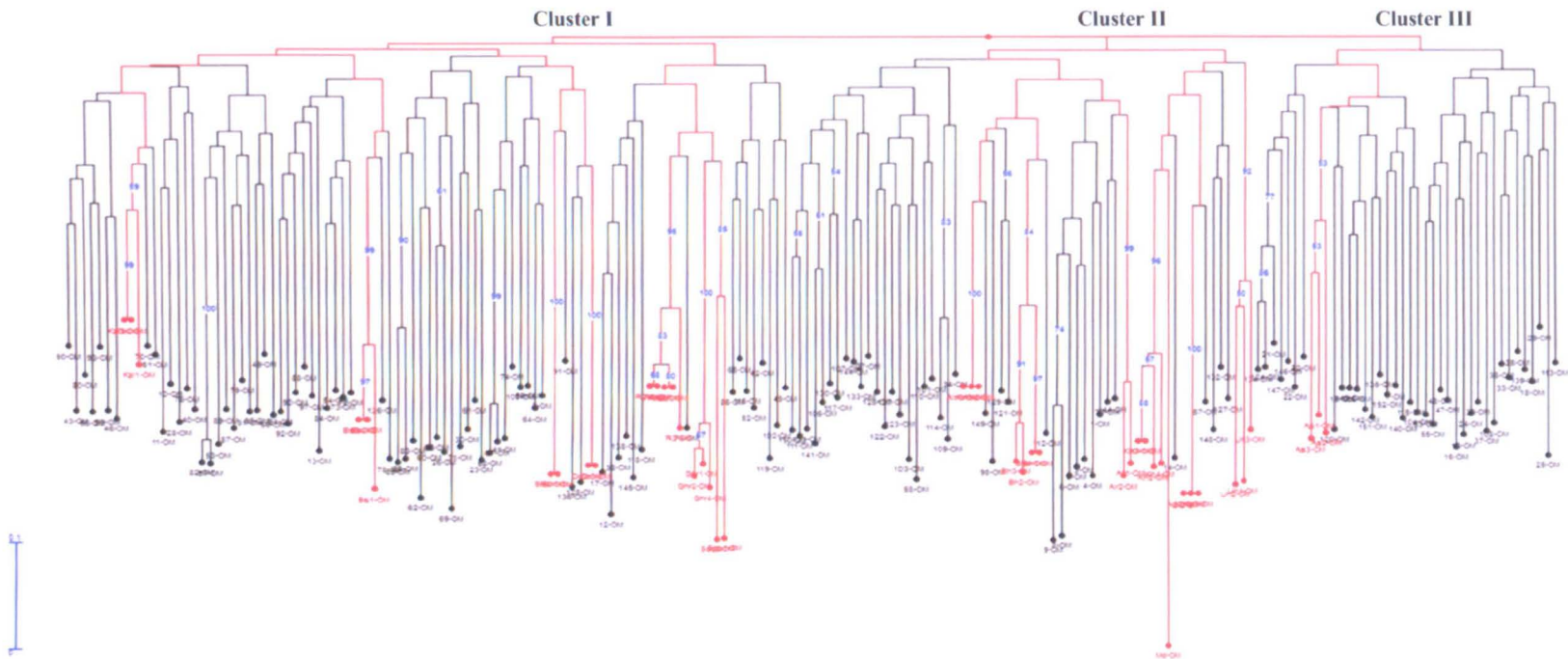


Figure 5.4: Phenetic analysis generated by Neighbour-Joining analysis based on Nei's genetic distance between 194 Omani date palm accessions representing 151 female and 43 male. Bootstrap values have been computed over 1000 replications and were placed at the major nodes. Codes correspond to the name of accessions in Table 5.2 and 5.3. Black colour is used for female palms, while red is used for male palms.

5.3.2 Diversity of Omani germplasm accessions and the germplasm from other countries

5.3.2.1 Allele number, percentage of polymorphic loci and allele frequency of the combined date palm germplasm

Table 5.8 summarizes the allele number for the twelve SSR loci used to analyse 194 accessions from Oman (Section 5.3.1) and 48 accessions from other countries including, Italy (Sanremo, Bordighera), USDA-ARS, France, Iraq, Libya, Sudan and Iran. A total of 246 alleles were observed with an average of 20.5 alleles per locus. The number of alleles per locus varied between 9 in locus mPdCIR057 and 27 in locus mPdCIR010.

Table 5-8: Allele number per locus for 242 accessions from Oman, Italy (Sanremo, Bordighera), USDA-ARS, France, Iraq, Libya, Sudan and Iran.

Locus name	Allele number	Locus name	Allele number
mPdCIR010	27	mPdCIR093	14
mPdCIR015	21	PDCAT2	24
mPdCIR016	15	PDCAT11	22
mPdCIR025	22	PDCAT12	20
mPdCIR050	25	PDCAT14	23
mPdCIR057	9	PDCAT20	24
Total number = 246			

The percentage of polymorphic loci for accessions from the nine different populations ranged from 91.67% to 100.00% with a mean of 99.17% (Table

5.9). Loci mPdCIR010, mPdCIR015, mPdCIR025 and PDCAT14 showed the highest allele numbers for accessions from Bordighera, whereas loci mPdCIR093, PDCAT2, PDCAT11, PDCAT12 and PDCAT20 contained the highest allele numbers for accessions from Iran. In addition, loci mPdCIR016, mPdCIR050 and mPdCIR057 had the highest allele numbers for accessions from Libya, Sudan and Oman, respectively (Figure 5.5).

Table 5-9: Percentage of polymorphic loci in 242 date palm accessions for nine populations: Oman (194 accessions), Sanremo (eight accessions), Bordighera (five accessions), USDA-ARS (five accessions), France (two accessions), Iraq (eleven accessions), Libya (seven accessions), Sudan (five accessions) and Iran (five accessions) using 12 SSR markers.

Population	% Polymorphic Loci	Population	% Polymorphic Loci
Oman	100.00%	Iraq	100.00%
Italy/Sanremo	100.00%	Libya	100.00%
Italy/Bordighera	100.00%	Sudan	100.00%
USDA-ARS	100.00%	Iran	91.67%
France	100.00%		
Mean = 99.17%			

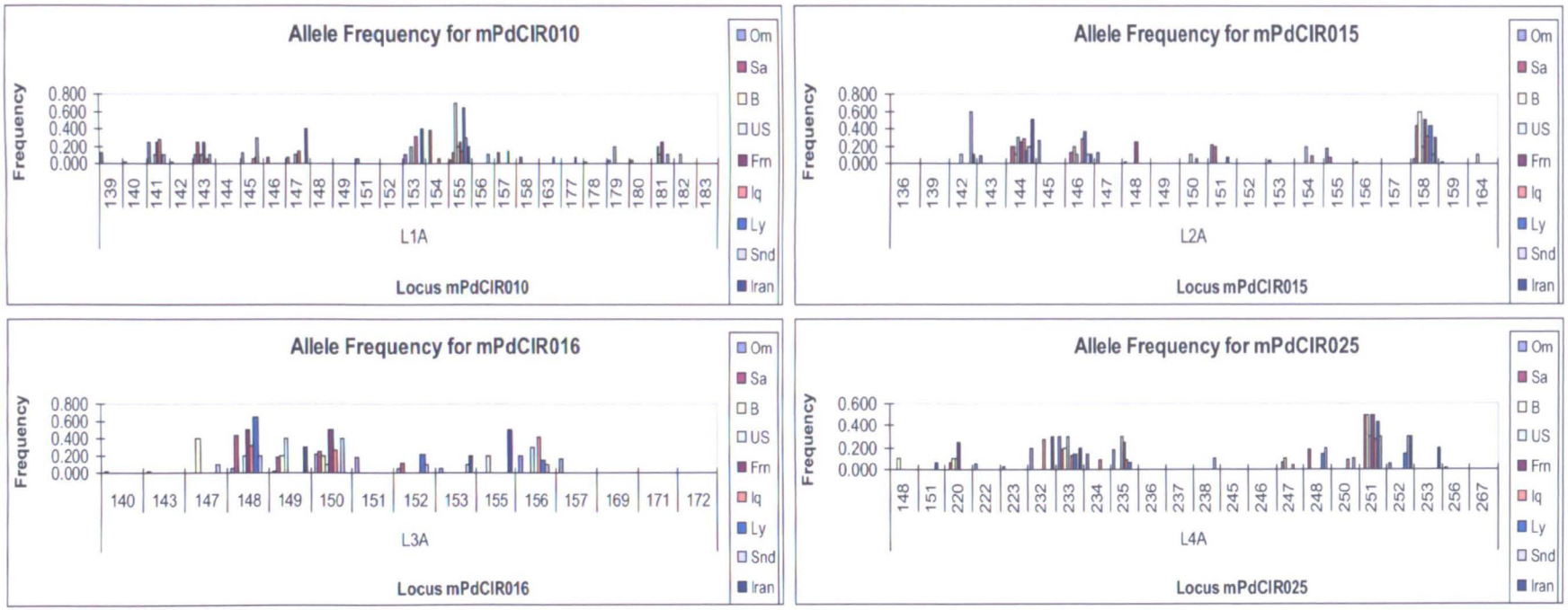


Figure 5.5: Histograms illustrating the microsatellite allele frequency distributions in 242 date palm accessions from nine different populations: Oman, Italy (Sanremo, Bordighera), USDA-ARS, France, Iraq, Libya, Sudan and Iran using 12 SSR markers.

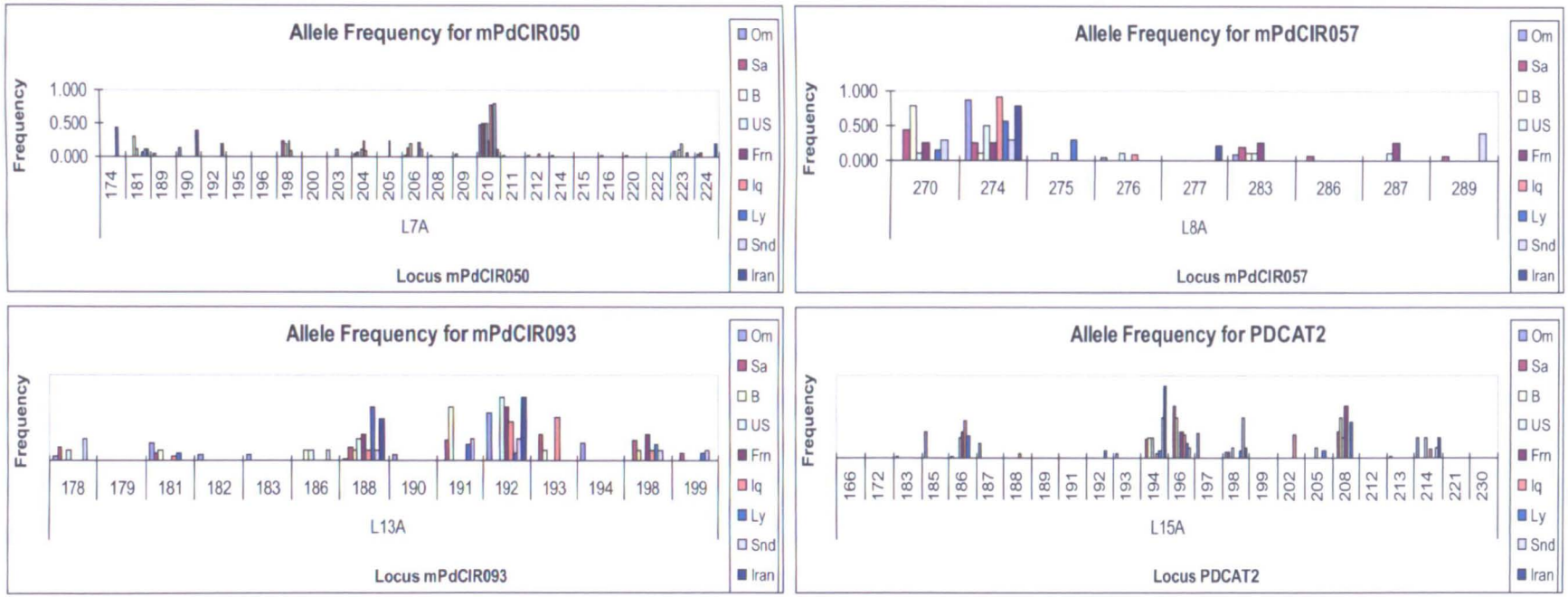


Figure 5.5 (Continued)

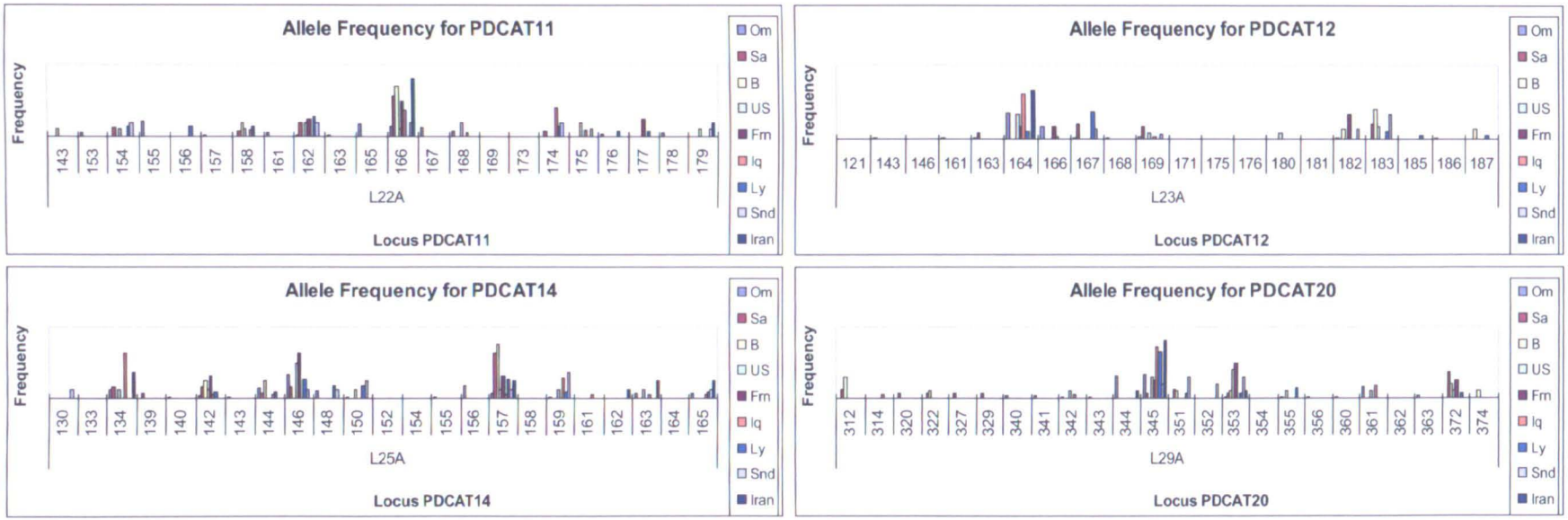


Figure 5.5 (Continued)

5.3.2.2 Heterozygosity and fixation index

The total rate of heterozygosity for 242 date palm accessions from Oman, Sanremo, Bordighera, USDA-ARS, France, Iraq, Libya, Sudan and Iran was high ranging from 0.679 (mPdCIR057) to 0.873 (PDCAT14). The mean of expected heterozygosity (m*He*) ranged from 0.493 to 0.740 in mPdCIR057 and PDCAT14, respectively. The mean of observed heterozygosity (m*Ho*) varied from 0.283 (PDCAT12) to 0.820 (PDCAT2). The m*Ho* for mPdCIR010, mPdCIR015, mPdCIR050, mPdCIR057, mPdCIR093, PDCAT2, PDCAT11 and PDCAT14 was higher than m*He* (Table 5.10).

The *Fis* values ranged from -0.032 (PDCAT11) to 0.454 (PDCAT12) with a mean of -0.005, whereas the *Fit* was 0.188 on average, and varied from 0.046 (PDCAT2) to 0.641(PDCAT12). Between accession genetic variation accounted for 19.7% of total, however the within accession genetic variation was 80.3% (Table 5.10).

Table 5-10: Heterozygosity and fixation index calculated with GenAlex 6.4 for 242 date palm accessions from Oman, Italy (Sanremo, Bordighera), USDA-ARS, France, Iraq, Libya, Sudan and Iran using 12 SSR markers.

Locus name	*Ht*	m*He*	m*Ho*	*Fis*	*Fit*	*Fst*
mPdCIR010	0.864	0.716	0.787	-0.099	0.090	0.172
mPdCIR015	0.813	0.686	0.755	-0.100	0.070	0.155
mPdCIR016	0.839	0.668	0.536	0.198	0.361	0.204
mPdCIR025	0.842	0.729	0.689	0.055	0.182	0.135
mPdCIR050	0.787	0.643	0.703	-0.093	0.106	0.183
mPdCIR057	0.679	0.493	0.553	-0.122	0.186	0.275
mPdCIR093	0.828	0.685	0.732	-0.069	0.116	0.173
PDCAT2	0.860	0.689	0.820	-0.189	0.046	0.198
PDCAT11	0.818	0.674	0.696	-0.032	0.150	0.176
PDCAT12	0.789	0.519	0.283	0.454	0.641	0.342
PDCAT14	0.873	0.740	0.818	-0.105	0.063	0.153
PDCAT20	0.809	0.644	0.616	0.044	0.239	0.204
Mean				-0.005	0.188	0.197

Key:

Ht = Total expected heterozygosity = 1 - Sum (tpi^2) where tpi is the frequency of the ith allele for the total & Sum tpi^2 is the sum of the squared total allele frequencies

m = mean

m*He* = mean of expected heterozygosity

m*Ho* = mean of observed heterozygosity

Fis = (Mean *He* - Mean *Ho*) / Mean *He*

Fit = (*Ht* - Mean *Ho*) / *Ht*

Fst = (*Ht* - Mean *He*) / *Ht*

5.3.2.3 Genetic similarity

Nei's genetic distance was used to estimate the genetic relationship between the 242 date palm accessions from nine populations: Oman, Italy (Sanremo, Bordighera), USDA-ARS, France, Iraq, Libya, Sudan and Iran. The genetic similarity ranged from 0.150 to 0.722, (Table 5.11). The highest similarity was observed between accessions from Sanremo and Bordighera, while the lowest similarity (most diversity) was observed between Oman and Bordighera accessions.

Table 5-11: The average genetic similarity between date palm accessions from Om: Oman (194 accessions), Sa: Sanremo (eight accessions), B: Bordighera (five accessions), US: USDA-ARS (five accessions), Frn: France (two accessions), Iq: Iraq (eleven accessions), Ly: Libya (seven accessions), Snd: Sudan (five accessions) and Iran (five accessions).

Pop.	Om	Sa	B	US	Frn	Iq	Ly	Snd	Iran
Om	1.000								
Sa	0.298	1.000							
B	0.150	0.722	1.000						
US	0.628	0.486	0.381	1.000					
Frn	0.403	0.627	0.531	0.634	1.000				
Iq	0.691	0.460	0.274	0.687	0.488	1.000			
Ly	0.375	0.579	0.468	0.557	0.535	0.531	1.000		
Snd	0.366	0.568	0.471	0.555	0.449	0.453	0.500	1.000	
Iran	0.566	0.320	0.301	0.603	0.379	0.666	0.404	0.350	1.000

5.3.2.4 Analysis of molecular variance of the combined accessions

The Analysis of Molecular Variance (AMOVA) indicated that (79%) of the molecular variation exists between accessions within populations, whereas (21%) exists between populations (Figure 5.6). Permutation tests (based on 999 permutations) showed that the overall ΦPT (ΦPT = 0.210, P = 0.001) was significantly different from the null distribution (Table 5.12).

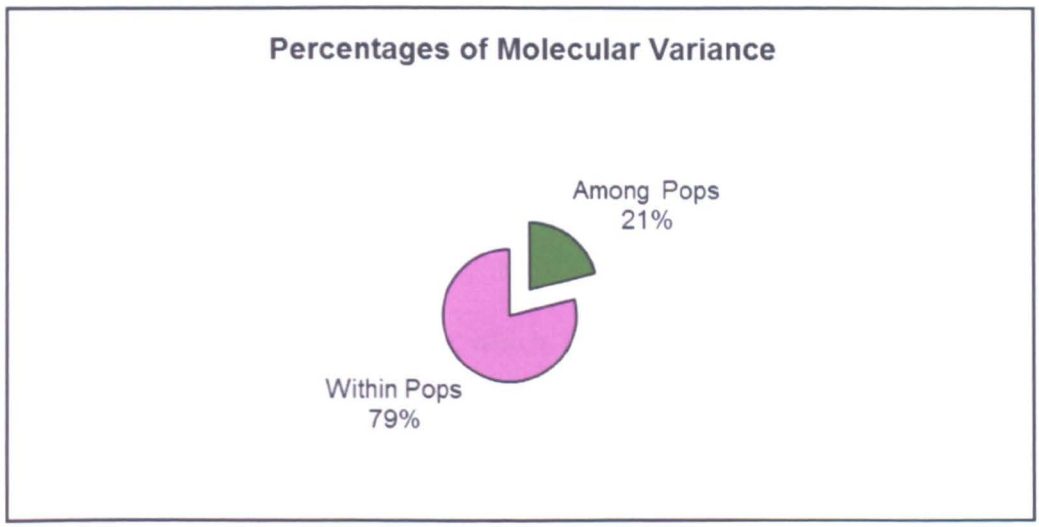


Figure 5.6: Distribution of the molecular variance within and among date palm accessions from Oman, Italy (Sanremo, Bordighera), USDA-ARS, France, Iraq, Libya, Sudan and Iran, obtained by microsatellite analysis

Table 5-12: Analysis of molecular variance for date palm accessions from Oman, Italy (Sanremo, Bordighera), USDA-ARS, France, Iraq, Libya, Sudan and Iran, obtained by 12 SSR primer pairs

Source	d.f.	Sum of Squares	MS	Est. Var.	Percentage of variation	*P* value
Among Pops	8	337.274	42.159	2.929	21%	0.001
Within Pops	234	2572.788	10.995	10.995	79%	
Total	242	2910.062		13.924	100%	

Stat	Value	P(rand >=data)
ΦPT	0.210	0.001

5.3.2.5 Associations among date palm accessions

Principal Coordinates Analysis (PCA) for molecular data from the 242 date palm accessions from Oman and other countries revealed a clear distribution of variation between the accessions as defined by the PC1 and PC2 which accounted for a combined 45.7% of the total variation. Figure 5.7 shows the

distribution of accessions on PC1 and PC2, accounting for 25.6 % and 20.2 % of the total variation, respectively.

Accessions from Europe-Africa: Sanremo, Bordighera, France, Libya and Sudan are completely separated by the first axis, except for one accession from Libya and one from Sudan which lie close to Iraqi accessions. The accessions from West-Asia: Oman, Iraq and Iran were closer to each other and well separated on the second axis. Furthermore, accessions from USDA-ARS were scattered according to their origin (Table 5.4). Medjool and Thory from Morocco and Algeria were placed within the Europe-Africa group, while Hilali, Barhee and Khalas from Oman, Iraq and Arabia, respectively, and were placed within the West-Asia group.

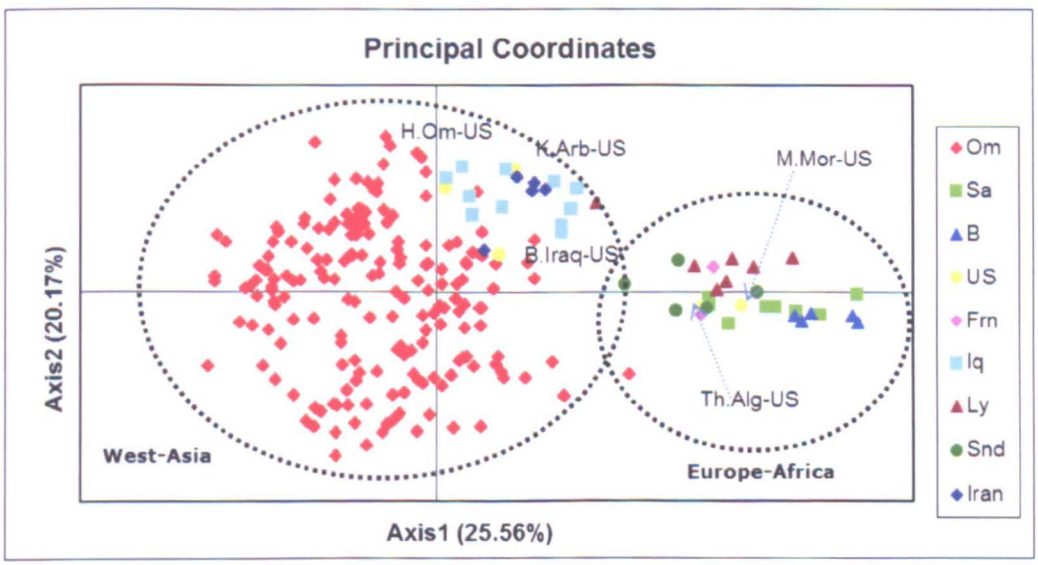


Figure 5.7: Principal Coordinates Analysis (PCA) of date palm accessions from Om: Oman, Sa: Sanremo, B: Bordighera, US: USDA-ARS (M.Mor: Medjool-Morocco, Th.Alg: Thory-Algeria, K.Arb: Khalas-Arabia, H.Om: Hilali-Oman and B. Iraq: Barhee-Iraq), Frn: France, Iq: Iraq, Ly: Libya, Snd: Sudan and Iran. Axis 1 accounted for 25.56% and Axis 2 accounted for 20.17% of the total molecular variation.

5.3.2.6 Cluster analysis

Unrooted UPGMA tree (Figure 5.8) and a bootstrap consensus tree (Figure 5.9) for 242 date palm accessions from Oman, Italy (Sanremo, Bordighera), USDA-ARS, France, Iraq, Libya, Sudan and Iran was generated by the UnWeighted Neighbor-Joining method using the DARwin 5.0 software and PowerMarker Ver. 3.25 (Liu and Muse, 2005), respectively. The results obtained from the cluster analysis (Figure 5.8 and 5.9) revealed a similar relationship when compared with PCA analysis with the same data set grouped by geographical region. The tree (Figure 5.8) displays two main groups: Europe-Africa and West-Asia showing a clear division between the two groups. All accessions from Italy (Sanremo, Bordighera), France, Libya, and Sudan were placed in the same group (Europe-Africa), while accessions from Oman, Iraq and Iran were placed in the West-Asia group (Figure 5.8). Some accessions from Oman (Unknown Male2, Unknown Male1, Unknown Male3, Rghad2, Rghad1, Al Maquidha2, Al Maquidha1, Rghad 3, Qash Al wa'b, Ghareef 2, Ghareef 1, Ghareef 4, Bu'Sab'ah 2, Bu'Sab'ah 1, Qash Hamreiyah, Qadmi, Do'wairah2, Do'wairah1) were closely located to accessions from Iraq and Iran. Two accessions from Libya and Sudan were placed close to Iraqi accessions, while accessions from USDA-ARS were placed according to their origin as shown in the PCA analysis. Similar results were also observed on bootstrap consensus tree (Figure 5.9).

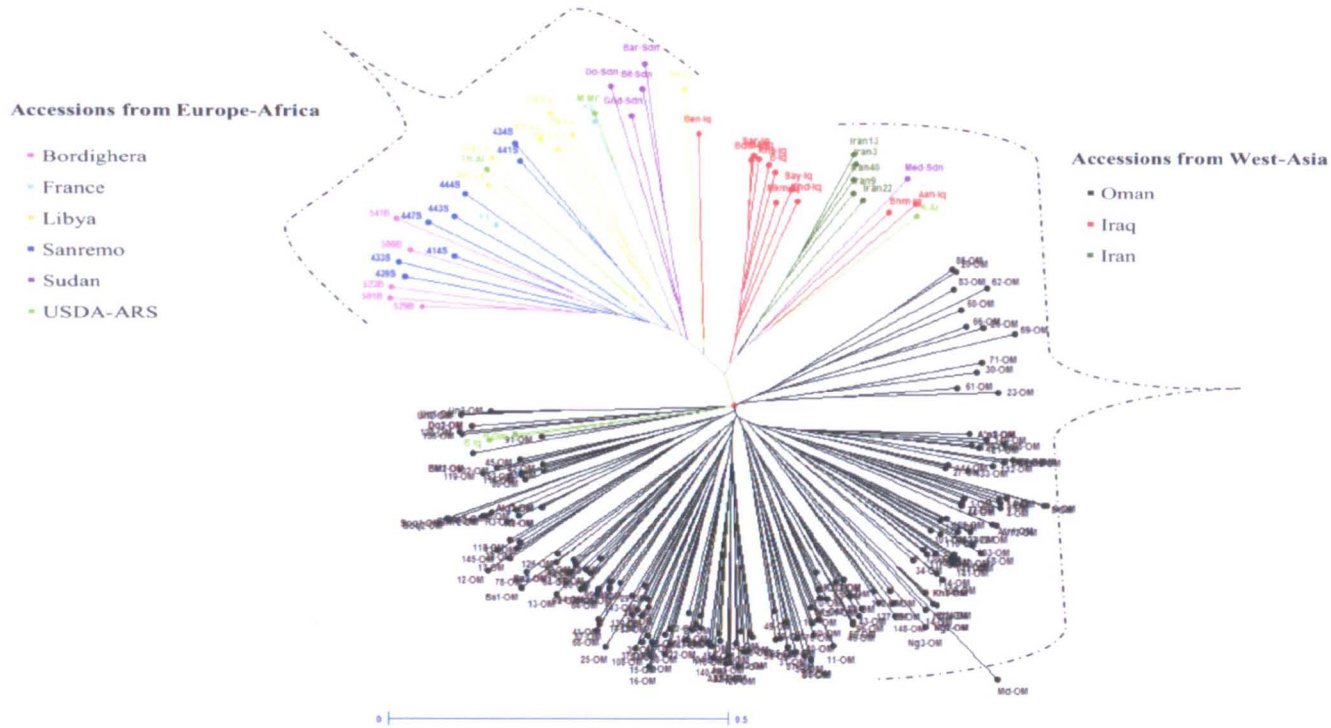


Figure 5.8: Unrooted UPGMA tree of 242 date palm accessions from Oman (female and male), Sa: Sanremo, B: Bordighera, Us: USDA-ARS, Frn: France, Iq: Iraq, Ly: Libya, Snd: Sudan, and Iran generated by the UnWeighted Neighbor-Joining method.

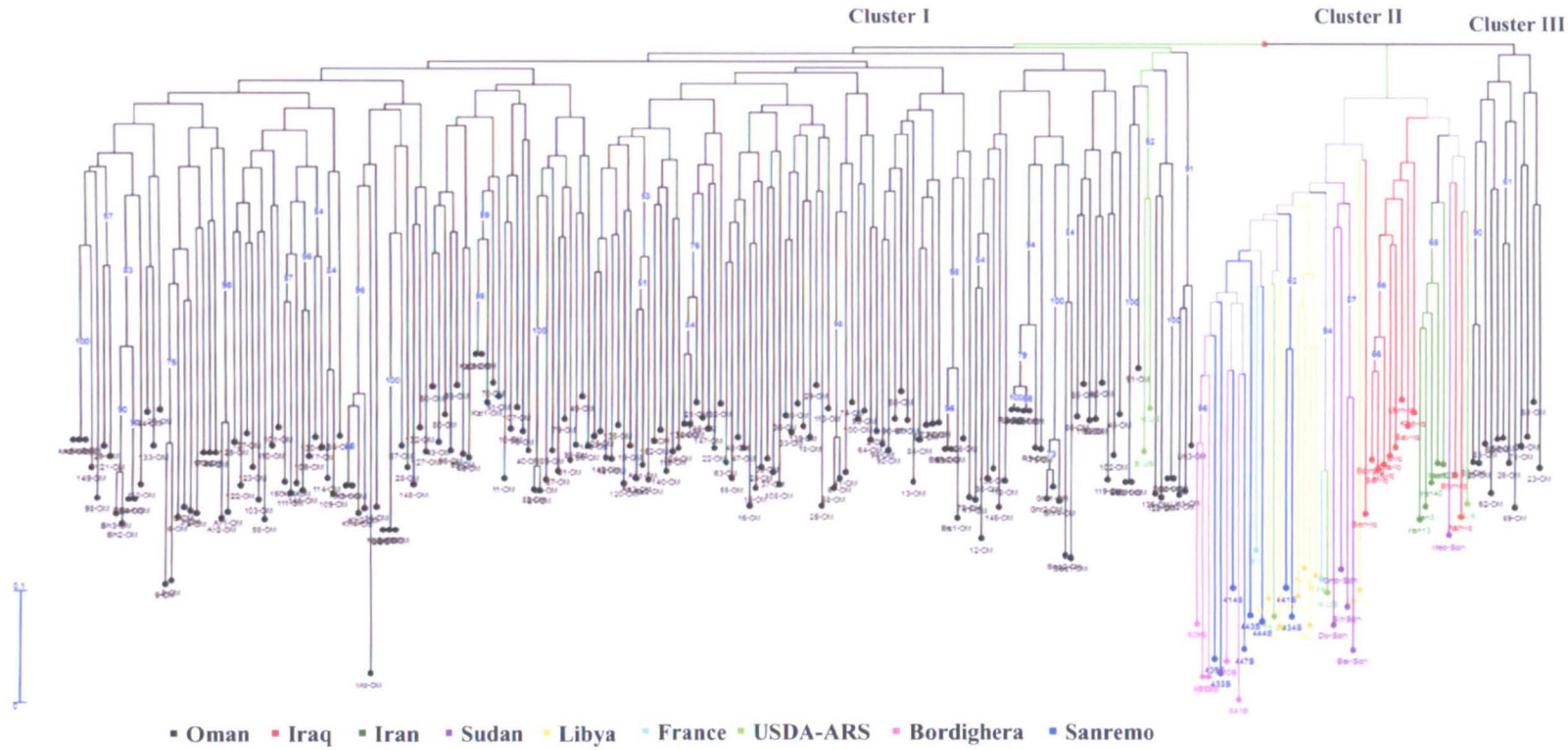


Figure 5.9: Phenetic analyses among 242 date palm accessions from Oman, Sanremo, Bordighera, USDA-ARS, France, Iraq, Libya, Sudan, and Iran was generated by the Neighbour-Joining analysis based on Nei's genetic distance. Bootstrap values have been computed over 1000 replications and were placed at the major nodes. Codes correspond to the names of accessions in Table 5.2, 5.3 and 5.4.

5.4 Discussion

5.4.1 Diversity analysis of Omani date palm cultivars

Genetic diversity and relatedness were measured between 151 female cultivars and 43 male trees of Omani date palm using 12 SSR markers. The twelve microsatellites used in this study were highly polymorphic revealing a total number of 188 alleles with a mean of 15.7 alleles per locus. The highest number of alleles (21) was amplified by locus PDCAT2, while the lowest (2) by locus mPdCIR057. The total number of alleles detected in this study were also higher than those found by Bodian *et al.* (2012), who noted 107 alleles with an average of 7.13 per locus from 128 date palm samples (11 female cultivars and 7 male trees) from different plantations areas of Figuig in Morocco, using 15 selected primers developed by Billotte *et al.* (2004) and Akkak *et al.* (2009). It is also high compared to Ahmed and Al-Qaradawi (2009) and Zehdi *et al.* (2004), who scored 40 and 100 alleles, respectively, when examining 15 Qatari and 46 Tunisian date palm accessions using 16 and 12 microsatellite loci. However, the allele number identified in this study was lower than those detected by Elshibli and Korpelainen (2008), who identified a total number of 343 alleles using 16 SSR markers from 37 female and 23 male accessions from Sudan with an average of 21.4 alleles per locus. The higher number of alleles in Sudanese date palm could possibly be attributed to the accessions from Sudan are more divergent than the Omani ones. The percentage of polymorphic loci in both female and male accessions was higher at 100% compared to the 96.11% reported by Bodian *et al.* (2012). A quite high level of heterozygosity was observed in accessions, varying from 0.241 to 0.870. Additionally, Bodian *et al.* (2012) remarked on a high level of

heterozygosity in Figuig oasis cultivars, varying from 0.684 to 0.930 using the same set of SSRs markers, thus confirming that these SSRs markers are useful tools for genetic diversity analysis of date palm germplasm.

For all SSR markers the observed heterozygosity values were less than the expected heterozygosity, with the exception of only one locus (mPdCIR057) which had a higher observed heterozygosity. This suggests that the Omani germplasm may be, to some extent, isolated and has not received free external gene flow. In contrast, Bodian *et al.* (2012) found that the observed heterozygosity in Figuig oasis cultivars was higher than the excepted heterozygosity. This finding seems to indicate that Figuig oasis cultivars have had greater than average external gene flow, which is likely the cause of the excess heterozygosity.

The average of *Fis* and *Fit* was 0.173 and 0.190, respectively, which suggests that genetic deviation from Hardy–Weinberg expectation has occurred between and within Omani female and male accessions. The *Fst* was 0.021 on average, implying that the within female plus male genetic variation (97%) was higher than that of the between female and male accessions (2.1%). This low level of *Fst* indicates that although some visible differences exist between cultivars, in general, Omani cultivars are very closely related and a high degree of genetic similarity (0.866) was observed between female and male accessions as presented in this study. This is not surprising considering the breeding histories and multiplication methods (seeds, offshoots and tissue culture) of these cultivars as well as the limited seed sources introduced from other germplasm to Oman, which could have shaped the current structure of Omani germplasm.

Six percent of molecular variation was between the Omani female and male populations, with the majority being within population with an estimated genetic variation component of 94% (AMOVA, $P < 0.001$), much lower than 59% of variability found between the date palm cultivars in Morocco (Bodian *et al.*, 2012). Furthermore, the value ΦPT obtained from AMOVA (ΦPT = 0.057) is analogous to Wright's *Fst* statistic and is in agreement with the values obtained from the direct estimation of *Fst* (*Fst* = 0.021) because this also indicates a small but significant genetic differentiation. The value ΦPT obtained from AMOVA (ΦPT = 0.057) is however much higher than the analogue F value (*Fst* = 0.021) and would predict little differentiation between male and female palms. The AMOVA (ΦPT) takes into account correlations among loci which is not considered when calculating the locus x locus *Fst* (Peakall & Smouse, 2006). AMOVA is based on individual x individual analysis and tries to find correlations from genotypes rather than from independent loci. Excoffier *et al.* (1992) reported that AMOVA is more sensitive to differences between populations due to co-evolution of genes and adaptation to local selection pressures. AMOVA does not take into account the expected heterozygosity for each locus but it takes individuals as haploid, therefore the obtained estimates of genetic differentiation between populations will be higher (the AMOVA ΦPT will be higher than the Wright's *Fst* estimate).

The Principal Coordinates Analysis (PCA) identifies the main trends within the marker data, rather than utilising the complete data. The first and second principal component (PC1 and PC2) accounted for 44.32% of the variation and the male palms are more constrained on axis 2 than the female palms. The

PCA graph corresponded very well to the cluster analysis (NJ tree-bootstrap 1000 replication), indicating that the Omani accessions were closely related to each other and the genetic polymorphism among them was found to be relatively low. There was no clear genetic differentiation between female and male cultivars. This is in accordance with the findings of Bodian *et al.* (2012) who found that there was no genetic differentiation between male and female cultivars in Morocco. Haider et al. (2012) justified that by the exchange of accessions between the different plantation areas, development of new males by seedling selection, clonal propagation of ecotypes and limited sexual reproduction as well as farmer selection for specific genotypes.

5.4.2 Diversity analysis of Omani accessions and comparison with germplasm from other countries

In this section we combined accessions from Oman, Sanremo, Bordighera, USDA-ARS, France, Iraq, Libya, Sudan and Iran together to examine genetic diversity and studied the genetic relationships between date palms from different origins.

The SSR primer pairs selected were highly polymorphic. A total of 246 alleles with a mean of 20.5 alleles per locus were scored. These results are slightly higher when compared to analysis of Omani accessions in Section 5.3.1 which possessed a total of 188 alleles with an average 15.7 alleles per locus.

Kotzé and Muller (1994) have defined heterozygosity as a measure of genetic variation within a population. High levels of heterozygosity were observed in this study with a total rate of 0.817 ranging from 0.679 to 0.873 for each marker. The high level of heterozygosity for these populations could be explained by one of the following reasons: long-term natural selection for

adaptation, the mixed nature of the populations or historic mixing of individuals of different populations.

In this study, the average of observed heterozygosity for loci mPdCIR010, mPdCIR015, mPdCIR050, mPdCIR057, mPdCIR093, PDCAT2, PDCAT11 and PDCAT14 was greater than the average of expected heterozygosity. Based on Hardy-Weinberg equilibrium (HWE), if observed heterozygosity is higher than expected, it is highly likely to be due to an isolate-breaking effect. This could also be referring to number of samples for each population used in this study. Therefore, the results may then be changed by adding an equivalent number of genotypes in each population.

There was also a quite high degree of genetic differentiation observed among all populations as measured by *Fst* (19.7 %) compared with variation among the Omani accessions at only 2.1% of the total variation.

The AMOVA analysis indicates significant genetic differentiation between populations (21%), although most genetic variation still existed within populations 79%. The value ΦPT obtained from AMOVA (ΦPT = 0.210, P = 0.001) is analogous to Wright's *Fst* statistic and is in agreement with the values obtained through direct estimation of *Fst* (*Fst* = 0.197) indicating that there is significant genetic difference between populations. The value ΦPT obtained from AMOVA (ΦPT = 0. 210) is however higher than the analogue F values (*Fst* = 0.197) suggesting a relatively high level of gene flow exists between these populations.

The highest similarity value was observed between accessions from Sanremo and Bordighera. The high similarity indicates that although some visible differences obviously exist between accessions, in general, Sanremo and

Bordighera accessions are very closely related and share a high degree of genetic similarity. Furthermore, the greatest divergence was detected between Omani and Bordighera cultivars.

The PCA analysis showed that accessions from Europe-Africa: Sanremo, Bordighera, France, Libya and Sudan are completely separated from the other accessions, except for one accession from Libya and one from Sudan which placed close to Iraqi accessions. However, accessions from West-Asia: Oman, Iraq and Iran were placed close to each other and well separated on the second axis. Accessions from USDA-ARS were placed in accordance to their origin; Medjool and Thory from Morocco and Algeria were placed within the Europe-Africa group, while Hilali, Barhee and Khalas from Oman, Iraq and Arabia, respectively, and were placed within the West-Asia group. A comparable region-specific cluster was also evident in the unrooted dendrogram tree, which showed two main groups Europe-Africa and West-Asia. The similar result of the unrooted dendrogram, bootstrap consensus tree and PCA plots reflect the geographic relationships between the studied accessions and suggest that Sanremo, Bordighera, France, Libya and Sudan (Europe-Africa) have different germplasm pools than Oman, Iraq and Iran (West-Asia) which should each be conserved.

5.5 Conclusions

The following main conclusions can be drawn from this analysis:

- The present study has demonstrated that the selected twelve SSR markers (Billotte *et al.*, 2004; Akkak *et al.*, 2009) provided an effective tool for assessing genetic diversity and relationships among and within

date palm germplasm and were useful in differentiating between closely related germplasm sources.

- Approximately 250 varieties of date palm are grown throughout the Sultanate (MAF, 2005; Al-Khatri, 2004) and this study has provided the first molecular identification key, which enables the unambiguous discrimination of 194 Omani date palm accessions (151 female cultivars and 43 male trees).
- The average of *Fst* was 0.021, implying that the genetic variation within Omani female and male accessions was 97%, while between female and male accessions was only 2.1%, indicating that the Omani female and male accessions have little consistent divergence, compared to the large-scale divergence between the accessions themselves. However, the limited number of date palm male trees used in this study was not fully representative of the genetic diversity level present in date palm male trees in Oman. The close similarity could be attributed this to Oman being a unit regarding cultural practices, exchange of plant material, the breeding histories, multiplication methods used and limited seed sources introduced from other germplasm to Oman.
- The genetic differentiation between the Omani female and male accessions was also estimated with AMOVA analysis, resulted in Six percent of molecular variation between the female and male accessions, with the majority being within population with an estimated genetic variation component of 94%.
- In addition, the value ΦPT obtained from AMOVA (ΦPT = 0.057) was in agreement with the values obtained from the direct estimation of *Fst*

(*Fst* = 0.021) which also indicated a small but significant genetic differentiation.

- The Principal Coordinates Analysis (PCA) showed overlapping between the Omani female and male accessions, where the male accessions are more constrained on axis 2 than the female accessions.
- The bootstrap consensus phenetic trees was generated using 1000 replications showed that the Omani accessions were closely related to each other and there was no clear genetic differentiation between female and male cultivars.
- There was a quite high degree of genetic differentiation observed between germplasm from Oman, Sanremo, Bordighera, USDA-ARS, France, Iraq, Libya, Sudan and Iran as measured by *Fst* (19.7 %) compared with the genetic differentiation observed among the Omani accessions (2.1%) of the total variation, which probably reflects the homogeneous nature of the Omani date palm used in this study comparing to the divergent sets of other germplasm.
- The AMOVA analysis showed significant genetic variation between all populations (21%), although most genetic variation still existed within populations 79%.
- High similarity value was observed between accessions from Sanremo and Bordighera, while high divergence was observed between Omani and Bordighera accessions.
- The study confirms that the Europe-Africa (Sanremo, Bordighera, France, Libya and Sudan) accessions are distinguished from West-Asia

(Oman, Iraq and Iran) accessions, have their own autochthonous origin, a finding which was strongly validated by bootstrap consensus tree test.

- Furthermore, accessions from USDA-ARS were placed according to their origin. Medjool and Thory from Morocco and Algeria were placed within the Europe-Africa group, while Hilali, Barhee and Khalas from Oman, Iraq and Arabia, respectively, and were placed within the West-Asia group in accordance to their origin.
- The unrooted dendrogram and bootstrap consensus tree clustered the accessions into two and three main groups, respectively, which were in accordance with their geographic origin but not with their sexuality.
- Our results provide evidence for the possibility of using these markers as descriptors in the certification and control of origin labels for date palm material.

Chapter 6. GENETIC MAPPING OF DATE PALM

6.1 Introduction

Genetic mapping (also known as linkage mapping or meiotic mapping) of many plant species have been achieved and utilized for various applications. Markers in genetic maps can also be used in fingerprinting applications as well as providing a way to test and track the co-segregation of markers with traits in segregating populations. Such linked markers can be used in selection of genes responsible for agronomically important traits, therefore facilitating crop improvement. Markers can also be used in different comparative studies to understand the processes that led to the diversification and evolution of a species (Cone and Cone, 2009). High density maps could potentially be used as the starting point for isolation and cloning of genes of interest (Ma, 2003; Mohan *et al.*, 1997).

The first genetic map was presented back in 1911 when T.H. Morgan and his students were able to demonstrate the location of six different sex-linked chromosome genes in fruit fly (*Drosophila melanogaster*) (reported in Semagn *et al.*, 2006). Arrangement of these markers in relative order based on their genetic distances (due to strength of co-inheritance) is called genetic mapping.

Different molecular markers types (and combination of molecular markers) have been used to construct linkage maps of various plant species, for example; RFLP (maize; Coe *et al.*, 2002, sorghum; Draye *et al.*, 2001), AFLPs (*Arabidopsis*; Peters *et al.*, 2001, papaya; Blas *et al.*, 2009), SSRs (maize; Sharopova *et al.*, 2002, soybean; Fu *et al.*, 2006, wheat; Roder *et al.*, 1998, sunflower; Tang *et al.*, 2002, rice; Wu and Tanksley 1993). According to

Semagn *et al.* (2006) AFLP is the most commonly used marker system to generate large numbers of markers for the construction of high-density genetic maps. However, Simple Sequence Repeat (SSR) markers remain as a standard for map construction due to their transferability between populations and their high levels of polymorphism between closely related lines. More recently, DArT and SNPs have been used in a number of genetic maps, including; wheat (Akbari *et al.*, 2006), barley (Wenzl *et al.*, 2004), rice (Jaccoud *et al.*, 2001), cassava (Xia *et al.*, 2005), *Arabidopsis* (Wittenberg *et al.* 2005), sugarcane (Bundock *et al.*, 2009) and chickpea (Gaur *et al.*, 2012). DArT is a polymorphic and reproducible marker generation method, however, their inheritance as dominant markers is still a limitation for mapping (Semagn *et al.*, 2006), although they are often transferable between different crosses, due to the hybridisation-based detection.

The genetic mapping of date palm (*Phoenix dactylifera* L.) lagged behind that of many other plant species because of long generation times, which may require >30 years to generate a backcrossing program (Al-Dous *et al.*, 2011). In addition, very poorly developed date palm crossing programmes mean that only very few controlled crosses are available as the starting material for the construction of the maps. Date palm is diploid with 18 pairs of chromosomes having a genome size of approximately 658 Mb (Al-Dous *et al.*, 2011). In the Middle East and North Africa, date palm is considered as one of the most important woody crops as well as a good candidate for improving agricultural yields in arid environments. Biotic (disease and pest) and abiotic (drought and salinity) factors have been found to limit date palm production in many areas around the world. Date Palm is also threatened by genetic erosion because

most date palm growers tend to cultivate specific cultivars with high commercial value and ignore other less valuable cultivars. This could affect the genetic diversity available for date palm improvement (El Kharbotly *et al.*, 2006).

Considering the importance of date palm and to assist in conservation of the germplasm, it is important to construct a date palm genetic map as a first and an essential step for conducting extensive genomic research (and representative genetic sampling of germplasm) for this crop. Constructing a genetic map for date palm would be helpful to allow screening for genetic markers close to the genes that control traits of interest such as yield or disease resistance as well as to develop markers able to distinguish between male and female palms before flowering and reduce the duration of the breeding cycle.

The main objective of this study was to construct initial genetic maps of date palm based on the available (small) populations (BC_1 and F_1).

6.2 Materials and methods

6.2.1 Mapping population

The male parent KI-96-13 was selected due to synchronized flowering with the mother cultivar Khalas-4 and was used to generate a controlled cross. Khakas-4 was selected as the recurrent parent as a cultivar producing high quality date fruit. This gave rise to the BC_1 population which consists of 53 individuals and is suitable for genetic mapping studies. KI-96-13 was also crossed with an Um-Assela cultivar palm to generate an F_1 population of 30 individuals (El Kharbotly *et al.*, 2006). The Um-Assela cultivar is known for its low quality

date fruit, but it is well adapted to the conditions in the coastal regions of Oman (high salinity and humidity). The BC_1 and F_1 populations were developed in 1996 and maintained in the Date Palm Research Station at Wadi Quriate and the Research Farm at Barka, respectively (Table 6.1). Most individuals in the BC_1 and F_1 population have reached the flowering stage and gender is known for all except one palm.

6.2.2 DNA extraction

Total genomic DNA was extracted from the young leaves of the 83 individuals of the BC_1 and F_1 populations along with their three parents using the DNeasy plant Maxi kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions and quantified on 1% agarose gel as described earlier in Chapter 3; Section 3.4.

6.2.3 SSR and SNP marker analysis

Eleven new SSRs markers plus the sex determination marker PDK_30s101A (locus 145) developed in this current study, along with 61 SSRs markers developed by others and from SSRs developed in this study from supplied primer sequences (Table 6.2; Billotte *et al.*, 2004; Akkak *et al.*, 2009; Hamwieh *et al.*, 2010) were screened for polymorphism. All primer pairs were first screened for their polymorphism against the parents of the mapping populations. The selected polymorphic markers were then used to genotype the individuals of the mapping populations. PCR reactions for SSR marker analysis were performed in a total reaction mixture of 20 µl following the procedures described in Chapter 3; Section 3.9 and 3.10.

SNPs marker assays for BC_1 and F_1 individuals plus their parents were performed by DArT Pty. Ltd. (Yarralumla, Australia; www.diversityarrays.com) (Wenzl *et al.* 2004; Akbari *et al.* 2006; Semagn *et al.* 2006) using the DArT Seq approach.

Table 6-1: Lists of 86 date palm samples from the BC_1 and F_1 populations used for genetic mapping.

Lab Code	Accession Name	Gender	Lab Code	Accession Name	Gender
Khalas 4	Parent1 BC_1	F	46B	BC_1	F
Kl-96-13	Parent2 BC_1 & F_1	M	47B	BC_1	M
Um.SQU	Parent1 F_1	F	48B	BC_1	M
1B	BC_1	M	49B	BC_1	M
2B	BC_1	F	51B	BC_1	F
3B	BC_1	M	52B	BC_1	F
4B	BC_1	M	53B	BC_1	F
5B	BC_1	F	54B	BC_1	M
6B	BC_1	M	55B	BC_1	M
7B	BC_1	M	57B	BC_1	M
8B	BC_1	M	58B	BC_1	M
10B	BC_1	M	59B	BC_1	F
11B	BC_1	M	60B	BC_1	F
12B	BC_1	M	1F	F_1	M
13B	BC_1	M	2F	F_1	M
14B	BC_1	F	6F	F_1	F
15B	BC_1	M	7F	F_1	M
16B	BC_1	F	8F	F_1	M
17B	BC_1	F	9F	F_1	M
18B	BC_1	M	10F	F_1	F
19B	BC_1	F	13F	F_1	M
20B	BC_1	M	14F	F_1	F

23B	BC_1	F	15F	F_1	F
24B	BC_1	F	16F	F_1	F
25B	BC_1	M	17F	F_1	F
26B	BC_1	F	18F	F_1	F
27B	BC_1	F	19F	F_1	F
28B	BC_1	M	20F	F_1	F
29B	BC_1	F	21F	F_1	F
30B	BC_1	M	22F	F_1	F
32B	BC_1	M	23F	F_1	M
33B	BC_1	F	24F	F_1	M
34B	BC_1	M	25F	F_1	M
35B	BC_1	F	26F	F_1	M
36B	BC_1	M	27F	F_1	F
37B	BC_1	F	28F	F_1	F
38B	BC_1	F	29F	F_1	F
39B	BC_1	M	31F	F_1	M
40B	BC_1	F	34F	F_1	M
42B	BC_1	F	35F	F_1	F
43B	BC_1	Unknown	37F	F_1	M
44B	BC_1	F	41F	F_1	F
45B	BC_1	F	42F	F_1	F

F=female, M=male

Table 6-2: List of 73 microsatellite primers polymorphic within Omani germplasm and their annealing temperature used in this study.

Marker name	Annealing Tm (°)	Marker name	Annealing Tm (°)
DateS1	50°C	DPALM315	58°C
DateS8	55°C	DPALM319	55°C
DateS9	55°C	DPALM325	55°C
DateS12	55°C	DPALM327	55°C
DateS16	55°C	DPALM328	55°C
DateS17	55°C	DPALM332	55°C
DateS103	55°C	DPALM333	55°C
DateS110	52°C	DPALM336	55°C
DateS111	52°C	DPALM340	55°C
DateS130	52°C	DPALM341	55°C
DateS131	52°C	DPALM342	55°C
mPdCIR010	52°C	DPALM343	58°C
mPdCIR015	52°C	DPALM344	58°C
mPdCIR016	52°C	DPALM348	55°C
mPdCIR025	52°C	DPALM349	55°C
mPdCIR050	52°C	DPALM350	61°C
mPdCIR057	52°C	DPALM352	55°C
mPdCIR078	52°C	DPALM357	57°C
mPdCIR085	52°C	DPALM361	55°C
mPdCIR093	52°C	DPALM362	57°C
PDCAT2	55°C	DPALM363	58°C
PDCAT5	55°C	DPALM366	57°C
PDCAT10	55°C	DPALM369	55°C
PDCAT11	55°C	DPALM374	55°C
PDCAT12	55°C	DPALM377	61°C
PDCAT14	55°C	DPALM378	55°C
PDCAT17	55°C	DPALM379	55°C
PDCAT18	55°C	DPALM380	55°C
PDCAT20	55°C	DPALM388	55°C
PDCAT21	55°C	DPALM398	50°C
DPALM302	55°C	DPALM402	55°C
DPALM303	55°C	DPALM404	55°C
DPALM305	55°C	DPALM405	55°C
DPALM307	58°C	DPALM408	55°C
DPALM309	58°C	DPALM410	55°C
DPALM311	50°C	PDK_30s101a	60°C
DPALM312	58°C		

6.2.4 Map construction

While both populations are defined types (BC_1 and F_1) they are made by crossing between out-crossing parental lines which have high levels of heterozygosity. As such they do not fit the classical models for F_1 and BC_1 which are usually constructed from inbred parental lines. A linkage map for each population was constructed with the JoinMap4.1 software (Van Ooijen, 2006) by combining data from SSR and SNP markers.

Phase determination for SSR and SNP markers was carried out by analyzing both BC_1 and F_1 populations based on their heterozygous parents, however, mapping for these markers in both populations was carried out using the Cross Pollinator (CP) model (Tables 6.3 and 6.4) as the BC_1 model cannot be applied because the BC_1 cross does not fit the classical mode.

The "Locus genotype frequency" function was applied to calculate chi-square values for each marker to test for expected segregation patterns. Markers were placed into linkage groups using the "LOD groupings" and "Create groups for mapping" command with the Kosambi mapping function (Kosambi, 1944). Calculation parameters were set for a minimum LOD threshold of 3.0, and recombination fraction of 0.250. Markers order within groups was using the "Calculate Map" command.

Table 6-3: Segregation type codes for the population type CP.

Code	Description
<abxcd>	locus heterozygous in both parent, four alleles
<efxeg>	locus heterozygous in both parent, three alleles
<hkxhk>	locus heterozygous in both parent, two alleles
<lmxll>	locus heterozygous in the first parent
<nnxnp>	locus heterozygous in the second parent

Table 6-4: Genotype codes for a "CP" population, depending on the locus segregation type.

Seg. Type	Possible genotypes
<abxcd>	ac, ad, bc, bd, --
<efxeg>	ee, ef, eg, fg, --
<hkxhk>	hh, hk, kk, h-, k-, --
<lmxll>	ll, lm, --
<nnxnp>	nn, np, --
Remarks	
1	each character "a" to "p" represents a distinct allele; "-" means unknown allele
2	"h-" and "k-" are dominant genotypes: "h-" means either "hh" or "hk" "k-" means either "kk" or "hk"
3	"." and u are treated equivalent to "-"
4	the software is indifferent to the order of alleles in the codes, e.g. "hk" is equivalent to "kh"

6.3 Result

6.3.1 Polymorphism and markers for mapping

Polymorphism of SSRs and SNPs markers in the BC_1 population

SSR were first screened on the parents and a subset of the progeny of each cross, in order to identify the polymorphic markers. Parental samples were included with the populations for DArT Seq analysis, allowing phase to be determined in the generated population data.

SSR markers

A total of 73 SSR primer pairs were screened against the parents, Khalas-4 and KI-96-13. Among these, 50 primer pairs were polymorphic (68.5%) and discriminated between the parental alleles of the cross. Forty-two polymorphic primer pairs were used to screen all individuals in this population.

SNP markers

Different criteria have been used to select the best SNP markers to be considered in our analysis. The returned SNP data contain 10,557 pairs at different levels of confidence. For each cross, marker with more than five missing data were excluded from the input data before map construction in order to obtain a high quality data set. Each SNP marker should have two lines (reference and variant); combining both lines of the same marker should result in a co-dominant case for individual genotypes. Any marker having more than two lines (tri or tetra lines) was excluded from the analysis for the moment. After filtering data based on quality, polymorphism and missing data and subset of data was selected.

Out of potentially 10,557 SNPs pairs developed by DArT Pty. Ltd across the two populations and a limited numbers of cultivar genotypes, 973 (9.2%) were initially identified in the BC_1 population as being of highest quality. However, in an attempt to produce an initial outline map with minimal missing data 42 SSR and 285 SNP markers were used to construct an initial genetic map for the BC_1 population.

Polymorphism of SSRs and SNPs markers in the F_1 population

SSR markers

The same primer set of 73 SSR markers were tested for polymorphism in F_1 population, 50 polymorphic primers were detected (68.5%), in which 42 (84%) primers overlapped with the BC_1 population. Parental alleles of this cross were scored for all the individuals using 50 polymorphic SSR primers.

SNP markers

Similar criteria were used to select the best SNP makers for F_1 cross as mentioned in BC_1. Out of 10,557 SNPs pairs developed by DArT Pty. Ltd across both populations and a limited number of cultivar genotypes, 1,205 (11.4%) were identified as polymorphic markers in this cross and confirmed in the F_1 population. Twenty-two polymorphic SSRs plus 619 SNPs markers were used to construct a genetic map for F_1 population as an initial start, which consists only of 30 individuals.

6.3.2 Segregation distortion for BC_1 and F_1 populations

Marker segregation patterns were determined and their potential distortion observed in both populations by JoinMap 4.1 by performing a Chi-square test against expected segregation patterns ($p<0.05$ for significance). Five different allele patterns of segregation were observed for SSR and SNP markers in BC_1 and F_1 crosses (Table 5.6).

In the BC_1 population, the locus genotype frequency table (Appendix 5) suggested that out of 327 markers, 203 (62.1%) markers segregated in the expected Mendelian ratios of 1:1, 1:2:1 and 1:1:1:1 for both marker types SSR and SNP, while 124 markers (37.9%) deviated from the expected Mendelian ratio.

Out of 42 SSR markers used, 33 (78.6%) segregated in the expected Mendelian ratios (1:1, 1:2:1 and 1:1:1:1), where 9 (21.4%) were deviated from the expected ratio (Table 6.5). Out of 33 SSR markers, the total number of markers segregating 1:1 was 12 (36.4%), 15 (45.5%) segregated 1:2:1 and six (18.2%) segregated in a 1:1:1:1 ratio (Table 6.5).

Out of 285 SNP markers, 170 (59.6%) markers segregated in the expected Mendelian ratio 1:2:1, while 115 (40.3%) markers were distorted from this ratio.

In the F_1 population, 403 (62.9%) of the markers evaluated segregated in the expected Mendelian ratios 1:1, 1:2:1 and 1:1:1:1, while 238 (37.1%) showed deviation from the expected ratios for both SSR and SNP markers (Table 6.5, Appendix 6). Out of 22 SSR marker, 19 (86.4%) were segregated in the expected Mendelian ratio (1:1, 1:2:1 and 1:1:1:1) while 2 (9.1%) deviated from the expected ratio (Table 6.5). Out of 19 SSR, seven (36.8%) were segregating 1:1, four (21.1%) segregating 1:2:1, while eight (42.1%) segregated 1:1:1:1. The total number of SNP marker segregating 1:2:1 was 384 (62%) out of 619.

Table 6-5: Genotype configurations and distribution of segregating marker loci observed in BC_1 and F_1 populations.

Segregating marker alleles	Parent genotype		Progeny genotype classes		Number of segregating marker loci in the cross			
					BC1		F1	
	P1	P2	Allelic pattern	Segregation ratio	SSR	SNP	SSR	SNP
2 alleles	lm	ll	ll, ll, ml, ml	1:1	2	0	1	
	nn	np	nn, np, nn, np	1:1	10	0	6	
	hk	hk	hh, hk, hk,kk	1:2:1	15	170	4	384
3 alleles	ef	eg	ee, eg, fe, fg	1:1:1:1	6	0	6	
4 alleles	ab	cd	ac, ad, bc, bd	1:1:1:1	0	0	2	
Total					33	170	19	384
					203		403	

6.3.3 Linkage mapping and marker distribution

In the BC_1 map, out of 327 markers, 270 (82.6% grouped; 28 SSR and 242 SNP) could be assigned to 29 linkage groups (LG1-LG29; Figure 6.1) which had between 2 and 27 markers per group and a linkage group length that varied from 3.9 cM (LG26) to 101.8 cM (LG12). 0 to 2 SSR markers were present per linkage group, while the SNP marker number varied from 1 to 19 markers per linkage group (Table 6.6). The maps of linkage groups 1, 2, 3, 4, 6, 7, 9 and 27 required two to three rounds of mapping using the regression approach, but the round one (Map1 version) was used from these groups to assemble the combined linkage map in the segregating BC_1 population. These groups showed large *jump* threshold values for markers which were unmapped in round 1 (Map 1) and round 2 (Map 2), before they were forced into map 2 to create map 3. This indicates a poor fit of these added markers and so all added markers from round 3 were removed from the final map.

The linkage map spanned a total genetic distance of 1,486.7 cM with 57 (17%) markers remaining unlinked to at least one other marker. Markers were randomly distributed on the linkage groups. The distance between adjacent markers on the map also varied greatly across the different linkage groups. The average marker distance was 6.74 cM, with intervals between markers ranging from 1.87 cM to 16.2 cM (Table 6.6). The sex determination locus PDK_30s101A (coded as locus 145) was mapped in linkage group 'LG18' at 42.8 cM.

Table 6-6: Distribution of the markers, linkage group size and marker density in the genetic map constructed in the BC_1 population.

Linkage groups	Length (cM)	No. of markers mapped in the groups			Average marker interval (cM)
		Total marker	SSR marker	SNP marker	
1	75	21	2	19	3.57
2	84.1	20	1	19	4.21
3	39.3	10	1	9	3.93
4	75	17	2	15	4.41
5	51.5	13	1	12	3.96
6	50.6	27	0	27	1.87
7	67.6	10	1	9	6.76
8	81	13	1	12	6.23
9	66.3	14	2	12	4.74
10	56.1	13	1	12	4.32
11	70	13	2	11	5.38
12	101.8	11	0	11	9.25
13	51.6	11	2	9	4.69
14	33.2	9	0	9	3.69
15	77.3	7	1	6	11.04
16	52.7	6	0	6	8.78
17	41.6	6	1	5	6.93
18	50.8	6	2	4	8.47
19	25.3	6	1	5	4.22
20	42.8	4	1	3	10.70
21	35.9	4	1	3	8.98
22	46.2	4	0	4	11.55
23	15.4	3	1	2	5.13
24	38.2	3	0	3	12.73
25	15.9	2	0	2	7.95
26	3.9	2	1	1	1.95
27	41.2	7	1	6	5.89
28	31.6	4	1	3	7.90
29	64.8	4	1	3	16.20
Total	1486.7	270	28	242	-
Range	3.9-101.8	2-27	0-2	1-19	1.87-16.20

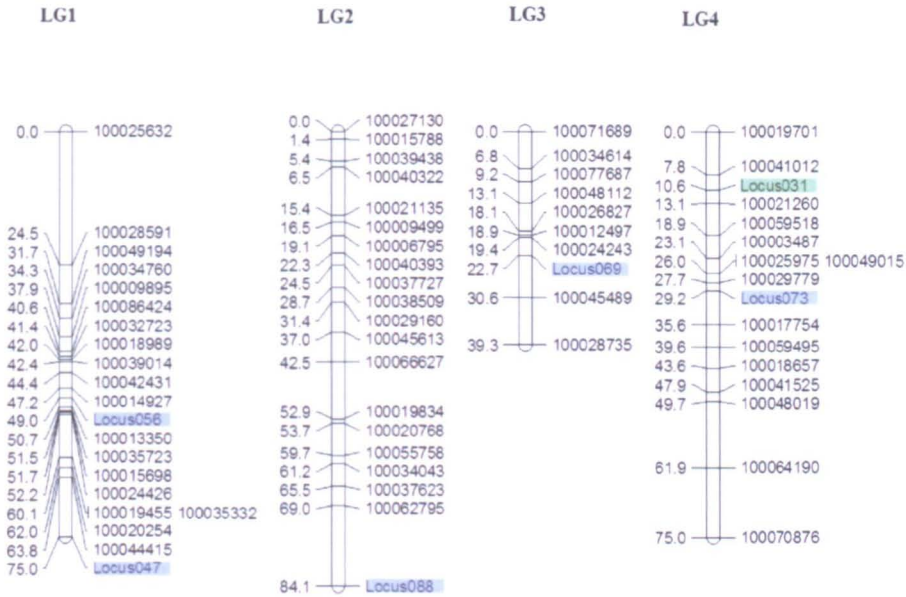


Figure 6.1: A genetic linkage map of 29 linkage groups. This was constructed in 53 BC_1 individuals derived from the cross between Khalas-4 and KI-96-13. The locations of 28 SSR and 242 SNPs markers are given. Positions are given in centimorgan (Kosambi units) to the left of the linkage groups and the name of markers to the right. The SSRs markers are color coded by their source (*green*; Billotte *et al.* (2004) and Akkak *et al.* (2009), *blue*; Hamwieh *et al.* (2010) primers sets tested in this study, *red*; SSRs developed in this study, *orange*; sex determination locus). All SNPs markers are in *black*.

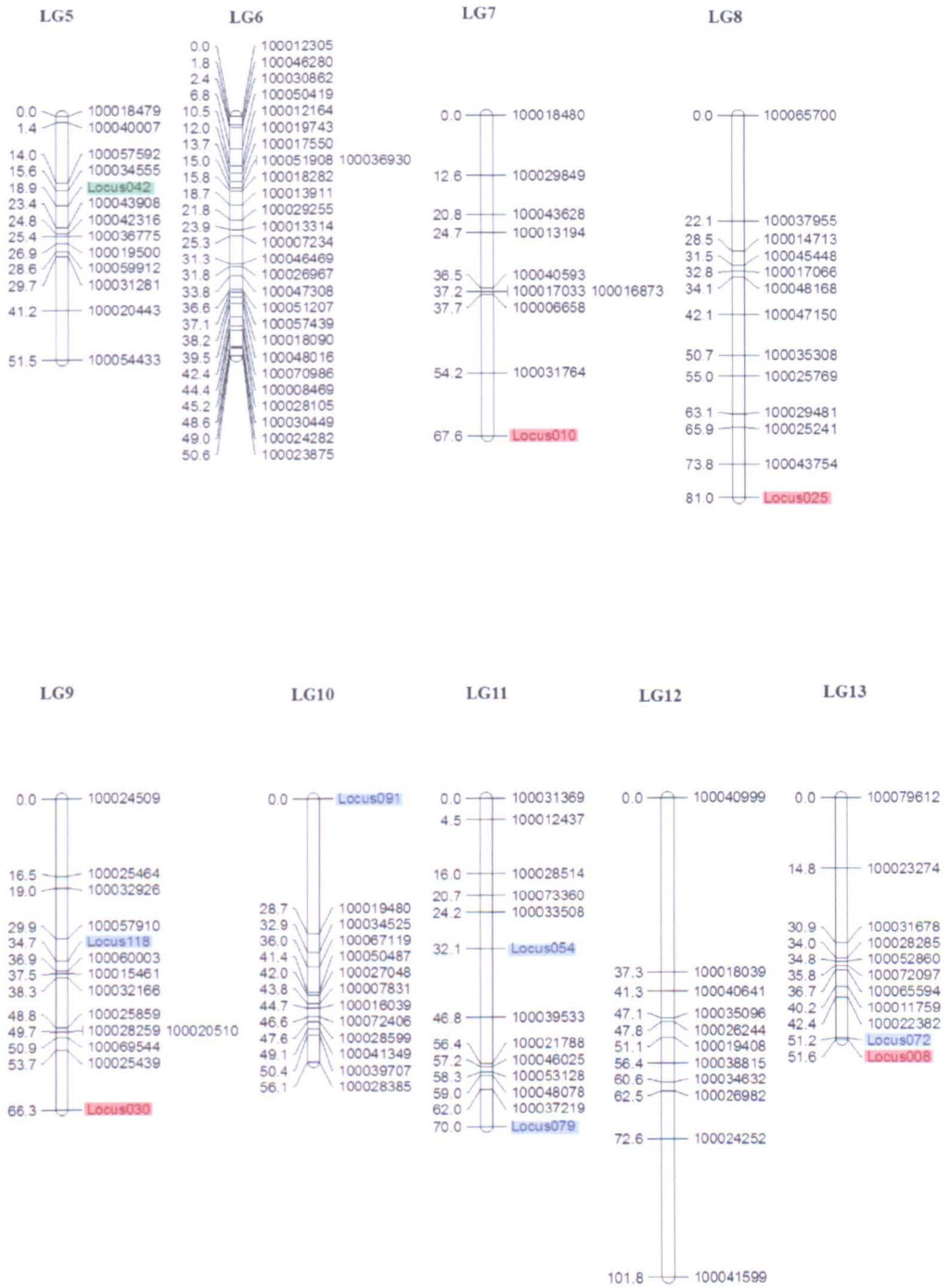


Figure 6.1 (Continued)

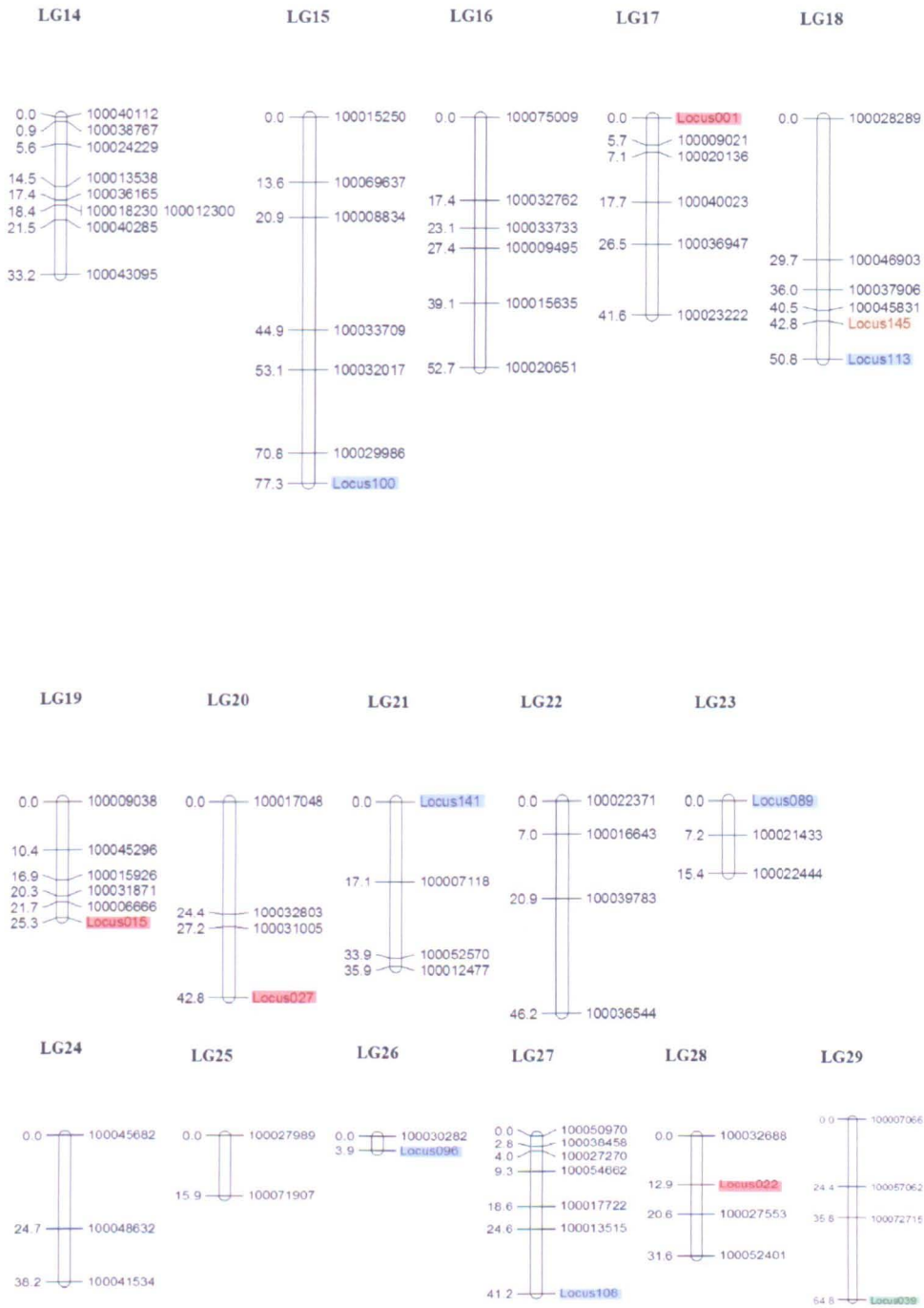


Figure 6.1 (Continued)

The F_1 population map was constructed using SSR and SNP markers. Out of 641 markers, 591 (21 SSR and 570 SNP; 92% mapped) could be assigned to 30 linkage groups (LG1-LG30; Figure 6.2), which had between 2-56 markers and a linkage group length varying from 8.5 cM (LG26) to 156.9 cM (LG12). The SSRs marker number varied from 0 to 2 markers per linkage group, while the SNP marker number varied from 2 to 54 markers per linkage group (Table 6.7). The maps of linkage groups 1, 2, 3, 4, 5, 6, 7, 8, 9, 11, 12, 13, 14, 15 and 27 required more than a single round of mapping (using the regression algorithm). The round one (Map1 version) was used for all groups in the reported map, due to an above threshold value for *jump* in the second (Map2) and the third maps (Map3) indicating a poor fit of the added markers. This is with the exception of groups 5 and 11 where Map2 versions were selected as the markers placed during the second round did not exceed the *jump* limits

The linkage map spans a total genetic distance of 2,385.6 cM with 50 markers remaining unlinked (7.8%). Markers were randomly distributed on the linkage groups. The average marker distance was 4.95 cM, with intervals between markers ranging from 1.6 to 11.9 cM (Table 6.7). The sex determination locus PDK_30s101A (locus 145) was mapped in linkage group 'LG29' which was located at 4.9 cM.

Table 6-7: Distribution of the markers, linkage group length and marker density in the genetic map constructed in the F_1 population.

Linkage groups	Length (cM)	No. of markers mapped in the groups			Average marker interval (cM)
		Total marker	SSR marker	SNP marker	
1	139.8	56	2	54	2.50
2	99.6	49	2	47	2.03
3	70.4	44	2	42	1.60
4	88.2	38	1	37	2.32
5(Map2)	96.5	31	2	29	3.11
6	89.3	28	1	27	3.19
7	139.5	29	1	28	4.81
8	78.9	29	3	26	2.72
9	140.9	26	0	26	5.42
10	143.7	20	0	20	7.19
11(Map2)	68.3	20	0	20	3.42
12	156.9	18	1	17	8.72
13	75.6	16	0	16	4.73
14	39.8	13	1	12	3.06
15	49.1	13	1	12	3.78
16	31.3	13	1	12	2.41
17	106.1	13	0	13	8.16
18	47.8	12	1	11	3.98
19	96.5	11	0	11	8.77
20	68	10	0	10	6.80
21	60.8	10	0	10	6.08
22	94.3	8	0	8	11.79
23	40.7	7	0	7	5.81
24	19.4	7	0	7	2.77
25	31.1	7	0	7	4.44
26	8.5	2	0	2	4.25
27	155	39	1	38	3.97
28	77	11	0	11	7.00
29	42.2	7	1	6	6.03
30	30.4	4	0	4	7.60
Total	2385.6	591	21	570	-
Range	8.5-156.9	2-56	0-2	2-54	1.6-11.9

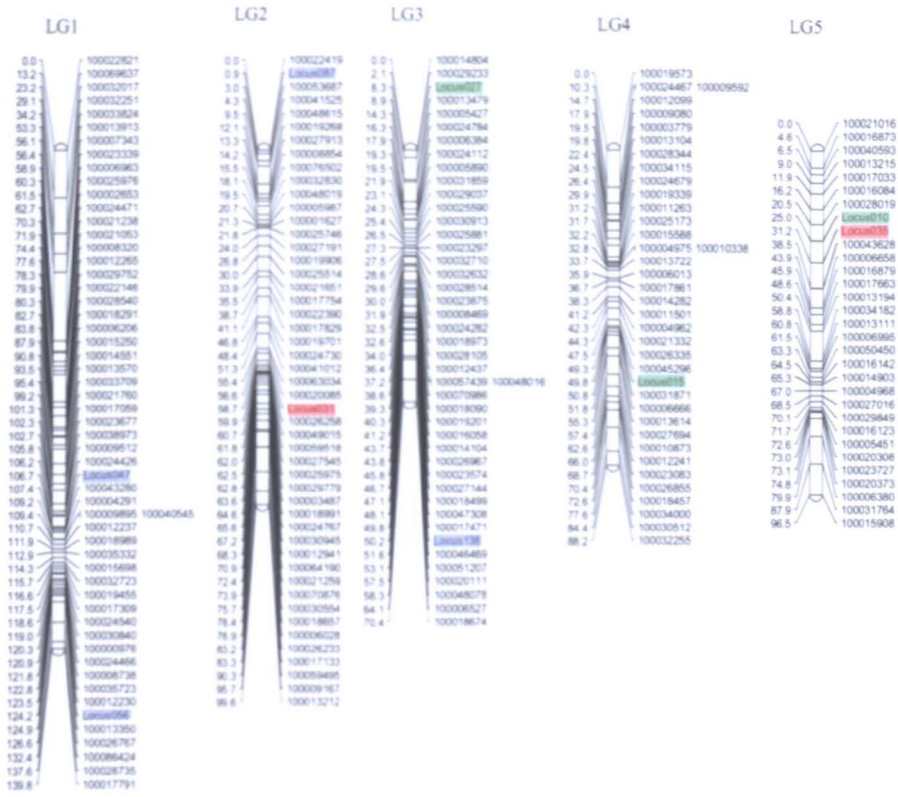


Figure 6.2: A genetic map of the 30 linkage groups. This was constructed in 30 F_1 individuals derived from the cross between Um-Assela and KI-96-13. The locations of 21 SSR and 570 SNPs markers are given. Positions are given in centimorgans (Kosambi units) to the left of the linkage groups and the name of markers to the right. The SSRs markers are color coded by their source (*green*; Billotte *et al.* (2004) and Akkak *et al.* (2009), *blue*; Hamwieh *et al.* (2010) primers tested in this study, *red*; SSRs developed in this study, *orange*; sex determination locus). All SNPs markers are in *black*.

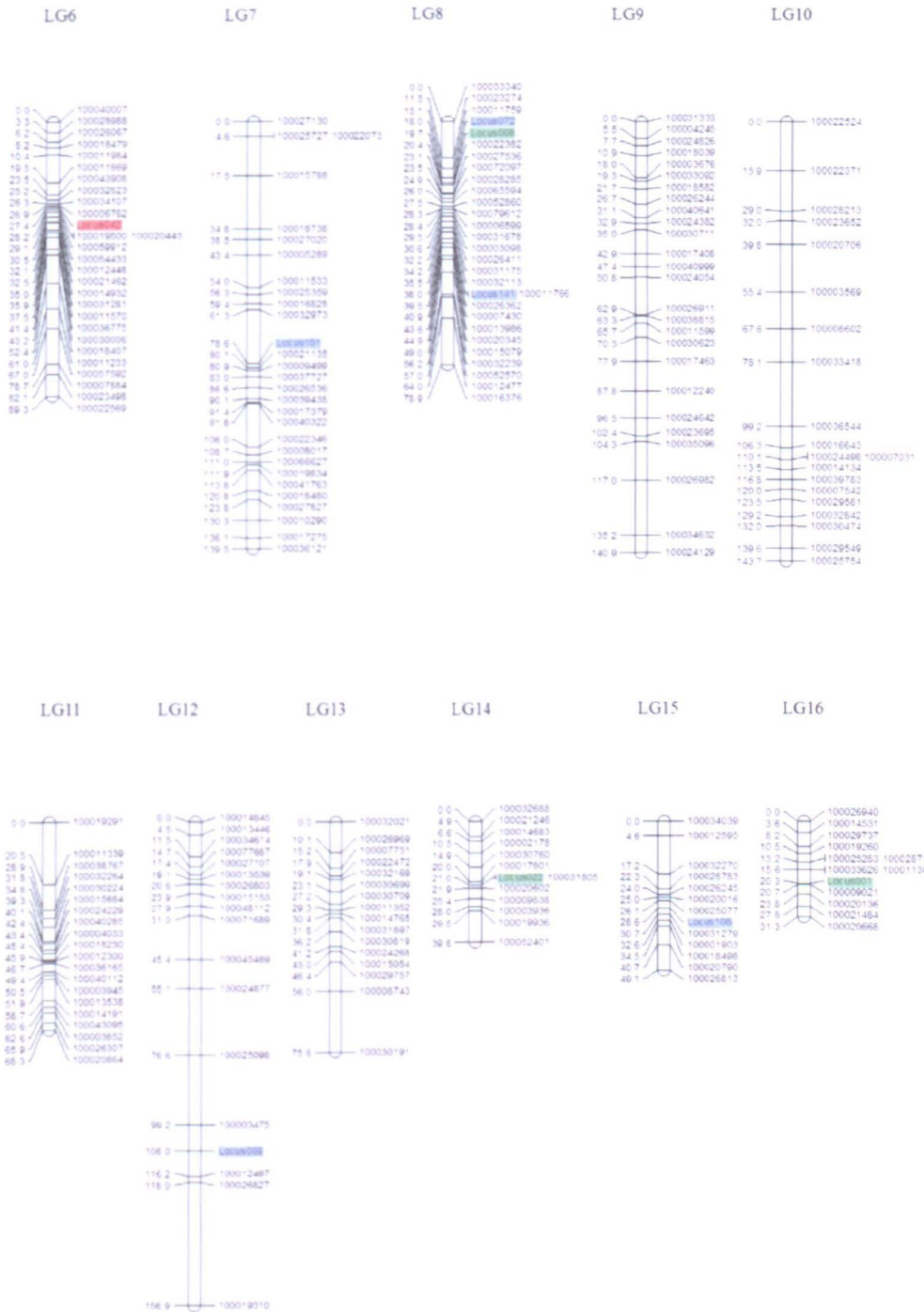


Figure 6.2 (Continued)

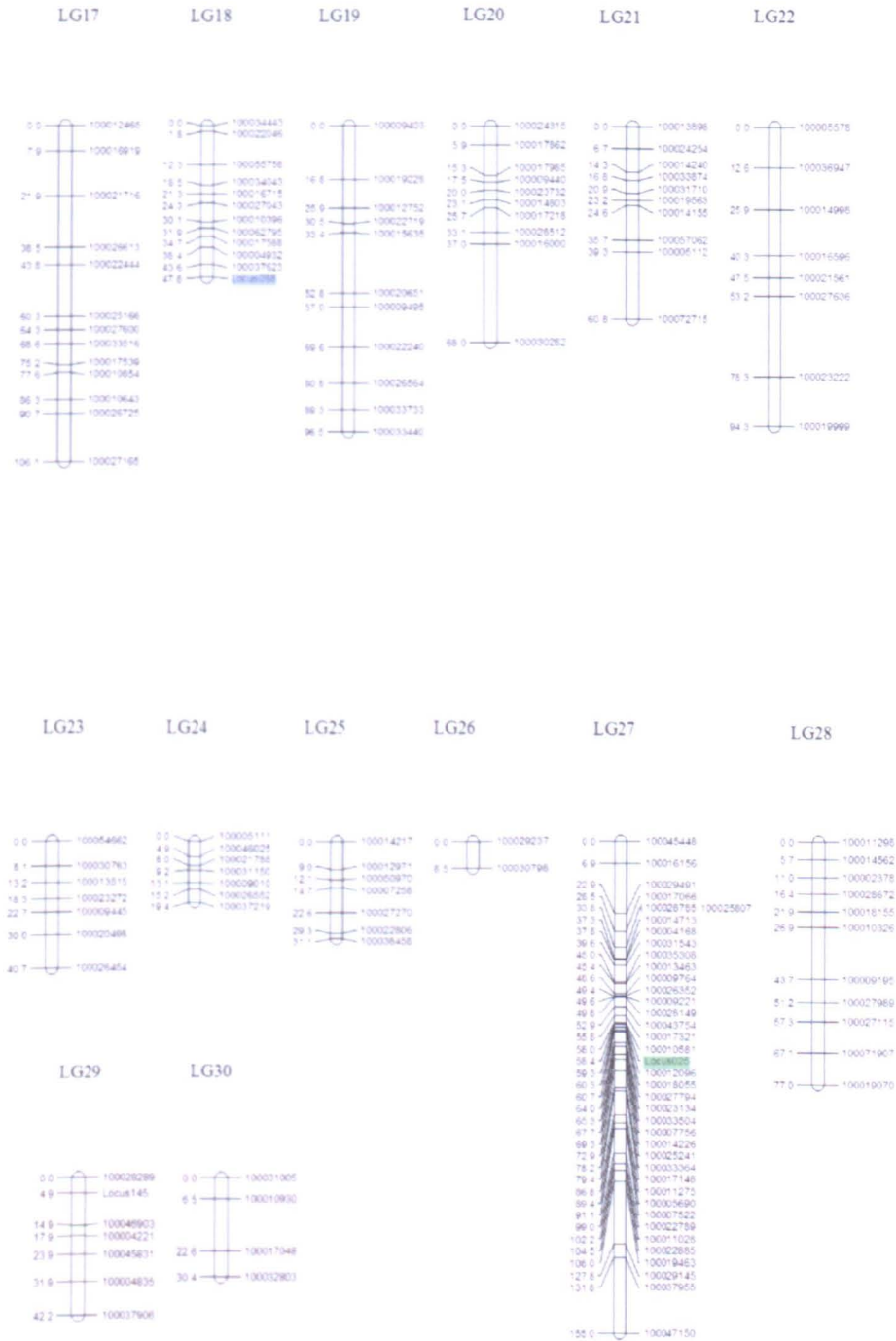


Figure 6.2 (Continued)

6.3.4 Combining the 'BC_1' and 'F_1' population maps

JoinMap4.1 was used to combine linkage groups from both maps. The "Combine Maps" command was used to align genetic maps obtained in different populations for a visual inspection of the marker order. The 'BC_1' and 'F_1' crosses carried a total of 15 common SSR and 142 SNPs markers, giving a total of 157 common markers (Table 6.8). These were used to combine the linkage groups from both populations, where two or more common markers existed, giving a total of 25 combined linkage groups. The JoinMap4 function 'combine groups for map integration' was applied for the pairwise data, followed by regression mapping under default conditions. Two sets of linkage groups could only be linked through the existence of one common marker each (Figure 6.4). While it was possible to confirm common groups through one common marker on the same chromosome, their relative orientations were not determined.

Table 6-8: Number and type of common markers in both 'BC_1' and 'F_1' cross population used to combine linkage groups.

Linkage group	Linkage groups		Common marker		
	BC_1	F_1	SSR	SNP	Total
combined group-1	1	1	2	10	12
combined group-2	2	7	0	9	9
combined group-3	2	18	1	4	5
combined group-4	3	1	0	1	1
combined group-5	3	12	1	7	8
combined group-6	4	2	1	14	15
combined group-7	5	6	1	10	11
combined group-8	6	3	0	12	12
combined group-9	7	5	1	8	9
combined group-10	7	7	0	1	1
combined group-11	8	27	1	8	9
combined group-12	11	24	0	3	3
combined group-13	12	9	0	8	8
combined group-14	13	8	2	9	11
combined group-15	14	11	0	9	9
combined group-16	15	1	0	4	4
combined group-17	16	19	0	4	4
combined group-18	17	16	1	2	3
combined group-19	18	29	1	4	5
combined group-20	19	4	1	3	4
combined group-21	21	8	1	2	3
combined group-22	22	10	0	4	4
combined group-23	27	23	0	2	2
combined group-24	28	14	1	2	3
combined group-25	29	21	0	2	2
non grouped 1	20	3	1	0	1
non grouped 2	27	15	1	0	1

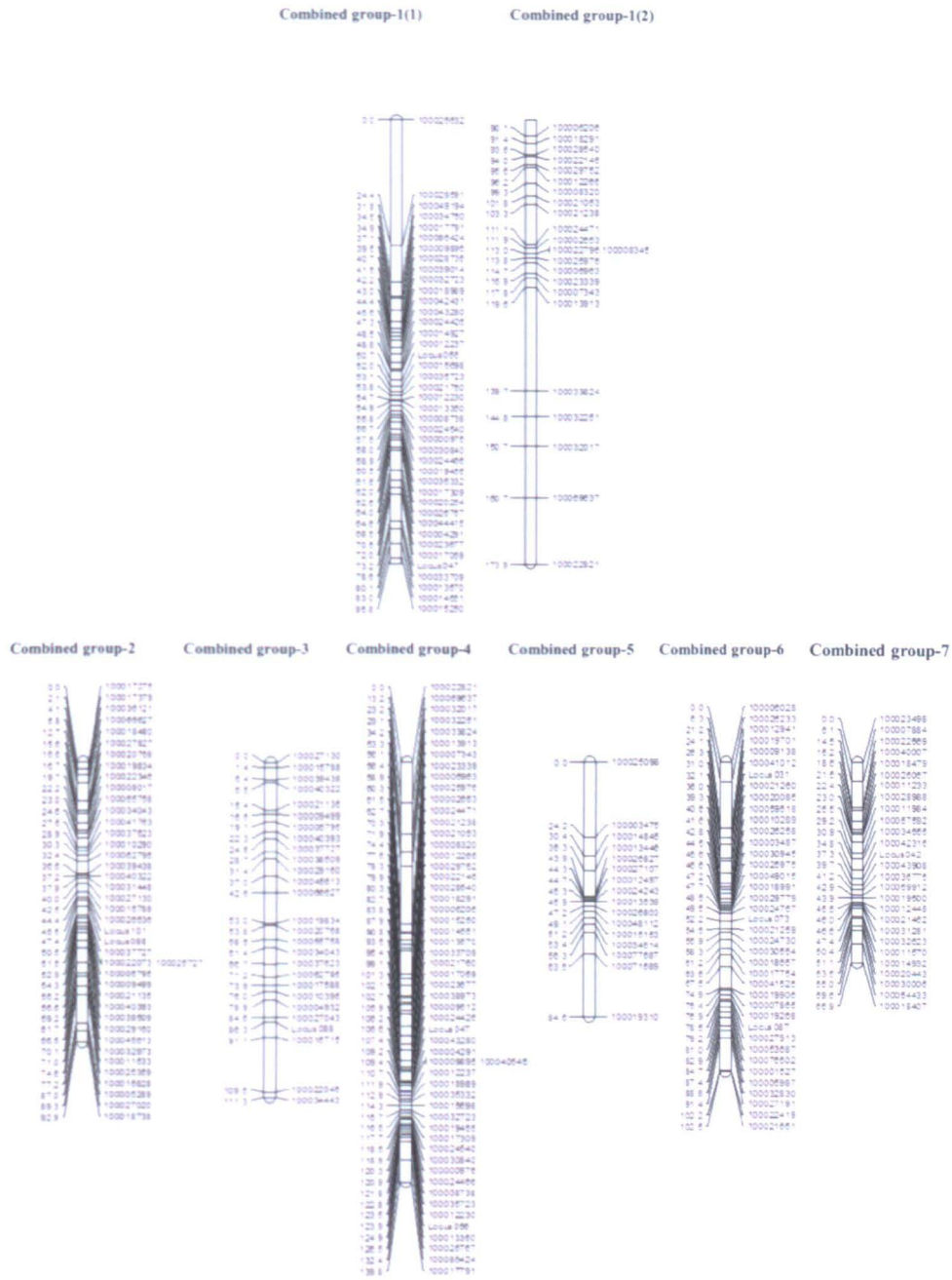


Figure 6.3: Combined linkage groups (combined group-1 to combined group-25) from 'BC_1' and 'F_1' maps where one or more common marker exists and the 'combined' linkage groups where a single common marker exists and relative orientation cannot be determined. Positions are given in centimorgans (Kosambi units) to the left of the linkage groups and the name of markers to the right.

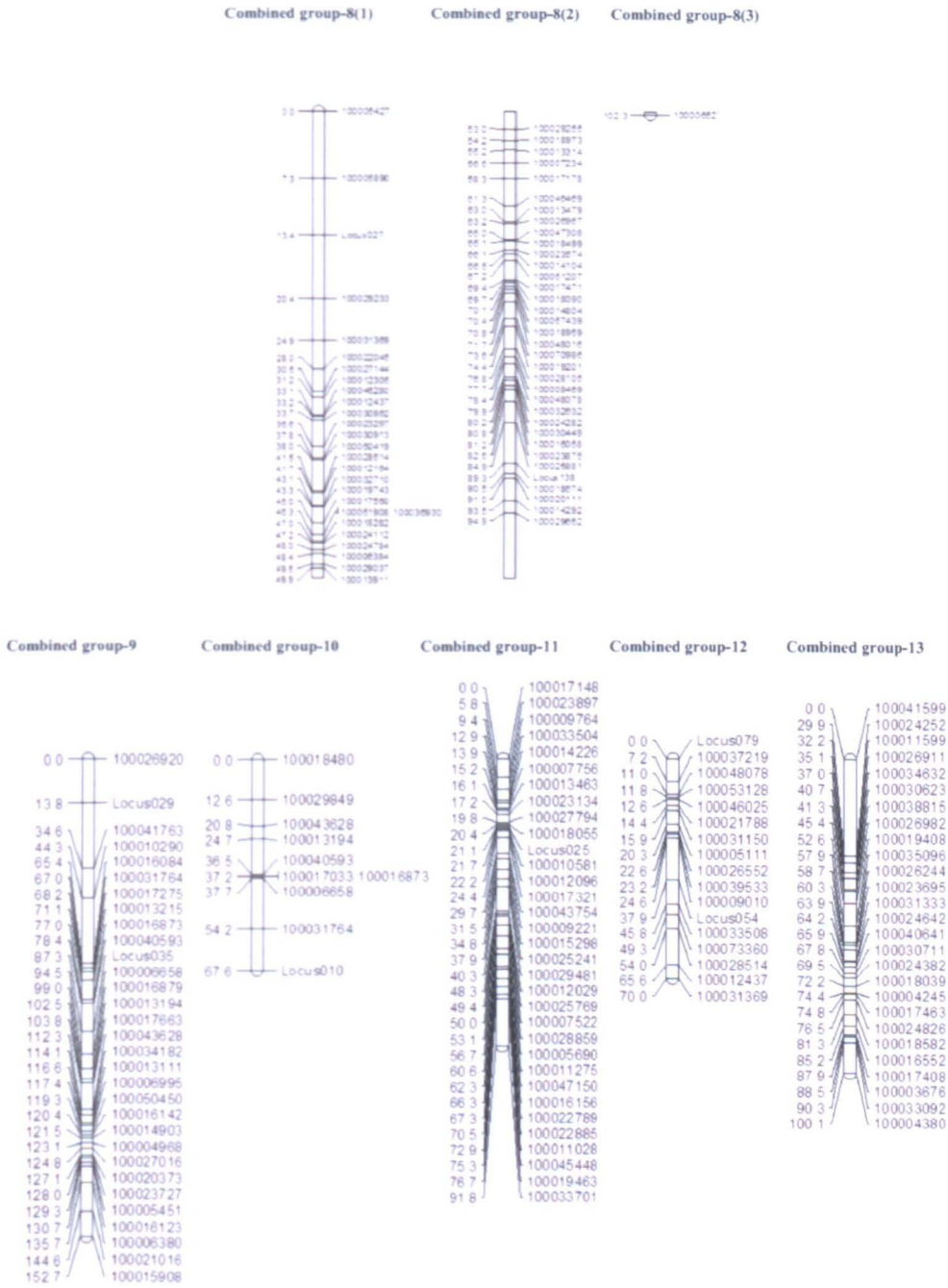


Figure 6.3 (Continued)

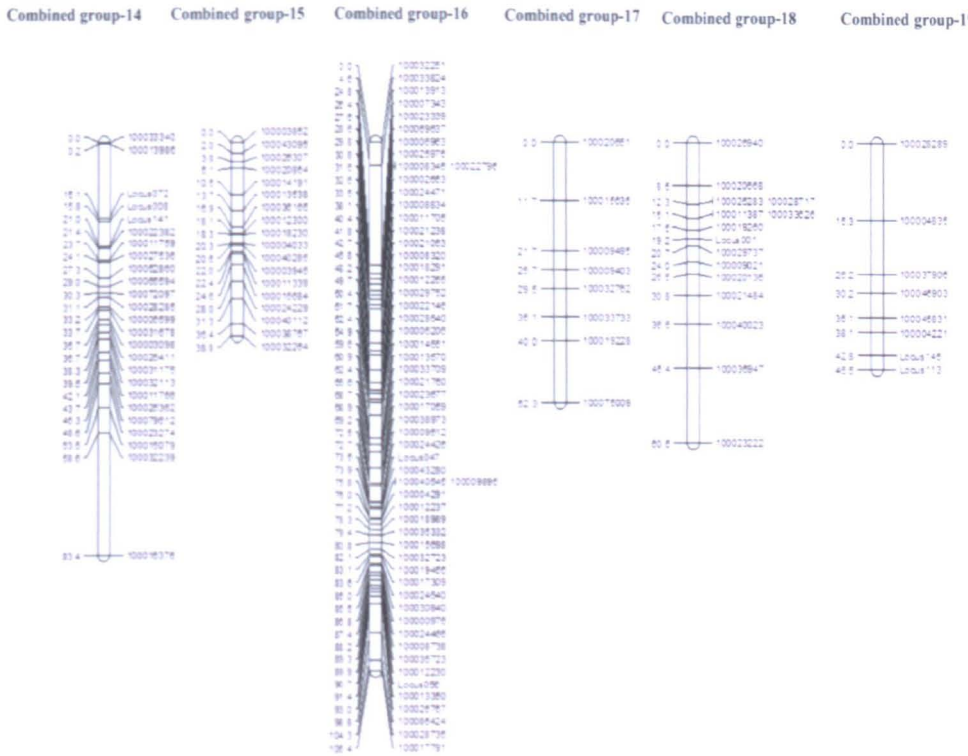


Figure 6.3 (Continued)

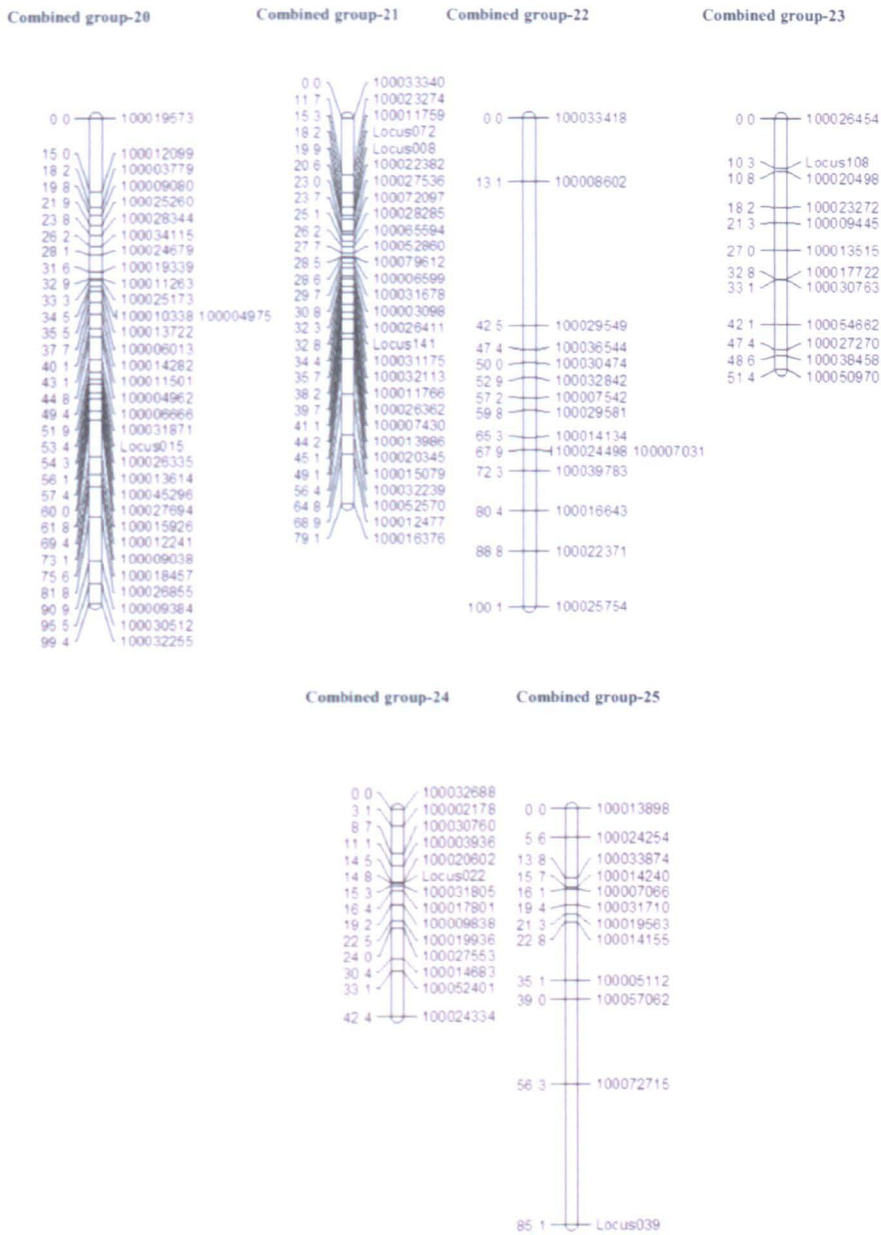


Figure 6.3 (Continued)

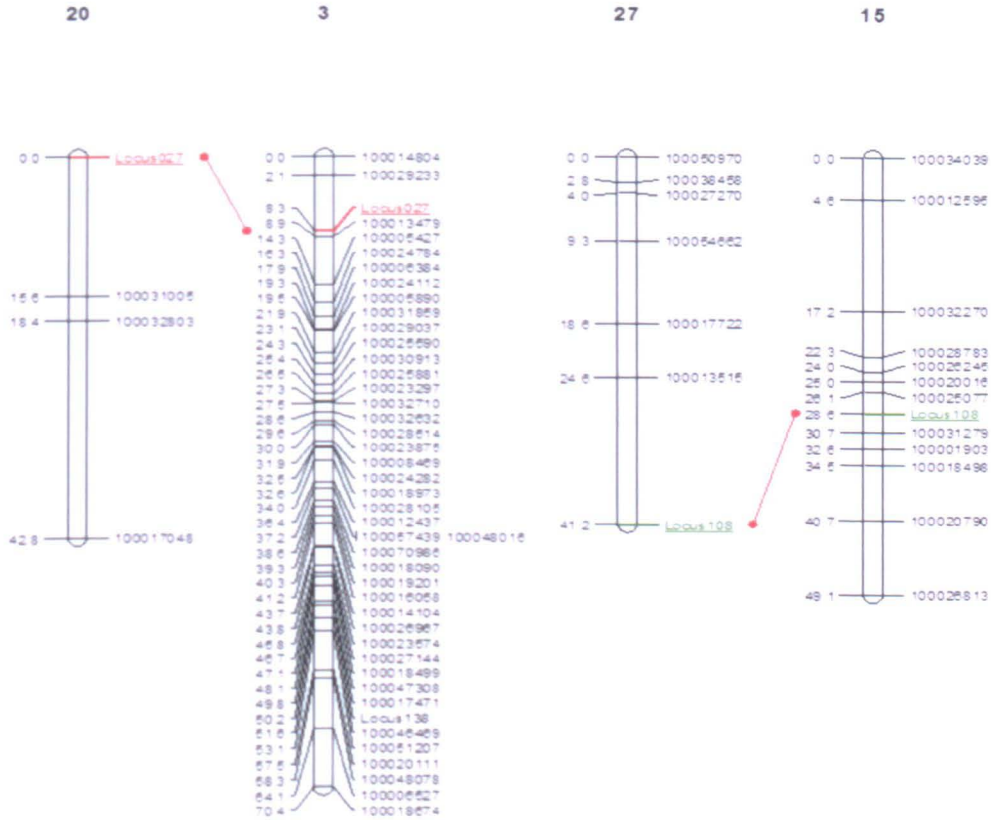


Figure 6.3 (Continued)

6.4 Discussion

6.4.1 Population size

Inheritance in date palm is poorly understood due to the absence of enough segregating populations with sufficient time-depth following their establishment to allow a detailed assessment. No physical or linkage maps have yet been constructed for this crop (Jain *et al.*, 2011).

The main problem with date palm genome mapping is the construction of an appropriate mapping population because of the long time required and substantial space, given the long period to reach maturity and the physically large size of the crop. A large mapping population size should be available so that it includes enough genetic information from many segregating gametes to produce a reliable map. Therefore, it will definitely take a long time to obtain a segregating population of date palm with a reasonable size with trait data. The only available mapping populations for date palm are a BC_1 (KI-96-13 was crossed with its mother Khalas-4) comprising 53 individuals, and an 'F_1'(a progeny from a cross between two heterozygous parents; KI-96-13 and Um-Assela cultivar palm) comprising 30 individuals which were also used in this study. In date palm generation of a classical F_2 is impractical because homozygous lines are not available as each date palm is highly heterozygous and could itself be considered equivalent to an F_1. In addition, as date palm is dioecious, the option often followed in oil palm of self-pollination is not possible. The small population sizes represent a potential problem for development of a comprehensive map in date palm, with the BC_1 (n=53) reasonable, but still below the desired numbers of palms (Young 1994). However, a total of 52 genotypes (a defined F_1) descending from a cross

between a Malayan Yellow Dwarf genotype (MYD20) and a Laguna Tall genotype (LAGT07) was sufficient to generate a linkage genetic map with 16 independent linkage groups in coconut (*Cocos nucifera* L.; $2n = 2x = 32$) (Herrán *et al.*, 2000).

The mapping in out-breeding heterozygous perennial crops is more complicated due to the absence of complete homozygosity in the parents with most of these species not tolerating inbreeding (Semagn *et al.*, 2006).

6.4.2 Segregation distortion

Segregation distortion has been reported in many plants including rice, wheat, barley, and maize. Different explanations have been reported for distortion of segregation ratios in plants, including: competition among gametes, chromosome loss, the presence of viability genes (Zamir and Tadmor 1986; Kasha and Kao 1970; Bradshaw and Stettler 1994), sampling and scoring errors (Faris *et al.*, 1998; Echt and Nelson 1997, Nikaido *et al.* 1999) as well as small population sizes (Millan *et al.*, 2010).

The proportions of distorted markers found in this study were 37.9% in the BC_1 and 37.1% in the F_1, which is higher than 10.6% found in coconut (*Cocos nucifera* L.) (Bandaranayake and Kearsey, 2005), however, they were lower than the 73% found in an interspecific cross of tomato (Paran *et al.*, 1995). The ratios of distortion found in this study are consistent with the reports from other plants like *Medicago tornata* (40.6%; Janczewski *et al.*, 1997), and *Coffea sp.* (30%; Ky *et al.*, 2000). Sampling, scoring errors and small population size may have contributed to these apparent levels of distortion found in the BC_1 and F_1.

6.4.3 Map construction

In this study, construction of the first genetic map (based on SSR and SNP markers) in date palm (*Phoenix dactylifera* L.) is reported. The maps presented here will provide a genetic framework for mapping the qualitative and quantitative traits in this crop.

There is no obvious clustering of markers in both BC_1 and F_1 maps. In the BC_1 map, the number of markers ranged from 2 to 27 markers per group, while 2 to 56 markers per group in F_1 map.

The genetic map of the BC_1 population had 29 linkage groups. These 29 linkage groups had between 2 and 27 markers and a linkage group length varying from 3.9 cM to 101.8 cM. While the genetic map of the F_1 population had 30 linkage groups with 2 to 56 markers and a linkage group length varying from 8.5 cM to 156.9 cM. The number of linkage groups in the BC_1 and F_1 maps exceeded the expected number of 18 pairs of chromosomes ($2n = 2x = 36$) in date palm (Jain *et al.*, 2011). However, some of the smaller linkage groups with a few markers are likely to be derived from the same chromosome, suggesting that the excess of linkage groups might be due to incomplete coverage of the genome with this number of markers. Adding more markers could bring the smaller groups together.

A draft sequence of the entire genome of date palm have been recently published, estimating the size to be around 658 Mbp (Al-Dous *et al.*, 2011), which appears to be relatively small in comparison to other monocotyledons and perennial species such as oil palm and coconut with genome sizes of 1,800 Mb and 2,150 Mb, respectively (Feuillet *et al.*, 2011). In fact plant genome varies significantly in DNA content between species. 'C-value Paradox' has been used as a term for difference in genome size between such similar

species (Mayes *et al.*, 2008). Most of the differences in genome size between such species are highly likely to be due to repetitive DNA, mainly a class of DNA element termed as retrotransposons. Retrotransposons have the ability to make an RNA copy of themselves, convert it into DNA, and insert the new copy into the plant genome in a different place, therefore increasing the genome size. However, some techniques including: molecular genetics, mapping, and QTL analysis are not significantly affected by genome size, in which the genetic distance depends on the number of DNA crossovers occurs during meiosis which is likely to be dependent on chromosome number (with at least one cross-over needed per chromosome to ensure correct disjunction during meiosis) but less dependent on genome size (Mayes *et al.*, 2008).

The BC_1 map had a genetic distance of 1,486.7 cM, while the F_1 map had 2,385.6 cM, with the latter longer than the genetic maps reported for oil palm (1,743 cM) and coconut (1,971 cM). The observed differences could be explained by population size, population type, number of markers, missing data and scoring errors.

The total length of the linkage maps differ for the BC_1 and F_1 populations. Relative to the F_1 map, the BC_1 map was approximately 62% shorter, which could partly reflect the number of recombination events being observed. In theory, in a BC_1 developed between two inbred lines there is only detectable recombination in the F_1 parent as the recurrent parent is homozygous, so only a single recombination event is observed per plant. In the F_1 cross; there will be observable recombination events in both parents.

Genetic maps with different total lengths have been reported for many crops using different types and sizes of mapping population. This is true even if

maps are generated for different populations of the same species and can partly be explained by variation in the level of recombination that occurs in different crossings (Semagn *et al.*, 2006). For example, Herrán *et al.* (2000) compared parental linkage maps in coconut (*Cocos nucifera* L.) in an F_1 populations derived from a cross between Malayan Yellow Dwarf (MYD) and Laguna Tall (LAG). The total length of Malayan Yellow Dwarf (MYD) map at 1,266 cM was 58.9% shorter than Laguna Tall (LAG) map 2,226 cM.

The sex determination locus PDK_30s101A (locus 145) developed in this study was also localized in both BC_1 and F_1 maps at 42.8 cM and 4.9 cM in linkage groups 18 and 29, respectively. This marker has been widely tested on date palm samples from Oman and in different genetic origins and was found to predict a high level of discrimination between male and female date palms among multiple varieties distributed across a wide range of cultivation, with an accuracy of 100% in the crosses, 96% in the Omani material and 86% in the broadest date palm germplasm (Chapter 7). Sex differentiation has been studied in a number of plant species such as papaya, *Silene latifolia*, melon and grapevine (Ming *et al.*, 2007; Farbos *et al.*, 1999; Boualem *et al.*, 2008; Martin *et al.*, 2009; Marguerit *et al.*, 2009). Such markers for sex determination in crops like date palm are very important. The sex-determination marker can help breeders and producers of date palm identify and eliminate non-productive male trees in the nursery before planting on a field scale, where it is known that male flowers from a single tree can be used to pollinate 40–50 female date palms (Jain *et al.*, 2011).

6.4.4 Combined linkage maps

The use of common markers to link genetic maps to one another has been applied over the years to generate a number of genetic-to-genetic map linkages. If markers are located as common markers on two maps, the comparative locations can be easily determined by extrapolating from the normalized distance between the common markers on the two maps (Cone and Cone, 2009).

The 'BC_1' and 'F_1' maps were combined based on common SSR and SNP markers present in both maps to form the final genetic map in date palm. Combining linkage groups with two to 15 common markers was possible for 25 groups, however due to the lack of a physical map for date palm the orientation of other groups with single linkages could not be determined.

Out of a total of 270 SSR and SNP markers mapped in BC_1, 157 (58%) markers were common, which is actually a high percentage and helps to validate the linking of the BC_1 and F_1 maps.

These common markers are useful to help bridge individual linkage maps within one species. Although the total number of markers used in this study is still limited, however, we could increase the number of markers in future, which will help to increase the accuracy of the estimates of linkage between markers, especially for the combined map.

6.5 Conclusion

- This is the first report of the construction of a medium density genetic map in date palm. The BC_1 population allowed the construction of a linkage map with total genetic length of 1,486.7 cM, consisting of 270

markers (28 SSR and 242 SNP) distributed into 29 linkage groups. While the F_1 population allowed the construction of a linkage map with total genetic length of 2,385.6 cM, consisting of 591 markers (21 SSR and 570 SNP) distributed into 29 linkage groups.

- This study showed that SNPs generated through the DArT Seq approach are a useful molecular marker technique to obtain many polymorphic loci to allow development of a framework map for a crop like date palm, as the number of currently available SSR markers is not adequate to allow map construction.
- It was possible to map the sex locus PDK_30s101A (coded as locus 145) in both BC_1 and F_1 maps at 42.8 cM and 4.9 cM in linkage groups 18 and 29, respectively and on combined group 19 at 42.8cM, which represents an important step towards validating this molecular marker tightly linked to the gene controlling the determination of sex in this dioecious crop.
- The presence of common markers in BC_1 and F_1 made it possible to identify a total of 25 combined linkage groups, in this way the number of markers available per linkage group can be increased and an attempt to refine the numbers of groups to closer the chromosomal number begun.

Chapter 7. DEVELOPMENT OF NEW MICROSATELLITE PRIMERS (SSRs) FOR GENDER DISCRIMINATION IN DATE PALM *(Phoenix dactylifera* L.)

7.1 Introduction

Sex determination is an important developmental process in the life cycle of many sexually reproducing plants. It is defined as the physical process that leads to separation of gamete-producing structures in both males and females of different plant species (Tanurdzic and Banks, 2004). The bisexual pistillate and staminate flowers either develop in separate individuals (dioecious species; e.g. papaya and date palm) or on the same individual (monoecious species; e.g. maize and oil palm) (Charlesworth, 2002).

Sex differentiation has been investigated in a number of plant species including papaya (Ming *et al.*, 2007), *Silene latifolia* (Farbos *et al.*, 1999), melon (Boualem *et al.*, 2008; Martins *et al.*, 2009) and grapevine (Marguerit *et al.*, 2009). This has been 'fruitful' in identifying the factors that could be involved in determining the sex of the flower or individual.

Date palm (*Phoenix dactylifera* L.) is a diploid with 18 pairs of chromosomes having a genome size of 658 Mb as documented recently by Al-Dous *et al.*, (2011). It is a dioecious plant, with male and female flowers on separate trees (Elmeer and Mattat, 2012; Al-Khalifah and Askari, 2007; Ainsworth *et al.*, 1998; Kgazal *et al.*, 1990). The structure of the staminate and pistillate flowers in date palm was first described by De Mason and Tisserat (1980) and De Mason *et al.* (1982). Masmoudi *et al.* (2008) have further reported identification of eight developmental stages in female date palm flowers. More recently, Daher *et al.* (2010) have provided a detailed analysis of inflorescence

and flower development in date palm. The inflorescences of date palm are produced in the axils of their subtending leaves. In general, both male and female plants have the same inflorescence structures and axis organization (Daher *et al.*, 2010). A single prophyll encloses the whole inflorescence and each inflorescence contains a rachis (main axis) bearing numbers of rachillae where the individual flowers occur (Figure 7.1).

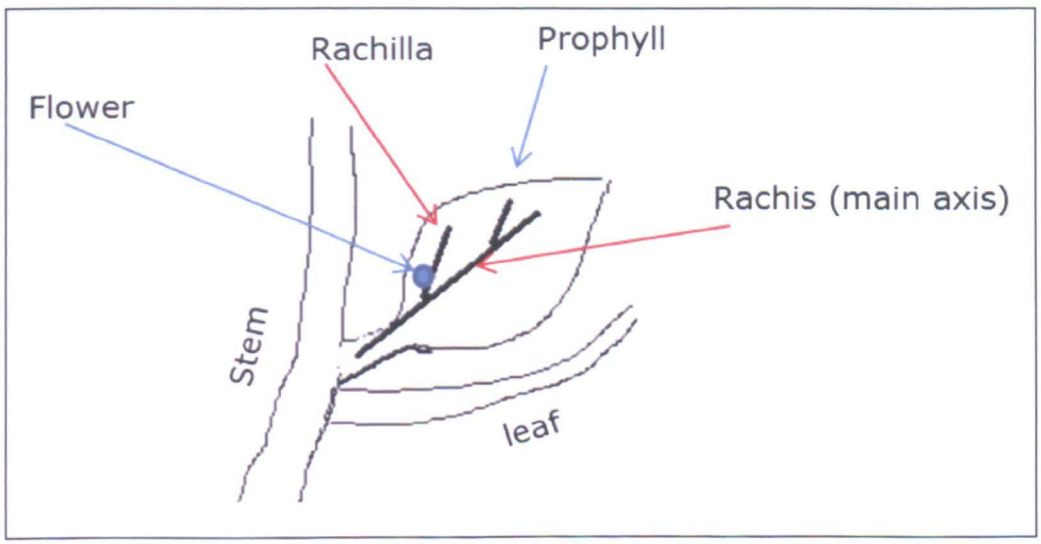


Figure 7.1: Structure of inflorescences in date palm; representing the rachis (main axis), and rachilla where individual flowers occur.

At anthesis, the inflorescences of male and female palms show differences in shape and size. Female inflorescences have large, elongated peduncles while the male have shorter peduncles (Figure 7.2). The rachis of female plants has fewer branches or rachillae and each rachilla bears about 40 solitary flowers. In contrast, the rachis of male plants has many shorter rachillae, each bearing about 50 solitary flowers (Daher *et al.*, 2010).

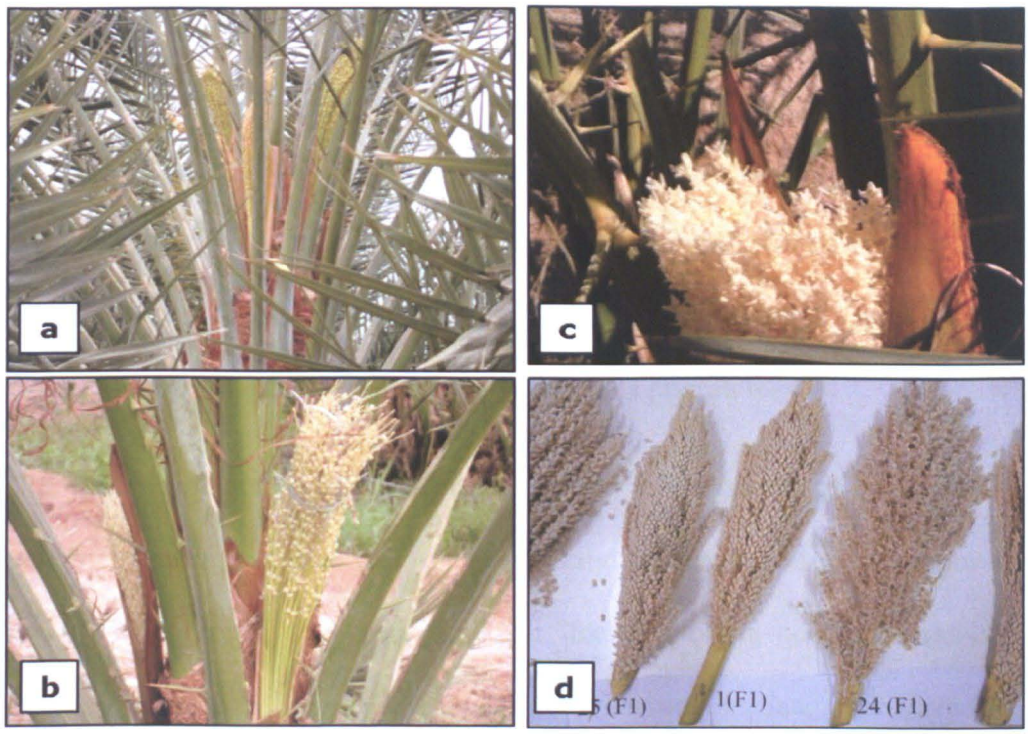


Figure 7.2: (a) and (b) Female inflorescences, (c) and (d) Male inflorescences of date palm.

Furthermore, Bekheet & Hanafy (2011) have reported that male and female progeny palms are produced in a 1:1 ratio. It is almost impossible to distinguish between male and female palms until the onset of fruiting, which normally takes 5 to 7 years and by the time farmers are able to differentiate between male and female palms they have invested significant water, fertilizer and planting space for male palms that will not give any return on this investment. This is one of the major reasons that date palms are not commercially propagated through seed and also is a serious constraint on the development of breeding programs.

A number of attempts have been made to identify sex-specific DNA markers for date palm cultivars using different molecular techniques including; RAPD and ISSR (Ahmed *et al.*, 2006; Younis *et al.*, 2008), and SSRs (Elmeer and

Mattat, 2012). Al-Dous *et al.* (2011) were able to identify a region strongly linked to sex determination in date palm. Their investigation revealed that the date palm has an XX XY system with the male being the heterogametic sex. Al- Dous *et al.* (2011) also suggested that further studies should be conducted on this region to identify a specific mutation or any other gene content difference that would determine palm gender in a precise way.

The specific objective of this study was to benefit from the published draft genome sequence of date palm in general and gender-linked region reported by Al- Dous *et al.* (2011) in particular, for developing new microsatellite markers that would help in early sex determination in date palm breeding programme.

7.2 Plant materials

The study involved 290 date palm accessions representing 151 female and 43 male trees collected from the National Germplasm Collection at Wadi Quriat Research Station, Buhla, Oman. It also included 90 palms from the BC_1 and F_1 populations (Figure 7.3; Appendix 7). The BC_1 and F_1 populations were developed in Oman in 1996 and planted at the Date Palm Research Station at Wadi Quriat and at the Research Farm at Barka, respectively (El Kharbotly *et al.*, 2006) and were used to construct a genetic map (Chapter 6).

In addition, DNA samples from Sanremo, Bordighera, USDA-ARS and France were used in this study together with date palm samples of various origins including Iraq, Libya, Sudan and Iran, giving a total of 96 additional samples (Figure 7.3; Appendix 8).

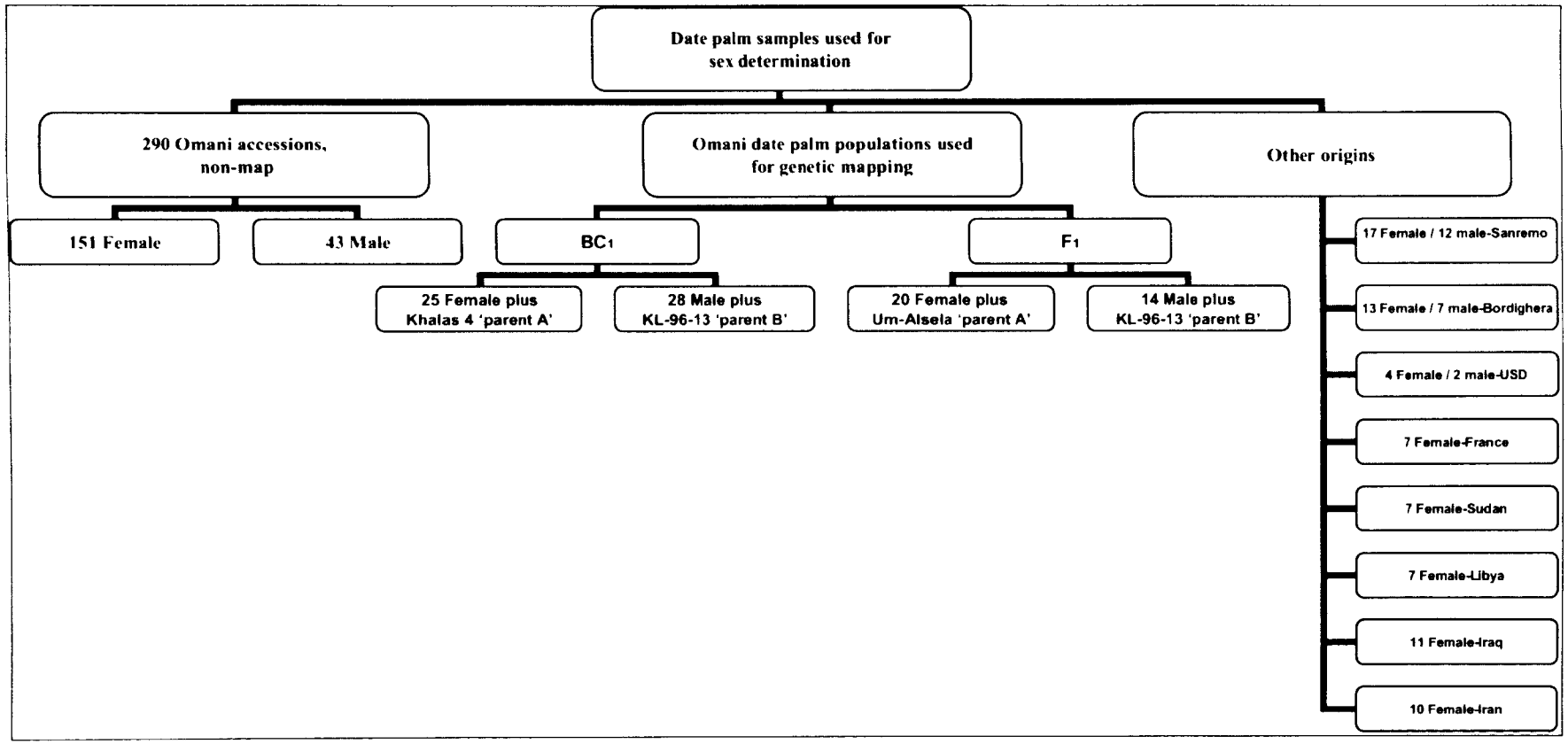


Figure 7.3: A flow diagram illustrating date palm samples used for sex determination.

7.3 Results

Sequences for five microsatellite primer pairs (PDK_30s101A, PDK_30s131, PDK_30s231, PDK_30s101B, PDK_30s101C; Table 7.1) were obtained from the sequences of three scaffolds showing SNP segregation associated with gender in date palm reported by Al-Dous *et al.* (2011; Figure 7.4). The annealing temperature for the five primer pairs was determined to be 60°C (Figure 7.5) and the expected allele sizes ranged between 211 bp and 367 bp.

Table 7-1: List of five new microsatellite primers, their sequences and M13-extention (in parentheses), motif repeat, annealing temperature and expected allele size in date palm (*Phoenix dactylifera* L.).

Oligo's Lab Code	Oligo Name	Sequences (5'-3')	Motif repeat	Annealing T_m (°C)	Expected size (bp)
1	PDK_30s101AF	(CACGACGTTGTAAAACGAC)CTCAAGAGAGTACCCCAAGCAT	$(TA)_7$	60	317
	PDK_30s101AR	GGGATAATGTTGTTGCTCCG			
2	PDK_30s131F	(CACGACGTTGTAAAACGAC)TTTGGAGCTACCTTTTCTGTGA	$(A)_{11}$	60	367
	PDK_30s131R	GAAGAATGTGGGGATGGATTT			
3	PDK_30s231F	(CACGACGTTGTAAAACGAC)CTCTCCTCCGTTCCTCCTAGAT	$(A)_{11}$	60	341
	PDK_30s231R	CAGGGTAGATGGGTAAATCCAA			
4	PDK_30s101BF	(CACGACGTTGTAAAACGAC)CTCGGCTTATTTGGTGGAAA	$(TA)_7$	60	278
	PDK_30s101BR	CTTCTCTGGGATAATGTTGTTGCT			
5	PDK_30s101CF	(CACGACGTTGTAAAACGAC)CTCGATATTTTCCTTGGATCAC	$(TA)_7$	60	211
	PDK_30s101CR	AGACTCCTCCTTCACATAGAACAA			

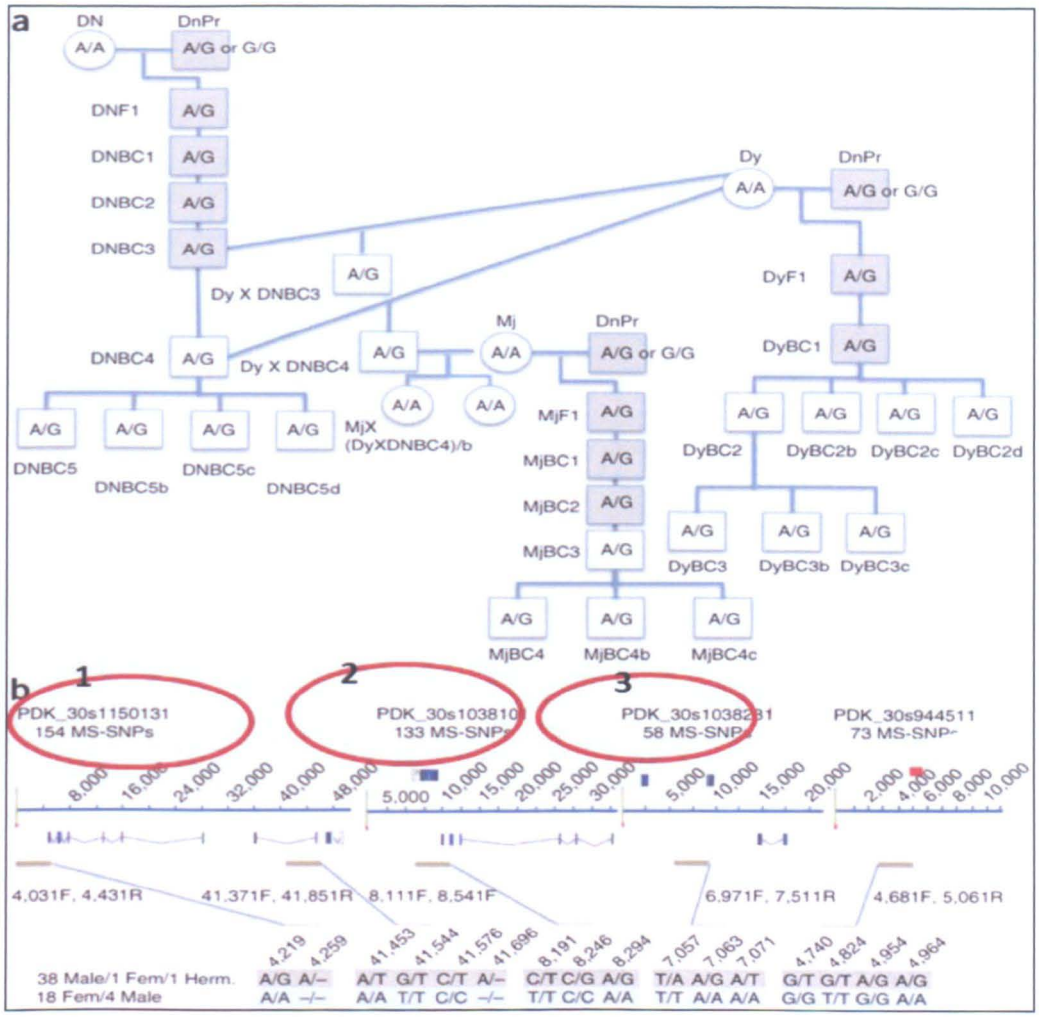


Figure 7.4: Shows the selected three scaffolds (PDK_30s1150131, PDK_30s1038101, and PDK_30s1038231) used for designing SSR primers. This figure was published by Al-Dous *et al.* (2011) and summarizes the pedigree and genotype information for the gender- discriminating regions.

Three SSRs primer pairs PDK_30s101A, PDK_30s131 and PDK_30s231 were first screened and tested on eight backcrossed males and females (1B, 7B, 16B, 35B, 36B, 38B, 43B and 44B; Table 7.1) with the aim of identifying polymorphic co-dominant markers, specific enough to distinguish sex in multiple varieties and backcrossed males and females. The results indicated that all three SSRs primer pairs amplified successfully all samples (Figure 7.6).

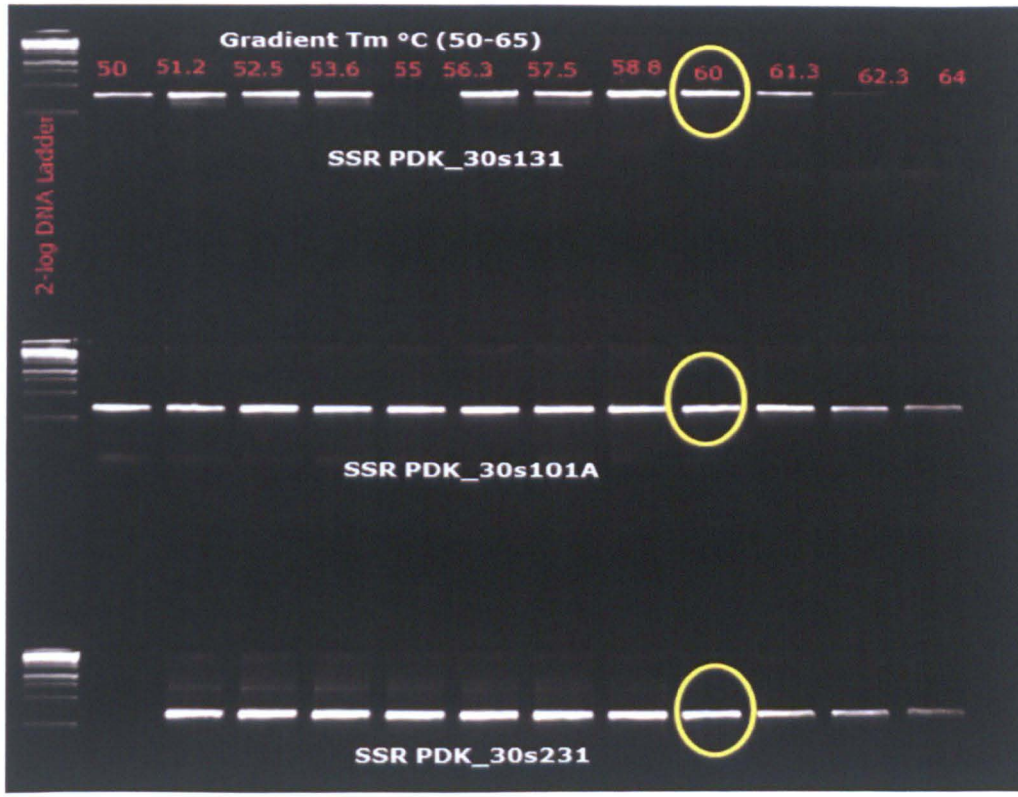


Figure 7.5: Annealing temperature optimisation for the three microsatellites primers PDK_30s131, PDK_30s101A, and PDK_30s231. The selected annealing temperature is indicated with yellow circle. The missing sample for 55oC with SSR PDK_30s131 is likely to be due to be due to a pipetting error.

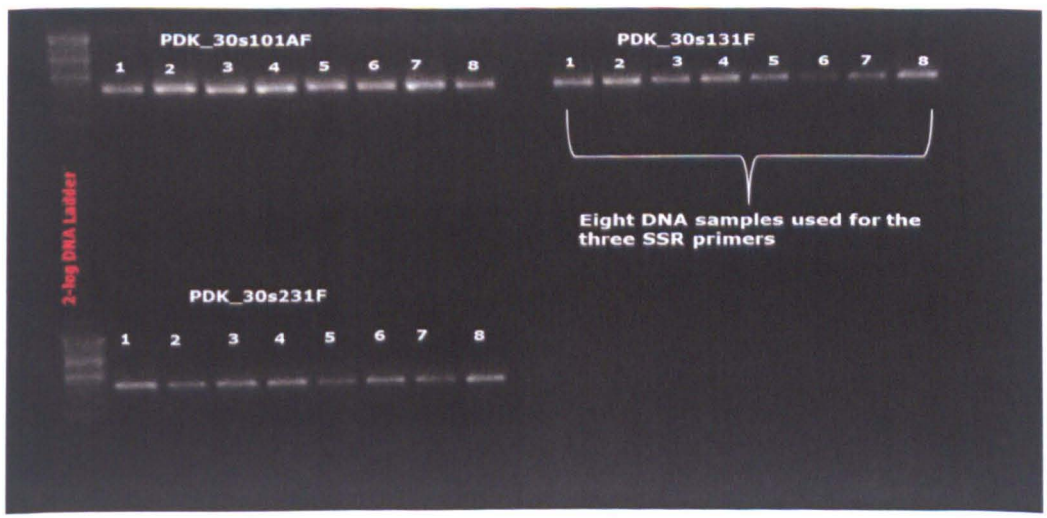


Figure 7.6: PCR products of the three SSR primers (PDK_30s101A, PDK_30s131 and PDK_30s231) on 2% (w/v) agarose gels using eight date palm samples.

After CEQ analysis, two microsatellite loci PDK_30s131 and PDK_30s231 were excluded due to monomorphic bands in all samples with allele sizes of 389 and 359 bp, respectively.

The most interesting locus was PDK_30s101A as it revealed polymorphism amongst samples. The fragment analysis of this locus gave two distinct alleles, one was shared between males and females with allele size 339 bp, while the second allele appeared strictly limited to the male phenotype with size 346 bp (Figure 7.7).

In this case we decided to test this locus with multiple date palm varieties. One hundred and fifty-one female accessions from Oman were amplified using the PDK_30s101A SSR primers and all samples produced a homozygous allele at size 339 bp except for six female accessions that showed both female and male alleles at 339 bp and 346 bp (Table 7.1). These accessions were Qash Hareer (34), Hessas (47), Naghl Lulu (72), Huzaifah (125), Qash Baloobiya (137), and Beliaq (143). Additionally, all forty-three male trees produced two alleles, one allele at 339 bp, while the size for the second allele was 346 bp, as reported earlier. This indicates that the female palm is homozygous with two copies of the same allele, yielding a single band or peak. However, the male has two different alleles and produces two distinct peaks.

The PDK_30s101A SSR primer was also tested on 90 palms from the BC_1 and F_1 populations. The BC_1 population consisted of 28 male and 25 female palms from a cross between the Kl-96-13 male and Khalas 4 female, whereas the F_1 included 14 male and 20 female palms derived from a cross between the male Kl-96-13 and Um-Alsela female. The other three samples were Khalas 4, Kl-96-13 and Um-Alsela, parents of both populations. As expected, we found that

in both populations females and males shared one allele with a size of 339 bp, while the second allele was strictly limited to male palms with size 346 bp (Figure 7.8).

In an attempt to quantify the ability of the PDK_30s101A SSR primers to differentiate consistently between male and female plants across the genetic diversity of the date palm, 96 accessions of diverse origins were used (Table 7.2). The PDK_30s101A SSR primer amplified successfully all 96 accessions (Figure 7.9 and 7.10). Fifty-six accessions from Sanremo, Bordighera and USDA-ARS amplified a single allele of size 339 bp in the majority of females and an additional band of 346bp in the majority of males. Overall, four males did not contain the male specific allele, and four females showed a male pattern with both alleles at 339 and 346 bp.

In addition, forty female accessions from France, Iraq, Libya, Sudan and Iran showed the homozygous female specific allele 339 bp, except for a single sample from France and another from Iran that had the two band male pattern.

Another two microsatellite primer pairs PDK_30s101B and PDK_30s101C were designed using the same sequence for the same scaffold PDK_30s1038101 and both forward and reverse sequences are listed in Table 7.3. The two primer pairs were used to amplify the same eight date palms (1B, 7B, 16B, 35B, 36B, 38B, 43B and 44B). From the fragment results, we observed that all samples were amplified successfully and both primers producing a single peak with female and two distinct peaks with male. The allelic size for the female was 301 bp with PDK_30s101B and 232 bp with PDK_30s101C. All males analyzed with PDK_30s101B produced 301/308 bp product sizes, and 232/238 with PDK_30s101C.

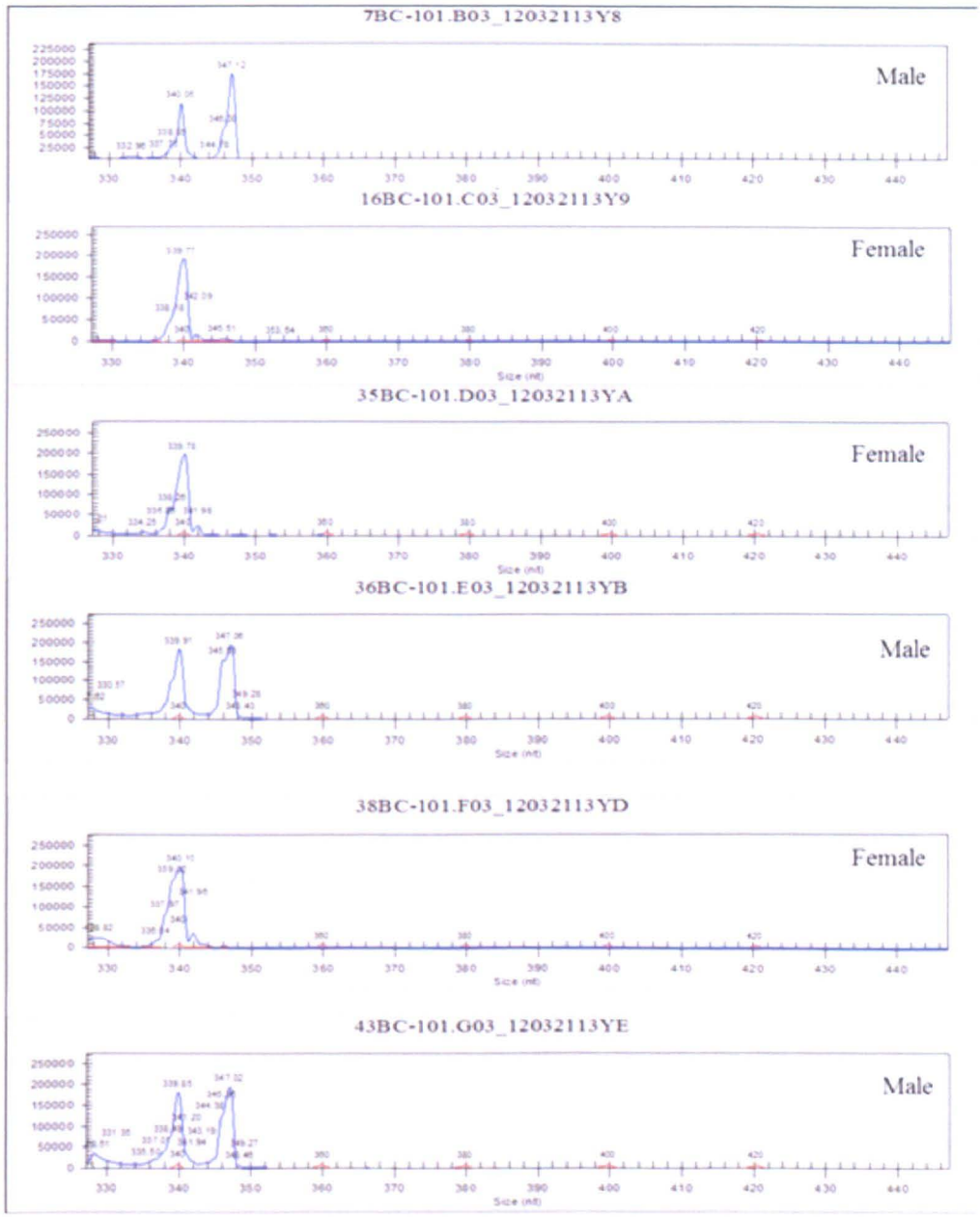


Figure 7.7: Fragment analysis results obtained from CEQ 8000 showing PCR products amplified by PDK_30s101A SSR primer from a selection of BC1 yielding fragment ranging from 339 bp for female and 339/346 bp for male.

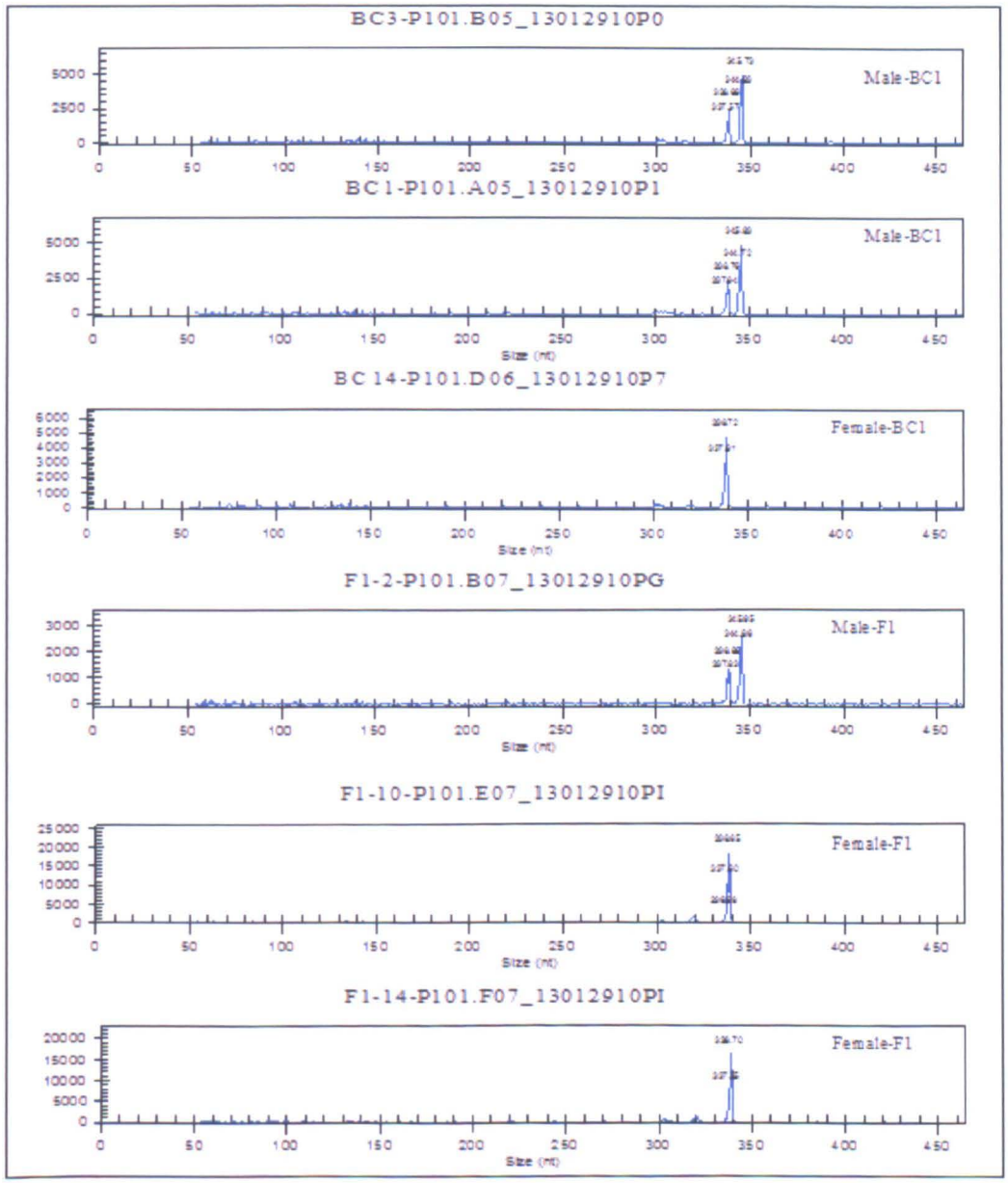


Figure 7.8: Fragment analysis results obtained from CEQ 8000 showing PCR products amplified by PDK_30s101A SSR primers using a selection of samples from BC1 and F1 populations.

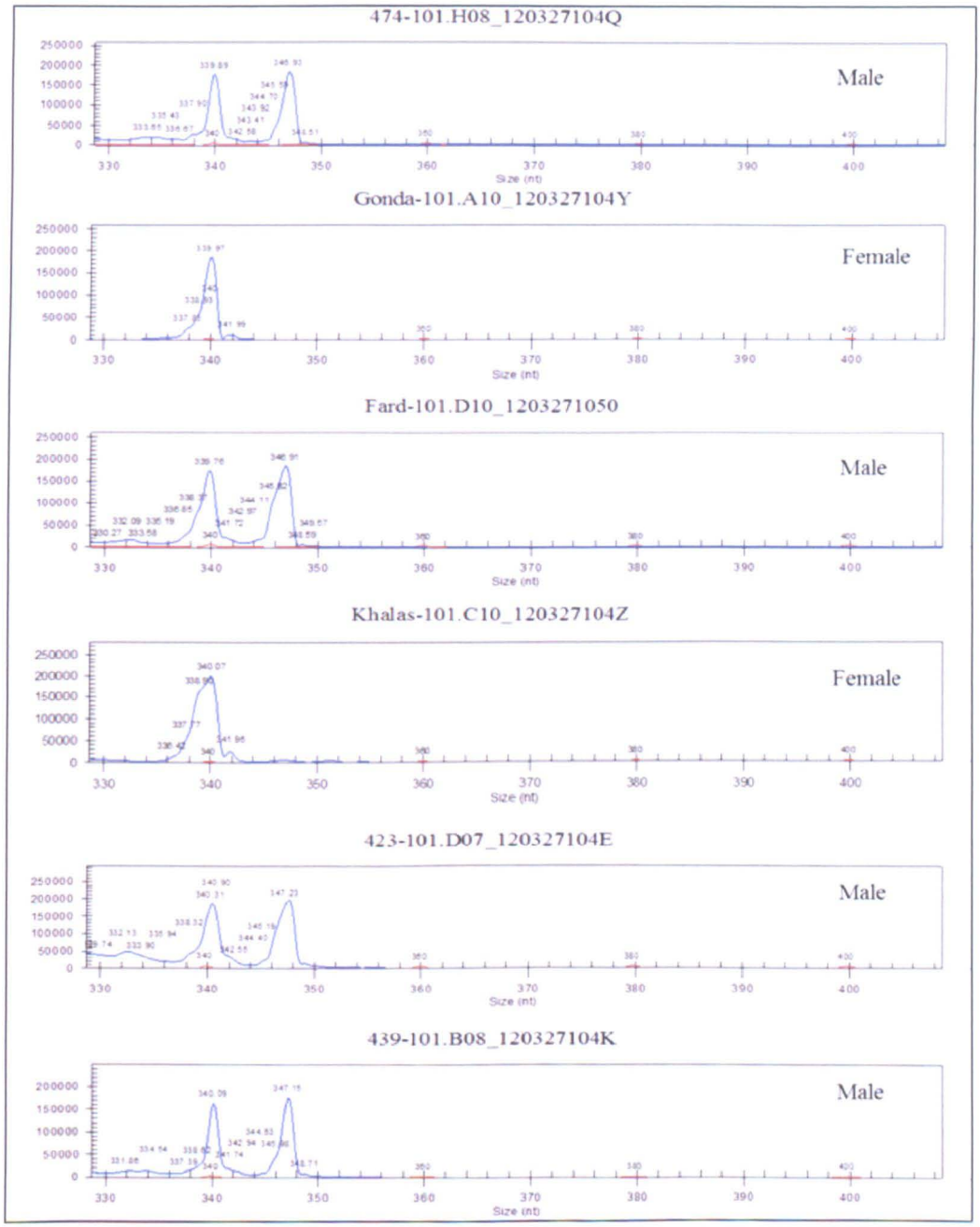


Figure 7.9: Fragment analysis results obtained from CEQ 8000 showing PCR products amplified by PDK_30s101A SSR primer using a range of accessions from Sanremo, Bordighera and USDA-ARS.

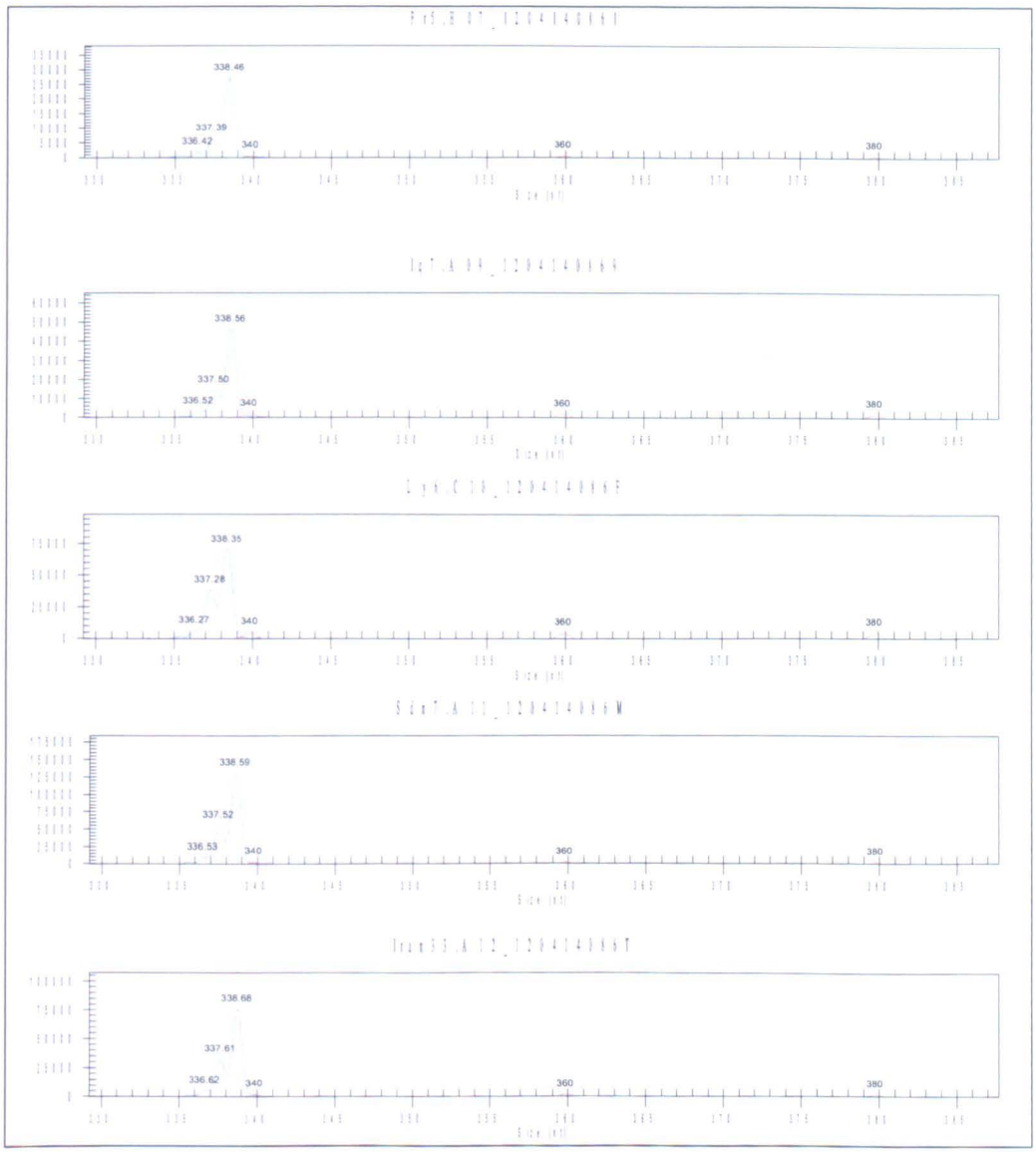


Figure 7.10: Fragment analysis results obtained from CEQ 8000 showing PCR products amplified by PDK_30s101A SSR primer using a range of female accessions from France, Iraq, Libya, Sudan and Iran.

7.4 Discussion

Attempts to identify the sex determining locus in date palm have until recently remained frustratingly unsuccessful. There have been several efforts to identify sex in date palm using different molecular markers. Specific primers for ribosomal RNA genes and RAPD molecular markers were used to detect the similarities and genetic relationships among three unknown males and four known females (Bartamoda, Sakkoty, Dagana, and Malkabi cultivars) of the Egyptian date palm (Ahmed *et al.*, 2006). However, results obtained by the primers specific for 18S rRNA gene showed no difference between the tested male and female plants and the seven varieties grouped into two clusters (Ahmed *et al.*, 2006). Ahmed *et al.* (2006) also found that the RAPD results were not discriminative enough to identify unknown male palms and identification of the male palms needed more advanced molecular studies.

RAPD and ISSR markers have been also used to try to identify sex-specific DNA markers for four female date palm cultivars (Sakoty, Bertmoda, Malkabi, Dagana) and three males (Dagana, Malkabi,Sakoty) from Aswan in Egypt (Younis *et al.*, 2008). Out of seven RAPD primers, three primers A10, A12 and D10 consistently yielded amplicons of 490, 750 and 800 bp, respectively from female plant samples, whereas A12 and D10 consistently yielded amplicons of 370 and 675 bp, respectively from all male samples.

Younis *et al.* (2008) also have reported that out of seven ISSR, five primers yielded clear amplification only in males but not in females, suggesting the possibility of using these markers to determine sex in date palm. However, so far no more efforts have been made to test these markers on a large number of date palm varieties to confirm their ability to differentiate between male and female.

Al-Dous *et al.* (2011) scanned more than three million SNP genotypes in the male and female genomes to detect polymorphisms that segregated with gender and found that the date palm uses an XY system of gender inheritance, which was the starting point for the researchers making use of this information for sex determination in date palm.

Furthermore, Al-Mahmoud *et al.* (2012) tested two DNA-based assays for sex determination taking into consideration the benefits of the findings that have previously been reported by Al-Dous *et al.* (2011). The first application was a PCR based restriction fragment length polymorphism that required a process of amplification followed by a restriction digestion with fragment detection by gel electrophoresis. In the second approach Al-Mahmoud *et al.* (2012) designed a unique PCR method in order to take advantage of the high level of heterogeneity present in the sex-linked region which removed the need for the restriction digestion step. Al-Mahmoud *et al.* (2012) were able to reduce the steps and simplify the entire process. It needs also to be noted that in their study they have used date palm samples that were collected from several farms throughout Qatar as well as from the U.S Department of Agricultural Research Service (USDA-ARS) based in California, in order to test the developed assays on multiple varieties. Overall, Al-Mahmoud *et al.* (2012) reported that they were able to developed two different assays that would allow fellow researchers to differentiate the gender of date palms especially during their early stages of development. In addition, the authors have suggested that their tests have been proven to work on many different varieties of date palms, suggesting that the studied polymorphism is ancient and widespread. When the sex-linked area is properly mapped and the sex controlling mutation fully

elucidated, the assays can be altered to take in consideration such information. Al-Mahmoud *et al.* (2012) concluded that there is "at least a 90% discrimination levels using these approaches".

In an effort to discriminate date palm gender, Elmeer and Mattat (2012) have used 14 microsatellite primer pairs (mPdCIR010, mPdCIR015, mPdCIR016, mPdCIR025, mPdCIR032, mPdCIR035, mPdCIR044, mPdCIR048, mPdCIR057, mPdCIR070, mPdCIR078, mPdCIR085, mPdCIR090, and mPdCIR093) developed by Bilotte *et al.* (2004) with 117 accessions representing 34 cultivars and 12 males from Qatar. They found 22 microsatellite loci that could be used to identify 9 out of 12 analysed males. The primer pair mPdCIR048 displayed heterozygous alleles with 160/190 allele size, which was recorded 4 times in the 12 individual male samples but not in any of the 117 female date palm (Elmeer and Mattat, 2012). In addition mPdCIR078/mPdCIR093 exhibited two different alleles 122/140 and 163/175, respectively, which were repeated twice among the 12 male but not in any of 117 females tested. The rest of the SSRs markers showed different heterozygous and homozygous alleles in both male and female. More recently, Cherif *et al.* (2013) found three genetically linked loci that are heterozygous only in males and the male-specific alleles were able to identify the gender in 100% of individuals tested. In their study, Cherif *et al.* (2013) also were able to confirm the existence of an XY chromosomal system with a non-recombining XY-like region in the date palm genome.

The three SSRs primer pairs PDK_30s101A, PDK_30s101B and PDK_30s101C developed in this study using the sequence of scaffolds PDK_30s1038101 (Al-Dous *et al.*, 2011) were found to be efficient showing

good amplification and high polymorphism in most accessions analysed. The PDK_30s101A primer pair produced two alleles, one having size 339 bp and shared between female and male. The second allele appeared limited to the male phenotype with size 346 bp. Minor allele size variation within PDK_30s101A was attributed to the analysis conditions carried out in two different laboratories located in Oman and UK, which affected the allele size calling; proven through using common control DNA preps in both sites. According to Deemer and Nelson (2010) variation in PCR product sizing with internal size standards does occur due to a variety of factors affecting denaturing capillary electrophoresis, such as polymer lot, buffer concentration, array quality, ambient laboratory temperature, and fluorescent label. Because of this variation, SSR markers alleles' size is subject to inconsistency in allele naming. Stewart *et al.* (2011) have reported that, even though SSRs are robust markers, however, detection of allele sizes can be difficult with some systems as well as consistency among laboratories.

Reproducibility of molecular markers such as RAPD, AFLP and SSR has been also tested in several European laboratory networks and differences in allele sizing were recorded across laboratories (Stewart *et al.*, 2011; Jones *et al.*, 1997). The allele size is not only dependent on the number of nucleotides but there are several factors affecting the allele size including: the mobility of the fragment in the electrophoresis, the distance of the allele from the standard used, the type of fluorescent label used, and the use of different instruments using different software (Stewart *et al.*, 2011).

The other two primer pairs PDK_30s101B and PDK_30s101C also showed only one allele size in females 301 bp and 232 bp, respectively. In contrast, all

males displayed two different alleles 301/308 bp (PDK_30s101B) and 232/238 bp (PDK_30s101C), one of the allele was observed in the females while the other was male specific (Table 7.2). Due to time constraints, the two primer pairs PDK_30s101B and PDK_30s101C were only tested on limited male and female accessions. Further confirmation of these results needs a larger number of accessions from different regions of the world.

Table 7-2: List of specific microsatellite primers, their name and observed allele size range for sex determination in date palm.

Oligo Name	Observed size (bp)	
	Female	Male
PDK_30s101A	339/339	339/346
PDK_30s101B	301/301	301/308
PDK_30s101C	232/232	232/238

In this study we reported three SSRs markers that could differentiate between female and male in date palm. All three loci produced two alleles; one allele was shared between female and male, while the other allele was strictly limited to the male. These results are in agreement with Al-Mahmoud *et al.* (2012), who reported that male date palms carry two alleles; one is male specific and the other is female allele using PCR-RFLP assay and PCR-only–based assay. Al-Mahmoud *et al.* (2012) also noted that females are homozygous with a single allele, which is also in agreement with our results. More recently, Cherif *et al.* (2013) have found three SSRs markers (mPdIRDP80, mPdIRDP50 and mPdIRDP52) to be potentially sex-linked and showed significantly high genetic differentiation between the male and female, as measured by the *Rst* index. However, these three loci showed more than two alleles compared to results in this study (Table 7.5). Four alleles were

reported in locus mPdIRDP80, two alleles (mPdIRDP80_311, mPdIRDP80_320) were shared between males and females, but allele's mPdIRDP80_213 and mPdIRDP80_329 appeared strictly limited to the male phenotype, suggesting Y-linkage. Locus mPdIRDP50 had two male-specific alleles, mPdIRDP50_199 and mPdIRDP50_201, while locus mPdIRDP52 yielded four male-specific alleles, with a duplicated allele in eastern males.

The validation and reliability of PDK_30s101A was tested using 380 samples representing Omani female and male date palm accessions, individuals from BC_1 and F_1 populations, as well as accessions from different origins (Table 7.3). The samples used in this study covered a broader genetic collection and examined larger numbers compared to samples used by Al-Mahmoud *et al.* (2012) and Cherif *et al.* (2013), suggesting that PDK_30s101A was specific enough to distinguish sex between backcrossed males and females and at the same time sensitive enough to distinguish sex in multiple varieties.

Table 7-3: Comparative studies for sex determination in date palm

Study	Technique used	Number of marker	Males and females shared alleles (X)	Male specific alleles (Y)	Used samples and their origin		Samples called incorrectly	
					Female	Male	Female	Male
This Study (95% accurate)	SSRs	PDK_30s 101A	339	346	151 Omani accessions -non map	43 Omani accessions-non map	6	0
		PDK_30s 101B	301	308	25 BC1-Oman-map + Khalas4 (parent A), 20 F1-Oman-map + Um-Alsela (parent A)	28 BC1-Oman-map, 14 F1-Oman-map + Kl-96-13 (parent B)	0	0
		PDK_30s 101C	232	238	17 Sanremo, 13 Bordighera, 4 USD, 7 France, 7 Sudan, 7 Libya, 11 Iraq, 10 Iran	12 Sanremo, 6 Bordighera, 2 USD	7	4
Cherif *et al.*, 2013	SSRs	mPdIRDP50	211	199	19 Tunisia, 1 Morocco, 2 Italy, 20 Djibouti, 6 Oman, 2 Syria, 2 Iraq	22 Tunisia, 2Italy, 20 Djibouti, 4 Oman, 1 Syria, 3 Iraq	0	0
			215	201				
			235	233				
			-	242				
			-	244				
			-	246				
		mPdIRDP52	220	207				
			224	209				
			226	210				
			228	212				
			230	216				
			234	218				
			240	-				
		mPdIRDP80	311	213				
			320	329				

Table 7.3 (Continued)

Study	Technique used	Number of marker	Males and females shared alleles (X)	Male specific alleles (Y)	Used samples and their origin		Samples called incorrect	
					Female	Male	Female	Male
Al-Mahmoud *et al.*, 2012	PCR-RFLP assay based on Bcl I digestion	Expected product sizes from digestion are 143 bp and 262 bp. In this assay, the female allele does not contain the restriction site and is not digested while the male allele is.			4 Qatar, 4 USD	1 Qatar, 2 USD, 5 backcross (USD)	0	2
	PCR-RFLP based on Hpa II digestion	Digestion of the female allele results in products of size 24, 59,180 & 189 bp, while Digestion of the male allele results in products of size 24, 59 & 369 bp.						
	PCR-RFLP based on Rsa I digestion	Expected product sizes from digestion of the female allele are 5, 205 & 283 bp, two males did not contain male-specific alleles						
	PCR-only based assay	Single band with female & two bands with male (not mentioned)			4 Qatar, 3 USD	7 backcross-USD		

In this study we have noted four males from Sanremo and Bordighera that did not contain the male specific allele and 13 female accessions that produced the male specific allele instead. Similar results were observed by Al-Mahmoud *et al.* (2012) using PCR-RFLP based on *Rsa*I digestion. They found that two males out of eight did not show the male-specific allele which could be due to the fact that this allele may not be common in the population; however this did not seem to affect the ability of the primer pair to anneal to other males. It is likely that there has been a recombination in this region in certain lineages, proving that the marker is not in the gene that controls sex determination in date palm. In addition, this primer pair can be potentially considered to predict a high level of discrimination between male and female date palms among

multiple varieties distributed across the wide range of cultivation, with an accuracy of 100% in the crosses, 96% in the Omani material and 86% in the broadest date palm germplasm, suggesting that this marker is close to the gene of sex determination (Table 7.3).

7.5 Conclusion

The results of this study complement the earlier investigations that have been conducted on sex determination of date palm (Al-Dous *et al.*, 2011; Al-Mahmoud *et al.*, 2012; Cherif *et al.*, 2013), and can be summarized as follows:

- Date palms contain an XY system of gender inheritance.
- The fragment analysis of PDK_30s101A locus gave two distinct alleles, one was shared between males and females, while the other allele appeared strictly specific to the male phenotype suggesting Y-linkage.
- Both linkage analysis and association analysis of large numbers of date palm confirm that the marker is close, but not in, the gene controlling sex determination in date palm.
- The SSR primer pair PDK_30s101A, has been more widely employed (194 accession from Oman and 96 accessions from different origin) and can be potentially considered to predict a high level of discrimination between male and female date palms among multiple varieties distributed across the wide range of cultivation, with an accuracy of 100% in the crosses, 96% in the Omani material and 86% in the broadest date palm germplasm.

- PDK_30s101A locus was also used to screen 90 individuals from BC_1 and F_1 population and was found to be specific to distinguish sex between all backcrossed and F_1 males and females.

- The linkage analysis suggests that the marker is within 1cM of the gene for sex determination.

- The other two primer pairs PDK_30s101B and PDK_30s101C showed the same results as PDK_30s101A, which provided evidence that these three SSRs markers are powerful and specific in distinguishing sex in multiple varieties as well as between backcrossed males and females. However, further confirmation of PDK_30s101B and PDK_30s101C loci needs a large number of accessions from different regions to be examined.

Chapter 8. GENERAL DISCUSSION AND CONCLUSION

8.1 Introduction

The date palm (*Phoenix dactylifera* L.) is an economically important crop grown over a million hectares worldwide, especially in the Middle East and North Africa. In 2003, the annual world production of dates was estimated to be 6.4 million metric tonnes (mt; El Hadrami and El Hadrami, 2009) and this production had increased to 6.8 million mt by 2010 (FAO Statistics, 2010).

The genetic pool of date palm is rich with over 5,000 cultivars known worldwide. However, this genetic pool is threatened by several biotic factors, such as disease and pests, and also by abiotic factors, such as salinity, drought, erosion, and excessive heat (Jain and Al-Khayri, 2011). Desertification and soil salinization are the most important factors affecting the natural habitat of date palms in the Middle East and North Africa. In certain parts of the Arabian Gulf (e.g. UAE), salt water intrusion has caused the loss of date palm in a number of production areas (Jain and Al-Khayri, 2011), while in Oman 'dubas' bug (*Ommatissus lybicus* DeBergevin) the red palm weevil (RPW; *Rhynchophorus ferrugineus* Olivier) and the lesser date moth (LDM; *Batrachedra amydraula*) are the most destructive pests of date palm (Al-Khatri, 2004). Additionally, the increased planting area of a limited number of cultivars with high commercial value risks neglecting less valuable cultivars and could also reduce the genetic diversity of date palm within regions of cultivation (El Kharbotly *et al.*, 2006).

In this study we aimed to develop and screen a set of high-quality microsatellite markers suitable for differentiation between and within Omani

date palm genotypes as well as between and within 'exotic' germplasm to obtain an accurate description and understanding of these genetic resources, along with tools for practical quality control within clonal propagation and breeding programmes. The developed and screened microsatellite markers along with SNP markers were also used to construct initial genetic maps of date palm based on the available (small) populations (BC_1 and F_1). An attempt to identify sex-specific SSR markers for date palm cultivars was also made using the published draft genome sequence of date palm. Evaluation of the levels of polymorphism and genetic diversity of date palm has become a prerequisite for the establishment of a research program aimed at rational germplasm conservation. In addition, constructing a genetic map would be very useful in a crop like date palm because of the disadvantages of breeding in dioecious tree crops. These include a long juvenile period, high levels of heterozygosity and the out-breeding nature, coupled with an inability to distinguish between male and female palms before flowering, 5-7 years after field planting.

8.2 Developing and screening microsatellite (SSR) markers for date palm

Microsatellites (SSRs) as DNA markers have proven to be useful tools (Khanam *et al.*, 2012; Ijaz, 2011). Three different sources of genomic microsatellite markers were used in this study (Figure 8.1). Constructing a new genomic DNA microsatellite-enriched repeat library sequenced with non-titanium reagents (Roche 454 Pyrosequencing) allowed us to identified two hundred simple sequence repeats for potential microsatellite construction (Chapter 4).

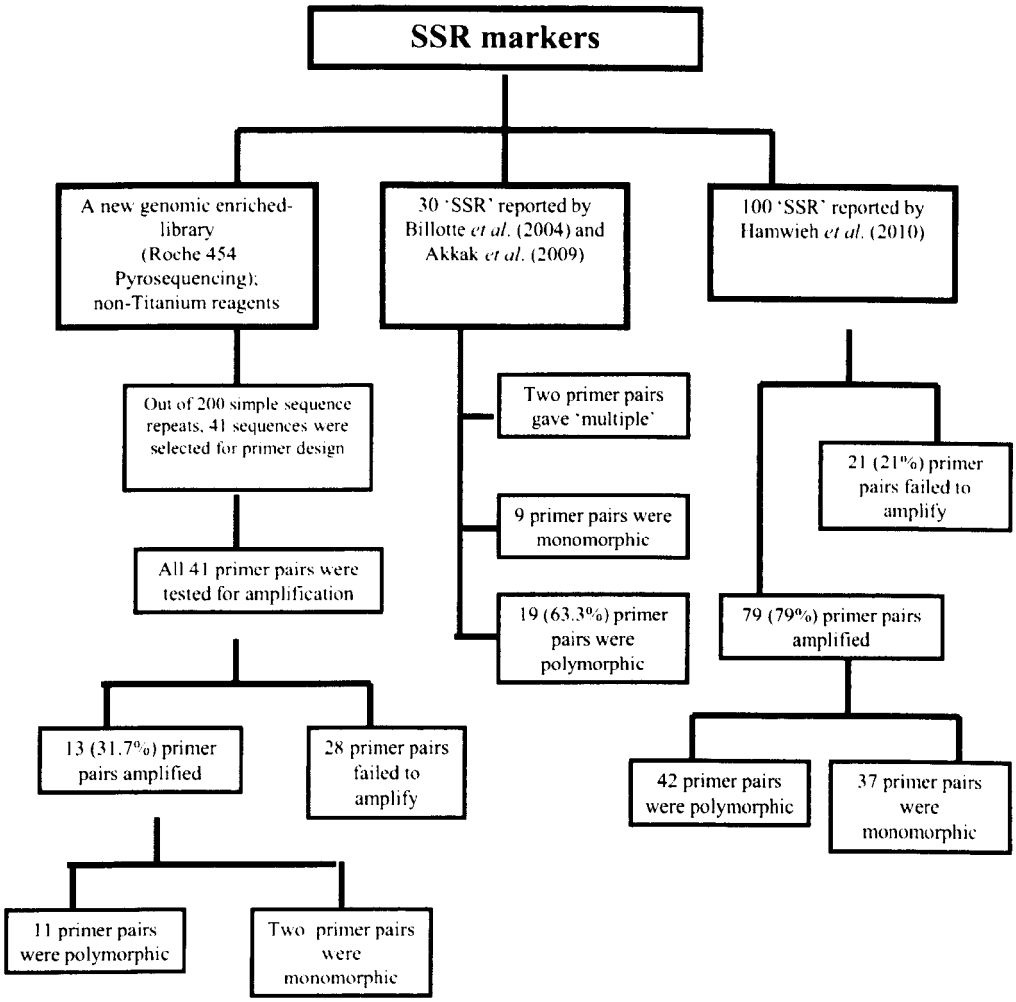


Figure 8.1: A flow diagram illustrating the three sources of microsatellite markers used in this study.

Out of the two hundred simple sequence repeats identified, 41 sequences were selected for primer design, while the remaining sequences did not have enough flanking sequence for successful primer design or had short repeat sequences below the minimum criteria chosen for use by the primer design software. Out of 41 primer pairs, a total of thirteen (31.7%) primer pairs amplified distinct bands, while 28 SSRs failed to amplify clean products from date palm DNA of the expected size. Out of thirteen SSR primer pairs, 11 were polymorphic in eight Omani genotypes while two were monomorphic.

The second source was 30 SSR markers reported by Billotte *et al.* (2004) and Akkak *et al.* (2009). Out of 30 SSR markers, 19 (63.3%) were polymorphic in the same eight Omani genotypes, while nine were monomorphic. The third source was 100 SSR primer pairs derived from a bioinformatics screen of the draft genome sequence which were obtained from the International Centre for Agricultural Research in the Dry Areas, Aleppo-Syria (ICARDA) and screened against the eight Omani date palms. Seventy-nine primers amplified a product of the expected size while 21 primers failed to amplify the genomic DNA. Out of 79 amplifying primer pairs, 42 primers were polymorphic, while 37 were monomorphic.

The transferability of microsatellite loci has been evaluated in different crops like cassava, rice, oil palm (Zaki *et al.*, 2012) as well as *Phoenix* species (Billotte *et al.*, 2004; Akkak *et al.*, 2009). Zaki *et al.* (2012) have reported that SSR markers developed for one species are known to detect homologous sites in related species. Eleven *Elaeis oleifera* SSR markers showed their ability to amplify DNA, not only in oil palm species, but also in coconut and other selected ornamental palms, thus confirming their ability to amplify across species and genera in the Arecaceae family (Zaki *et al.*, 2012).

Sixteen date palm specific SSR primer pairs revealed clear amplified bands within the expected allelic size range among 11 other *Phoenix* species as well as in species of four other palm genera, except for locus mPdCIR044, as reported earlier produced an erratic amplification (Billotte *et al.*, 2004). Akkak *et al.* (2009) have also reported cross-species amplification of 17 date palm specific SSR primer pairs in 14 other species across the genus *Phoenix*. The transferability across species and genera are advantageous as they save time

and cost in developing SSR markers for crops that have not been extensively studied and, in addition, facilitates comparative genetic studies (Zaki *et al.*, 2012). However, there is the potential for bias if microsatellites are used across long intraspecific genetic distances or even across species, as the genetically closer individuals to the source of the markers are more likely to produce consistent results, than genotypes distantly related to the source of microsatellite. Null alleles may also be more common, leading to an underestimation of heterozygosity.

The highest gene diversity or expected heterozygosity (0.75) among the eight Omani cultivars was observed using the 19 SSR markers from Billotte *et al.* (2004) and Akkak *et al.* (2009), suggesting that this set of markers are the most informative for the analysis of Omani date palm. However, no significant differences were noticed in the phenetic analysis of the eight Omani cultivars using data generated by 19 SSR, 42 SSR and combined 72 SSR markers. In addition, the phenetic analysis using 11 new genomic SSR markers produced different patterns by clustering cultivar Barni with Khori male and Nagla and creating separate branches for Bahlani and Khasab. This could be due to the source of these 11 markers, but perhaps is more likely to be an effect of the more limited sampling achieved by using 11 microsatellites, compared to 19, 42 and 72.

A subset of 12 SSRs makers derived from Billotte *et al.* (2004) and Akkak *et al.* (2009) were used to study the genetic diversity of Omani accessions as well as germplasm from different origins (Chapter 5).

The combined 72 SSR markers developed and screened in this study were used to construct a genetic map of date palm along with SNPs markers (Chapter 6; Figure 8.2).

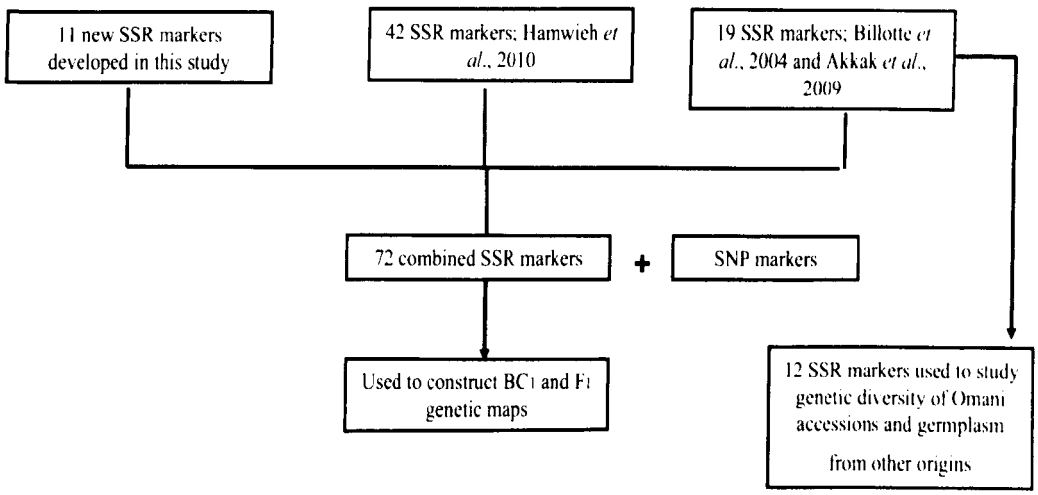


Figure 8.2: A flow diagram illustrating the polymorphic SSR from three sources and their use in this study.

8.3 Diversity analysis of Omani date palm cultivars

In Oman, there are more than 250 true 'varieties' of date palm propagated largely through offshoots at the base of mature date palms, in addition to a large number of genotypes originating from seed (MoA, 2011; MAF, 2005; Al-Khatri, 2004; Al-Ruqaishi, 2006). These could be an important component of plant improvement programs, providing plant breeders with sources of useful traits. These cultivars have been developed by continuous selection performed by date palm growers all over the Sultanate, mainly to improve crop yield and quality. The Omani date palm growers were keen to give names to either female or male clones to maintain their identity (El Kharbotly *et al.*,

2006). Most of these cultivars are recognized by their fruit characteristics such as size, shape, color and taste. To our knowledge, no detailed research has been conducted to study the genetic relationships among Omani date palm cultivars, except for 21 genotypes which were previously studied using SSRs markers (Al-Ruqaishi *et al.*, 2008).

In the present study, 12 microsatellite markers, chosen for their level of polymorphism, were used to examine genetic diversity and to study the genetic structure of most known date palm cultivars in Oman (151 female cultivars and 43 male trees). The analysis of Omani accessions resulted in a total of 188 alleles with a mean of 15.7 alleles per locus. However, this number of alleles is lower than that scored in 68 accessions from Sudan and Morocco at 343 alleles for 16 microsatellite loci (Elshibli and Korpelainen, 2008) which are probably because accessions from Sudan are more divergent than the Omani ones. The mean expected (*mHe*) and observed (*mHo*) heterozygosity detected for the Omani date palms were 0.744 and 0.606, respectively, while the respective values detected in the Sudanese date palm cultivars were 0.853 and 0.912 (Elshibli and Korpelainen, 2008). This result indicates that although there is clearly diversity between the Omani date palm accessions, however, Sudanese date palms have greater genetic diversity.

The AMOVA analysis in this study indicated that the majority of genetic variation at 94% (AMOVA, $P < 0.001$) was observed between individuals, within populations, while only 6% was observed between female and male populations. In contrast, the AMOVA analysis showed high genetic variation at 59% among 128 date palms sampled from the Figuig oasis (Morocco) using the 15 SSR markers developed by Billotte *et al.* (2004). Among these palms,

121 were females, belonging to 11 cultivars, while 7 were males suggesting that the Figuig cultivars were divergent from each other (Bodian *et al.*, 2012). Additionally, no genetic divergence was observed between Figuig male and female cultivars as confirmed by others (Sedra *et al.*, 1998; Sedra *et al.*, 2004; Zehdi *et al.*, 2004).

In the current study, the first and second principal components (PC1 and PC2) of a Principle Component Analysis accounted for 44.3% of the total molecular variation present. Although no clear differentiation between female and male accessions could be observed, however, the male palms are more constrained on axis 2 than the female palms. The derived Neighbor-Joining tree (NJ tree-bootstrap 1000 replication) clustered the female and male accessions into three main groups independently of their sex, which accords with the PCA, suggesting that the female and male Omani accessions are closely related to each other and that both male and female accessions have been derived from the same base populations. Sedra *et al.* (1998); Majourhat *et al.* (2002); Zehdi *et al.* (2004); Rhouma *et al.* (2008) have also reported independency of clustering patterns from the date palm cultivar sexuality in other regions.

This study suggested that the breeding histories in Oman and multiplication methods (seeds, offshoots and tissue culture) coupled with limited seed sources introduced from other germplasm to Oman, may have played an important role in the development of the current structure of Omani germplasm. In addition, most of the date palm growers tend to select a few cultivars with high commercial value and neglect other less valuable cultivars which affects the genetic diversity inside and between cultivation regions (El Kharbotly *et al.*, 2006) and could influence the overall genetic variation

present within Omani germplasm. Furthermore, the same cultivar is grown in different regions with different names or even different spelling of the name due to the different dialects from one region to the other which also could play a role in current population structure.

8.4 Diversity of Omani germplasm accessions and the germplasm from other countries

Using 12 SSR markers, we also studied the genetic variation of Omani date palm accessions alongside 'exotic' germplasm from Italy (Sanremo and Bordighera), USDA, France, Libya, Sudan, Iraq and Iran date palm cultivars and the patterns of genetic relationships among them were examined. The AMOVA analysis indicated that there were significant differences among populations as well as between accessions within populations (p = 0.001), but most of the total genetic variation was found within populations (79%) with only 21% of the variation found between populations. A similar observation has been made in various other out-crossing tree species which could be a general observation in long-lived wind pollinated tree species (Perera *et al.*, 2001).

An AMOVA analysis has been used to partition the genetic diversity of coconut in Sri Lanka which is taxonomically close to date palm and the result shows a very high percentage of within population variation (98.5%) for the tall coconut form *typical*. Perera *et al.* (2001) have suggested that the high level of within population variation in coconut which is an insect pollinated, out-crossing perennial species, is highly likely to be as a result of the common history of native Sri Lankan tall coconuts and their long generation time.

The 16 date palm specific SSRs primer pairs (Billotte *et al.*, 2004) plus one dodecanucleotide plastid minisatellite were used to investigate species delimitations in 308 accessions of *Phoenix* representing 12 species (Pintaud *et al.*, 2010). The genus *Phoenix* includes 14 species distributed in the Old World subtropics and tropics west of Wallace's line (Pintaud *et al.*, 2010). A high level of polymorphism was observed for all loci. All individuals from the same species clustered together, sometimes with high bootstrap support, supporting the existing taxonomy as well as confirming the good transferability of these SSR markers in most other species of *Phoenix*. Pintaud *et al.* (2010) have also reported high genetic diversity of *P. dactylifera* collected from various origins including: Tunisia, Italy, Oman, Djibouti, Niger, Senegal and Mauritania, giving a total of 146 accessions. The authors suggest that the high observed genetic diversity is probably due to the combined result of extensive natural variation and human selection. They also found that the three species with highest genetic diversity as measured by *He* were *P. dactylifera* (0.74), *P. reclinata* (0.73) and *P. loureiroi* (0.70), which were also characterized by the absence of monomorphic loci. While the other species *P. acaulis*, *P. canariensis*, *P. rupicola*, *P. pusilla*, *P. roebelenii*, *P. theophrasti* had low genetic diversity and some fixed private alleles, a pattern consistent with an evolution of small populations in isolation. Additionally, the cluster analysis placed *Phoenix atlantica* and *P. dactylifera* in one group, although without bootstrap support, suggesting that the two species share the same haplotype profile, at the shared plastid locus is different from the other species. This may also be indicative of a sister relationship.

The PCA analysis in this study which showed genetic differentiation between the Europe-Africa accessions (Sanremo, Bordighera, France, Libya and Sudan) and the West-Asia accessions (Oman, Iraq and Iran) was also in agreement with results obtained from the unrooted dendrogram and the bootstrap consensus tree which clustered the accessions in accordance with their geographic origin. However, one accession from Libya and one from Sudan were placed close to Iraqi accessions, which suggest that they or their ancestors could have been introduced from Iraq. In addition, five accessions were collected from USDA-ARS and the analysis placed them in accordance to their geographical origin; Medjool and Thory from Morocco and Algeria were placed within the Europe-Africa group, while Hilali, Barhee and Khalas from Oman, Iraq and Arabia were placed within the West-Asia group. This result is in agreement with Arabnezhad *et al.* (2011), who also reported the correspondence of genetic relationships and geographical location of date palm genotypes from Iran, Iraq and Africa. Zehdi *et al.* (2012) also obtained similar results with accessions from Iran, Iraq and North Africa using 14 SSRs markers, confirming that the North African cultivars are distinct to Middle Eastern ones. Arabnezhad *et al.* (2011) have suggested that this difference could have resulted from differences in domestication between the Middle-East and Africa, which is in opposition to the hypothesis that Mesopotamia (Fertile Crescent) is the only date palm domestication origin. In addition, the separation of Middle-Eastern accessions from other accessions supports the idea that 'increasing geographic distance between genotypes is associated with increasing genetic distance, on among-population diversity' (Arabnezhad *et al.*, 2011).

A similar study was conducted in Oil palm, in which the genetic relationship of four populations of *Elaeis oleifera* (Colombia, Honduras, Costa Rica and Panama) and two populations of *Elaeis guineensis* were studied using 18 *E. oleifera* genomic SSR markers. The six populations of oil palm clustered into two main groups, *E. oleifera* from Latin America and *E. guineensis* from Africa. The collections of *E. oleifera* from Costa Rica and Panama was very close to each other which is not surprising as Costa Rica and Panama are neighboring countries. The collection from Honduras was also close to Costa Rica and Panama fell into the same cluster. This could probably be due to the close proximity of Honduras to Costa Rica and Panama. However, the collections from Colombia were separated from the other three collections from Central America, although they fell into the same cluster. In addition, the two collections of *E. guineensis* from Africa fell into a separate cluster from the main group of *E. guineensis*. This suggested that the *E. oleifera* gSSR markers had the ability to reveal genetic diversity in the genus *Elaeis* in accordance to their origins and geographic distributions (Zaki et al., 2012).

It is important to understand the genetic makeup of date palm at the Regional level for efficient use of germplasm, classification, maintenance and conservation of date palm populations and their utilization in the improvement strategies. The agreement between the dendrogram, the Principle Component Analysis (PCA) and accessions geographical origins demonstrated that all accessions tested in this study were obtained from a reliable source with definite origin which is important for true estimates of genetic parameters in date palm molecular analysis

8.5 Genetic mapping of date palm

Genetic mapping in out-breeding heterozygous perennial crops is more complex due to the absence of complete homozygosity in the parents with most of these species not easily amenable to developing inbreed lines (Semagn *et al.*, 2006). Additionally, introgression of novel traits into elite cultivars would require extensive backcrossing which is hampered by long generation times and considerable land, given the long period to reach maturity and the physically large size of the crop.

In recent years, linkage maps have been constructed in coconut (Rohde *et al.*, 1999, Herrán *et al.*, 2000, Ritter *et al.*, 2000; Lebrun *et al.*, 2001) and oil palm (Mayes *et al.*, 1997; Moretzsohn *et al.*, 2000). However, no physical or linkage maps have yet been constructed for date palm because of the disadvantages mentioned above and the scarcity of controlled cross material from known provinces (Jain *et al.*, 2011).

This study reports the construction of a medium density genetic map in date palm using two populations, a small 'BC_1' and a smaller 'F_1'. Both populations are made by controlled pollination between out-crossing parental lines which have high levels of heterozygosity and they do not fit the classical models for F_1 and BC_1 which are usually constructed from inbred parental lines. The following are the basic steps of linkage maps construction for BC_1 and F_1 (Figure 8.3).

8.5.1 Population size

Identifying a suitable mapping population is the most critical decision made in construction of a linkage map for any crop. In order to develop a segregating population, selection of parental genotypes is very important as the segregation analysis and linkage mapping is difficult in the absence of extensive DNA polymorphism between selected parents.

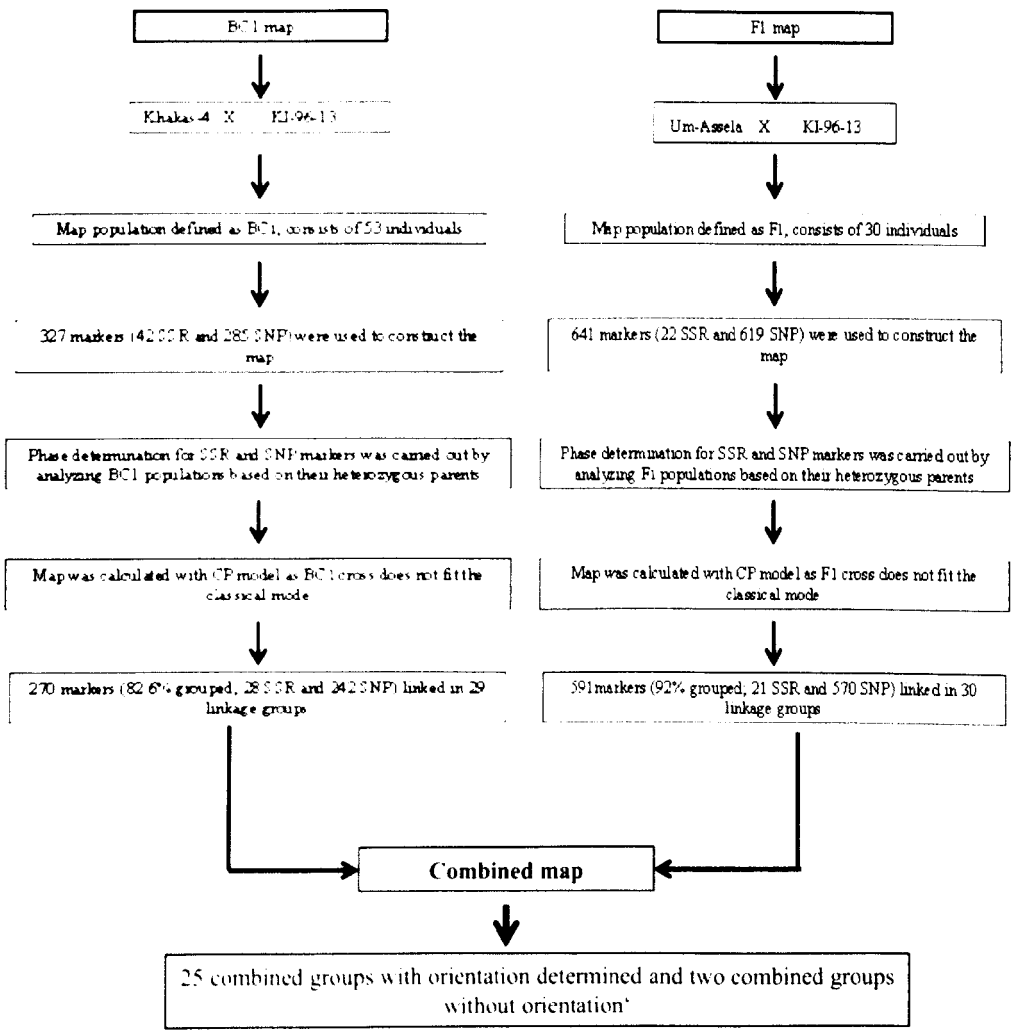


Figure 8.3: A flow diagram illustrating the process of construction of the genetic linkage maps in BC_1, F_1 population and their combination.

Therefore, polymorphism between parents must be present. The development of marker systems which generate large number of potential loci can partly

mitigate this problem and here we have used the high locus number DArT Seq approach to generate very large numbers of potentially polymorphic markers. The second important point is the size of mapping population since the accuracy of the genetic distance estimates is directly related to size and type of mapping population (Ferreira *et al.*, 2006) as the number of recombination cross-overs during meiosis determines the resolution of the map. Young (1994) has reported that a mapping population with less than 50 individuals would not be sufficient for construction of a reliable map. However, populations with a large number of individuals might not be practical - in date palm, they would require significant labor, time and space – or may simply not exist, which is also the case for date palm.

In the present study we used two populations with small numbers of individuals, as they are the only available mapping populations for date palm; a BC_1 population of 53 individuals and an F_1 population of 30 individuals. The BC_1 population was derived from the cross of Khalas-4 (selected as the recurrent parent as it produces high quality date fruits) by KI-96-13, while the F_1 population was derived from a cross between the Um-Assela cultivar by KI-96-13. The Um-Assela cultivar is well adapted to the conditions in the coastal regions of Oman (high salinity and humidity); however, it produces low quality date fruit. A similar mapping population (52 genotypes) has been used to generate a linkage genetic map in coconut (Herrán *et al.*, 2000). Our results showed that 53 and 30 individuals in BC_1 and F_1 populations, respectively, are good for the construction of initial genetic maps.

8.5.2 Marker distortion

In the current study the segregation patterns of the markers and detection of any distortion was determined in both populations by JoinMap 4.1 by performing a Chi-square test against expected segregation patterns ($p<0.05$ for significance). In the BC_1 population, out of 327 markers used , 203 (62.1%) markers segregated in the expected Mendelian ratios of 1:1, 1:2:1 and 1:1:1:1 for both marker types SSR and SNP, while 124 markers (37.9%) deviated from the expected Mendelian ratios. In the F_1 population, out of 641 markers used, 403 (62.9%) of the markers segregated in the expected Mendelian ratios 1:1, 1:2:1 and 1:1:1:1, whereas 238 (37.1%) showed deviation from the expected ratios for both SSR and SNP markers. Bandaranayake and Kearsey (2005) have reported 10.6% of markers deviation from Mendelian ratios in coconut, which is lower than 37.9% and 37.1% found in BC_1 and F_1, respectively.

8.5.3 Linkage map and marker distribution

The map constructed on the basis of the segregation data of the BC_1 population included 270 (82.6% mapped; 28 SSR and 242 SNP) markers and covered a total genetic distance of 1,486.7 cM with 57 (17%) markers remaining unlinked. These linked markers were assigned to 29 linkage groups (LG1-LG29), each containing from two to 27 loci per group and a linkage group length that varied from 3.9 cM (LG26) to 101.8 cM (LG12).

The F_1 population map included 591 (92% mapped; 21 SSR and 570 SNP) markers and covered a total genetic distance of 2,385.6 cM with 50 markers remaining unlinked (7.8%). The mapped markers were assigned to 30 linkage

groups (LG1-LG30) which had between 2 and 56 markers and a linkage group length varying from 8.5 cM (LG26) to 156.9 cM (LG12). The linkage groups numbers in both BC_1 and F_1 populations (29 and 30, respectively) was more than the expected number of 18 pairs of chromosomes ($2n = 2x = 36$) in date palm, as some linkage groups have few markers and these might be derived from the same chromosome. Similarly, Semagn *et al.* (2006) have reported 58 linkage groups, many more than the 21 haploid chromosomes of hexaploid wheat, suggesting that several areas of the genome remain undetected with the current set of markers. This large number of linkage groups can be reduced towards haploid chromosome numbers by increasing the number of markers or mapping populations.

The total length of the linkage maps in BC_1 was approximately 62% shorter than the total length of F_1 which could be partly explained by the number of recombination events being observed. Only a single recombination event can be observed per plant in BC_1, if this population had developed between two classical inbred lines, with recombination only detectable in the hybrid parent as the recurrent parent is homozygous. However, both parents in the F_1 population will have observable recombination events. In addition, the initial F_1 map had a genetic distance of 2,385.6 cM longer than the genetic maps reported for coconut (1,971 cM) and oil palm (1,743 cM), both of which have similar numbers of chromosomes ($2n = 2x = 16$). These differences could be explained by population size and type, number and type of markers, alongside potentially missing data and or incorrectly scored, with relatively minor scoring error for marker leading to map inflation.

8.5.4 Mapping the sex determination locus

After constructing a map, markers which co-segregate with important traits may then be identified like shell thickness in the oil palm (Mayes *et al.*, 1997; Moretzsohn *et al.*, 2000) and early flowering in coconut (Herrán *et al.*, 2000). The sex determination marker locus PDK_30s101A (locus 145) which was developed in the current study was also mapped in BC_1 at 42.8 cM, while in F_1 at 4.9 cM in linkage groups 18 and 29, respectively, and on combined group 19 at 42.8cM. The sex determination locus marker developed here was found to give a high level of discrimination between male and female date palms with an accuracy of 100% in the crosses, 96% in the Omani material and 86% in the broadest date palm germplasm among multiple varieties from Oman and other genetic origins. Sex discrimination has been studied in a number of species like papaya, *Silene latifolia*, melon and grapevine (Ming *et al.*, 2007; Farbos *et al.*, 1999; Boualem *et al.*, 2008; Martin *et al.*, 2009; Marguerit *et al.*, 2009). The development of markers close to the sex determination locus and preferably flanking markers, would allow date palm gender to be predicted before flowering, which normally takes 5-7 years. This would have relevance for breeding trials by eliminating non-productive male trees in the nursery before planting on a field scale planting.

8.5.5 Combined map

Based on common SSR and SNP markers present in BC_1 and F_1 maps, both maps were combined to form the final draft genetic map in date palm. Twenty-five linkage groups were combined with two to 15 common markers per group. However, due to the lack of a physical map for date palm the

orientation of the other two groups with single linkages could not be determined. 157 (58%) markers were common in BC_1 out a total of 270 markers, which is a high percentage and helps to validate the expected genetic linkage between the two populations used to generate the BC_1 and F_1 maps. It was also possible to localize the sex determination locus PDK_30s101A (locus 145) on combined group 19 at 42.8cM, which represents an important step towards validating this molecular marker tightly linked to the gene controlling the determination of sex in date palm.

8.6 Development of new microsatellite primers (SSRs) for gender discrimination in date palm

The gender of date palm was and still is one of the most important traits for date palm breeding, with a desire to determined gender as early as seedlings. Knowing gender early is of great importance in order to shorten the selection process of date palm selection and breeding. In addition, it has been one of the major practical impediments to establishing comprehensive breeding programmes in date palm. Very few controlled crosses are carried out because of the problem of 50% of the offspring being make and unproductive.

Al-Mahmoud *et al.* (2012) and Cherif *et al.* (2013) identified male-linked markers allowed them to identify date palm gender in 90% and 100% of individuals, respectively; however, this requires confirmation, as the markers were tested only in a small sample of date palms from limited numbers of origins. Their finding also confirmed the existence of an XY chromosomal system with a non-recombining XY-like region in the male date palm genome.

Three SSRs primer pairs PDK_30s101A, PDK_30s101B and PDK_30s101C were developed in this study using the sequence of scaffold PDK_30s1038101 (Al-Dous *et al.*, 2011) and were found to show good amplification and high polymorphism levels in most date palm samples analysed. The PDK_30s101A primer pair produced two alleles, one at 339 bp and common to female and male date palms. The second allele appeared strictly limited to male date palms with length 346 bp. The validation and reliability of this marker was tested using 380 samples representing Omani female and male date palm accessions, individuals from the BC_1 and F_1 populations, as well as accessions from a very broad range of different origins (Table 8.1).

The other two primer pairs PDK_30s101B and PDK_30s101C showed female alleles of 301 bp and 232 bp, respectively. While males displayed two additional alleles; 308 bp (PDK_30s101B) and 238 bp (PDK_30s101C). One of the alleles was shared with the females while the other was male specific with an accuracy of 100% in the tested controlled pollination samples. However, more samples from different origins are needed to confirm the accuracy of these two markers. The data presented in this study augments previous reports of the existence of an XY chromosomal system in date palm males and the availability of three markers with a high degree of confidence for gender differentiation in date palm from different origins. This will have a major impact on date palm breeding, reducing the time required for selecting female plants and facilitating the genetic improvement of this crop, through making seedling selection viable economically.

Table 8-1: Gender-specific marker in date palm, male and female alleles, type, origin, number of samples used and their accuracy.

Marker	Males and females shared alleles (X)	Male specific alleles (Y)	Samples used		
			Type and origin	Number	Accuracy (%)
PDK_30s101A	339	346	BC_1 and F_1 Omani populations	**90** (BC_1; 25 F + 28 M, F_1; 20 F + 14 M) and 3 parents	100%
			Omani accessions	**290** (151 F + 43 M)	96%
			Accessions from Sanremo, Bordighera, USD, France,Sudan, Libya, Iraq, Iran	**96** (76 F + 20 M)	86%

F: Female, M: Male

8.7 Major contributions made by this study

The foremost contributions made by this study are: the development and screening of microsatellite markers for date palm, studying the Omani germplasm in relation to the wider date palm germplasm, development of sex-linked markers and the construction of the first genetic map for date palm.

The markers developed and screened in this study form a good resource for phylogenetic analysis as well as germplasm conservation. They have been thoroughly tested for reliability in Omani accessions. They were shown to be reliable, polymorphic and able to distinguish date palm genotypes uniquely. A set of 12 SSR markers were able to differentiate between closely related germplasm and identify individuals uniquely (Omani germplasm) and allowed a comparison with 'exotic' germplasm. This study provides a picture of the population structure present in Oman and available for crop improvement,

suggesting that Omani germplasm is reasonably closely related and suggests that new foreign cultivars from different origins should be introduced to help diversifying the Omani germplasm. The unique identification of Omani date palm cultivars through the application of genetic markers would be of great interest in the description, registration and certification of planting material and for a rational management and germplasm conservation strategy to control genetic erosion of this important crop.

The study also provides a marker that can be used to distinguish between male and female date palms. The propagation of date palm has been primarily through off-shoots because of the economic consequences of using seed-derived material and the consequent 50% male palms that would be expected in field plantings. This marker will help to eliminate male palms that will not give any return, as well as use valuable water, fertilizer and planting space. In addition, constructing an initial genetic map for date palm is a very fruitful development, potentially allowing the identification of molecular marker tightly linked to traits of interest (here used for the formal confirmation that PDK_30s101A is very close to the sex determination locus genetically, with no recombinants in (2 x 53) + (30 x 1) = 136 detectable meiosis events) and also this will help address genetically controlled problems.

8.8 Summary of results

1. Development of a new set of SSR markers; 11 derived from a genomic library and a further 42 derived from untested primer sequences. The study also confirmed that the SSR from Billotte *et al.* (2004) and

Akkak *et al.* (2009) are highly informative among the sets of SSR primer pairs tested. (*Chapter 4*).

2. Together with existing SSR markers, these will supply the resources needed for the genotyping of date palm genotypes for quality control of breeding and clonal propagation programmes, genetic fingerprinting for origin studies and conservation as well as providing anchoring markers for constructing genetic maps and the identification of markers linked to various traits of interest, such as the sex determination gene, disease-resistant genes or genes involved in salinity stress tolerance (*Chapter 4*).
3. Omani accessions are related to each other, but show no clear genetic differentiation between female and male cultivars, suggesting that males have been derived from a broad range of origins (*Chapter 5*).
4. The West-Asia (Oman, Iraq and Iran) accessions were distinguished from Europe-Africa (Sanremo, Bordighera, France, Libya and Sudan) accessions and have their own autochthonous origin, reflecting a different phylogenetic history and perhaps even more than a single domestication origin (*Chapter 5*).
5. The Principal Coordinates Analysis (PCA) and cluster analysis (Unrooted UPGMA tree and a bootstrap consensus tree) placed accessions in accordance to their known origin, such as Medjool and Thory from Morocco and Algeria were placed within the Europe-Africa group, while Hilali, Barhee and Khalas from Oman, Iraq and Arabia, were placed within the West-Asia group, suggesting this

methodology has value for the placement of unknown origins and also for conservation and breeding (*Chapter 5*).

6. This study reports the first medium density genetic map in date palm. The available BC_1 population (53 individuals) allowed the construction of a linkage map with total genetic length of 1,486.7 cM, consisting of 270 markers (28 SSR and 242 SNP) distributed into 29 linkage groups. While the F_1 population (30 individuals) allowed the construction of a linkage map with total genetic length of 2,385.6 cM, consisting of 591 markers (21 SSR and 570 SNP) distributed into 30 linkage groups. Both crosses form the basis for further breeding work, with the F_1 potentially carrying important genes for salinity tolerance (*Chapter 6*).
7. The developed sex-determination marker locus PDK_30s101A (coded as locus 145) was mapped in both the BC_1 and F_1 maps at 42.8 cM and 4.9 cM in linkage groups 18 and 29, respectively and on the combined group 19 at 42.8cM (*Chapter 6*).
8. It was possible to combine BC_1 and F_1 maps through common SSR and SNP markers with a total of 25 combined linkage groups (*Chapter 6*).
9. This study suggested that the PDK_30s101A locus can discriminate between male and female date palms among multiple varieties distributed across the wide range of cultivation, with a high degree of accuracy; 100% in the crosses, 96% in the Omani material and 86% in the broadest date palm germplasm. For breeding within Omani materials, this level of accuracy can be directly applied to breeding programmes and should herald the start of considerable breeding effort

in date palm, which has not happened in date palm due to the high cost of unidentified males planted within any controlled cross. (*Chapter 7*).

8.9 Future directions

This work has made considerable progress in developing molecular genetics in date palm and especially Omani date palm germplasm. A range of work that could follow on from this study is outlined below:

- Further date palm genotypes should be amplified using the 11 SSRs primer pairs derived in this study and the 42 SSRs derived from untested primer sequences obtained from ICARDA and their transferability to other *Phoenix* species evaluated.
- Further work should focus on combining more cultivars from broader origins for delineating a core collection of date palm, both within Omani date palm and also in broader germplasm. This will allow the development of effective conservation strategies and breeding programs.
- Furthermore, association mapping between genotype and phenotype data should be carried out in date palm to investigate the relationship between molecular markers and traits of interest, which could increase the selection efficiency and complement QTL analysis in the limited number of controlled crosses which currently exist. Both could lead to marker that can be applied for selection in seedlings from controlled crosses.
- Seedling from the Um-Assela containing F_1 should be created in large numbers and tested for response to salinity, to determine whether this operates at the seedling level and could potentially be used for early selection of tolerant genotypes. Even if this is not the case, large-scale planting of controlled crosses for screening on saline soils should be

carried out, potentially representing the next generation of crosses for genetic mapping and trait determination.

- Further confirmation of sex-linked SSRs markers (PDK_30s101B and PDK_30s101C) needs a larger number of accessions from different origins, prior to an attempt to walk towards the gene directly responsible, based on the identified recombinant genotypes and the available sequence scaffolds.
- The transferability of microsatellites described in this study to other *Phoenix* species should be tested to elucidate the characteristics of these markers.
- The developed SNP markers can be further characterised and the associated 64N tags from the DArT-seq method compared with the available genome sequence to determine how many can be physically located on the genome, as a prelude to beginning to integrate genetic and physical maps at a crude scale.

References

1. Abdulla M., Gamal O. (2010) Investigation on molecular phylogeny of some date palm (*Phoenix dactylifera* L.) cultivars by protein, RAPD, and ISSR in Saudi Arabia. Australian Journal of Crop Science, 4: 23-28.

2. Adawy S.S., Hussein E.B.H., Ismail S.E.M.E., El-Itriby H.A. (2004) Genetic Diversity using AFLPGenomic diversity in date palm (*Phoenix dactylifera* L.) as revealed by AFLPs in comparison to RAPDs and ISSRs. Arab Journal of Biotechnology, 8: 99-114.

3. Agarwal M., Shrivastava N., Padh H. (2008) Review: Advances in molecular marker techniques and their applications in plant sciences. Plant Cell Reproduction, 27:617–631.

4. Ahmed M.M.M., Soliman S.S., Elsayed E.H. (2006) Molecular identification of some Egyptian date palm males by females varites (*Phoenix dactylifera* L.) using RAPD markers. Journal of Applied Sciences Research, 2: 270-275.

5. Ahmed O., Chokri B., Noureddine D., Mohamed M., Mokhtar T. (2009) Regeneration and molecular analysis of date palm (*Phoenix dactylifera* L.) plantlets using RAPD markers. African Journal of Biotechnology, 8: 813-820.

6. Ahmed T.A., Al-Qaradawi A.Y. (2009) Molecular phylogeny of Qatari date palm genotypes using simple sequences repeats markers. BioTechnology, 8: 126-131.

7. Ainsworth C., Parker J., Buchanan-Wollaston V. (1998) Sex determination in plants. Current Topics in Developmental Biology, 38:167 – 223.

8. Akagi H., Yokozeki Y., Inagaki A., Fujimura T. (1997) Highly polymorphic microsatellites of rice consist of AT repeats, and a classification of closely related cultivars with these microsatellite loci. Theoretical and Applied Genetics, 94: 61-67.

9. Akbari M., Wenzl P., Caig V., Carling J., Xia L., Yang S., Uszynski G., Mohler V., Lehmensiek A., Kuchel H., Hayden M.J., Howes N., Sharp P., Vaughan P., Rathmell B., Huttner E., Kilian A. (2006) Diversity arrays technology (DArT) for high throughput profiling of

hexaploid wheat genome. Theoretical and Applied Genetics, 113: 1409-1420.

10. Akkak A., Scariot V., Torello- Marinoni D., Boccacci P., Beltramo C., Botta R. (2009) ' Development and evaluation of microsatellite markers in *Phoenix dactylifera* L. and their transferability to other Phoenix species. Biologia Plantarum, 53: 164–166.

11. Al-Dous E.K., George B., Al-Mahmoud M.E., Al-Jaber M.Y., Wang Y., Salameh S.M., Al-Azwani E.K., Chaluvadi S., Pontaroli A.C., DeBarry J., Arondel V., Ohlrogge J., Saie J.I., Suliman-Elmeer K.M., Bennetzen J.L., Kruegger K.R., Malek J.A. (2011) 'De novo genome sequencing and comparative genomics of date palm (*Phoenix dactylifera*)', in 1-8, in Nature BioTechnology, Advance online publication, May Edition, doi:10.1038/nbt 1860Al-Jibouri A.A.M., Adham K.M., (1990), 'Biochemical classification of date palm male cultivars', Journal of Horticultural Sciences, 65: 725–729.

12. Alhammadi M.S., Kurup S.S. (2012) Impact of Salinity Stress on Date Palm (*Phoenix dactylifera* L) – A Review, Crop Production Technologies, Dr. Peeyush Sharma (Ed.), ISBN: 978-953-307-787- 1, InTech, Available from: http://www.intechopen.com/books/crop-production-technologies/impact-of-salinitystress-on-date-palm-phoenix-dactylifera-l-a-review.

13. Al-Khalifah N.S., Askari E. (2003) Molecular phylogeny of date palm (*Phoenix dactylifera* L.) cultivars from Saudi Arabia by DNA fingerprinting. Theoretical and Applied Genetics, 107: 1266-1270.

14. Al-Khalifah N.S., Askari E. (2006) Detection of genetic diversity in date palm (*Phoenix dactylifera* L.). Nova Science Publishers, Inc. chapter V, pp. 137-149.

15. Al-Khalifah N. S., Askari E. (2007) Early detection of genetic variation in date palms propagated from tissue culture and offshoots by DNA fingerprinting. Acta Horticulturae, 736: 105-112.

16. Al-Khatri S. (2004) Date palm pests and their control. Date Palm Regional workshop on Ecosystem Based IPM for Date Palm in the Gulf Countries UAE University, Al-Ain [28-30 March 2004].

17. Al-Mahmoud M.E., Al-Dous E.K., Al-Azwani E.K., Malek J.A. (2012) DNA-based assays to distingusish date palm (Areceae) gender, American Journal of Botany: e7–e10.

18. Al-Moharbi S. (2011) 'Date Palm in Oman', adapted from the Oman Observer. Source http://www.nizwa.net/agr/dates/datepalm/datepalm.html. Accessed on 18-11-11.

19. Al-Moshileh A. M., Motawei M. I., Al-Wasel A., Abdel-Latif T. (2004) Identification of some date palm (*Phoenix dactylifera* L.) cultivars in Saudi Arabia using RAPD fingerprints. SQU, Agriculture and Marine Sciences, 9: 1-3.

20. Al-Ruqaishi A.I., Davey M., Alderson P., Mayes S. (2008) Genetic relationships and genotype tracing in date palms (*Phoenix dactylifera* L.) in Oman, based on microsatellite markers. Short communication, Plant Genetic Resources, 6: 70-72.

21. Al-Ruqaishi I. A. (2006) Clonal propagation and molecular analysis of date palm (*Phoenix dactylifera L.*). PhD Thesis, School of Biosciences, Division of Agricultural and Environmental Sciences, University of Nottingham, UK.

22. Al-Sadi A.M., Al-Jabri A.H., Al-Mazroui S.S., Al-Mahmooli I.H. (2012) Characterization and pathogenicity of fungi and oomycetes associated with root diseases of date palms in Oman. Crop Protection. 37: 1-6.

23. Al-Yahyai R., Al-Khanjari S. (2008) Biodiversity of Date palm in the Sultanate of Oman. African Journal of Agricultural Research, 3: 389-395.

24. Amorim E.P., Vilarinhos A.D., Cohen K.O., Amorim V.B.O., Santos-Serejo J.A., Silva S.O., Pestana K.N., Santos V.J., Paes N.S., Monte D.C., Reis R.V. (2009) Genetic diversity of carotenoid-rich bananas evaluated by Diversity Arrays Technology (DArT). Genetics and Molecular Biology (BRA), 32: 96-103.

25. Anonymous (2009) Oman date production. Directorate General of National Statistics, Ministry of Agriculture, Muscat, Oman.

26. Arabnezhad H., Bahara M., Mohammadia H. R., Latifianb M. (2011) Development, characterization and use of microsatellite markers for germplasm analysis in date palm (*Phoenix dactylifera* L.). Scientia Horticulturae, 134: 150-156.

27. Arif I.A., Bakir M.A., Khan H.A., Al Farhan A.H., Al Homaidan A.A., Bahkali A.H., Al Sadoon M., Shobrak M. (2010) A Brief Review of

Molecular Techniques to Assess Plant Diversity. International Journal of Molecular Sciences, 11: 2079-2096.

28. Bader S. M., Baum M., Mhierallah H. S. M., Choumane W. (2007) The use of RAPD technique for the detection of genetic stability of date plantlets derived from In vitro culture of inflorescence. J. EDU. & Sci, pp 149-159.

29. Bagali P.G., Prabhu P.D., Raghavendra K., Bagali P.G., Hittalmani S., Vadivelu J.S. (2010) Application of Molecular Markers in Plant Tissue Culture. Asia-Pacific Journal of Molecular Biology and Biotechnology, 18: 85-87.

30. Bandaranayake C. K., Kearsey M. J. (2005) Genome mapping, QTL analysis and MAS: importance, principle, constraints and application in coconut. Plant Genetic Resources Newsletter, 142: 47-54.

31. Bekheet S.A., Hanafy M.S. (2011) Towards Sex Determination of Date Palm. Date Palm Biotechnology, pp 551-566.

32. Benaceur M., Lanaud C., Chevallier M-H., Bounaga N. (1991) Genetic diversity of the date palm (*Phoenix dactylifera* L.) from Algeria revealed by enzyme markers. Plant Breeding, 107: 56-57.

33. Bendiab K., Baaziz M., Brakez Z., Sedra M.H. (1993) Correlation of isozyme polymorphism and bayoud-disease resistance in date palm cultivars and progeny. Euphytica, 65: 23-32.

34. Bendiab K., Baaziz M., Majourhat K. (1997) Preliminary date palm cultivar composition of Moroccan palm groves as revealed by leaf isoenzyme phenotypes. Biochemical Systematics and Ecology, 26: 71-82.

35. Bhat Z.A., Dhillon W.S., Rashid R., Bhat J.A., Dar W.A., Ganaie M.Y. (2010) The role of Molecular Markers in Improvement of Fruit Crops. Notulae Scientia Biologicae, 2: 22-30.

36. Billotte N., Marseilla N., Brottier P., Noyer J.L., Jacquemoud-Collet J.P., Moreau C., Couvreur T., Chevallier M.H., Pintaud J.C., Risterucci A.M. (2004) 'Nuclear microsatellite markers for the date palm (*Phoenix dactylifera* L.): characterization, utility across the genus Phoenix and in other palm genera', Molecular Ecology Notes, 4:256-258.

37. Billotte N., Risterucci A. M., Barcelos E., Noyer J. L., Amblard P., Baurens F.C. (2001) Development, characterization and across-taxa

utility of oil palm (*Elaeis guineensis Jacq.*) microsatellite markers. Genome, 44: 413-25.

38. Blas A.L., Yu Q., Chen C., Veatch O., Moore P.H., Paull R.E., Ming R. (2009) Enrichment of a papaya high-density genetic map with AFLP markers. Genome, 52: 716-725.

39. Bodian A., El Houmaizi M.A., Ndoye Ndir K., Hasnaoui A., Nachtigall M., Wehiling P. (2012) Genetic diversity analysis of date palm (*Phoenix dactylifera* L.) cultivars from Figuig oasis (Morocco) using SSR markers. International Journal of Science and Advanced Technology, 2: 96-104.

40. Boualem A., Fergany M., Fernandez R., Troadec C., Martin A., Morin H., Sari M.A., Collin F., Flowers J.M., Pitrat M., Purugganan M.D., Dogimont C., Bendahmane A. (2008) A conserved mutation in an ethylene biosynthesis enzyme leads to andromonoecy in melons. Science, 321: 836–838.

41. Bradshaw H. D., Stettler R. F. (1994) Molecular genetics of growth and development in Populus. II. Segregation distortion due to genetics load. *Theoretical and Applied Genetics*, 89: 551-558.

42. Bundock P.C., Eliott F.G., Ablett G.A., Benson A.D., Casu R.E., Aitken K.S., Henry R.J. (2009) 'Discovering SNPs for gene mapping in sugarcane using deep sequencing', paper presented to the Plant and Animal Genomes XVII Conference, San Diego, California, USA, 10-14 January.

43. Callen D.F., Thompson A.D., Shen Y., Phillips H.A., Richards R.L., Mulley J.C., Sutherland G.R. (1993) Incidence and origin of null alleles in the $(AC)_n$ microsatellite markers. The American Journal of Human Genetics, 52: 922-927.

44. Cameron D. R., Moav R. (1957) Inheritance in Nicotiana tabacum. XXVII. Pollen killer, an alien genetic locus inducing abortion of microspores not carrying it. Genetics, 42, 326–335.

45. Cao B. R., Chao C. T. (2002). Identification of date (*Phoenix dactylifera* L.) cultivars in California using AFLP markers. Horticultural Science, 37: 966-968.

46. Chao CT., Krueger R.R. (2007) The date palm (*Phoenix dactylifera* L.): overview of biology, uses and cultivation. Horticulture Science, 42: 1077-1082.

47. Charlesworth D. (2002) Plant sex determination and sex chromosomes. Heredity, 88: 94–101.

48. Chen H., He H., Zou Y., Chen W., Yu R., Liu X., Yang Y., Gao Y-M., Xu J-L., Fan L-M., Li Y., Li Z-K., Deng X.W. (2011) Development and application of a set of breeder-friendly SNP markers for genetic analyses and molecular breeding of rice (Oryza sativa L.). Theoretical and Applied Genetics, 123: 869-79.

49. Cherif E., Zehdi S., Castillo K., Chabrillange N., Abdoulkader S., Pintaud J-C., Santoni S., Salhi-Hannachi A., Gl_emin S., Aberlenc-Bertossi F. (2013) Male-specific DNA markers provide genetic evidence of an XY chromosome system, a recombination arrest and allow the tracing of paternal lineages in date palm. New Phytologist, 197: 409–415.

50. Chetelat R. T., Meglic R.T., Cisneros V. (2000) A genetic map of tomato based on BC1 lycopersicon esculentum X solanum lycopersicoides reveals overall synteny but suppressed recombination between these homeologous genomes. Genetics, 154: 857–867.

51. Clauss M.J., Cobban H., Mitchell-Olds T. (2002) Cross-species microsatellite markers for elucidating population genetic structure in Arabidopsis and Arabis (Brassicaeae). Molecular Ecology, 11: 591-601.

52. Coe E., Cone K., McMullen M., Chen S., Davis G., Gardiner J., Liscum E., Polacco M., Paterson A., Sanchez-Villeda H., Soderlund C., Wing R. (2002) Access to the maize genome: an integrated physical and genetic map. Plant Physiology, 128: 9–12.

53. Collard B.C.Y., Jahufer, M.Z.Z., Brouwer, J.B., Pang, E.C.K. (2005) An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. Euphytica 142:169-196.

54. Cone K.H., Cone E.H. (2009) Genetic mapping and maps. In: Bennetzen J.l and Hake S. (eds.) *Maize Handbook - Volume II: Genetics and Genomic*, 507.

55. Cornicquel B., Mercier L. (1997) Identification of date palm (*Phoenix dactylifera* L.) cultivars by RFLP: partial characterization of a cDNA

probe that contains a sequence encoding for a zinc finger motif. International Journal of Plant Science 158: 152-156.

56. Corniquel B., Mercier L. (1994) Date palm (*Phoenix dactylifera* L.) identification by RFLP and RAPD. Plant Science, 101: 163-172.

57. Daher A., Adam H., Chabrillange N., Collin M., Mohamed N., Tregear J.W., Aberlenc Bertossi F. (2010) Cell cycle arrest characterizes the transition from a bisexual floral bud to a unisexual flower in *Phoenix dactylifera*. Annals of Botany, 106: 255–266.

58. Dakheel A. (2005) Date Palm Tree and Biosaline Agriculture in the United Arab Emirates. In: The Date Palm: From Traditional Resource to Green Wealth', in UAE Center of Studies and Strategy Researches, Abu Dhabi, UAE 247-263.

59. Dakin E.E., Avise J.C. (2004) Microsatellite null alleles in parentage analysis. Heredity, 93: 504–509.

60. DeGiorgio M., Rosenberg N. A. (2009) An unbiased estimator of gene diversity in samples containing related individuals, Molecular Biology and Evolution, 26: 501–512.

61. De Mason D.A., Stolte K.W., Tisserat B. (1982) Floral development in *Phoenix dactylifera* L. Canadian Journal of Botany, 60: 1439–1446.

62. De Mason D.A., Tisserat B. (1980) The occurrence and structure of apparently bisexual flowers in the date palm, *Phoenix dactylifera* L. (Arecaceae). Botanical Journal of the Linnean Society, 81: 283–292.

63. Deemer D.L., Nelson D.C. (2010) Standardized SSR allele naming and binning among projects. BioTechniques, 49: 835-836.

64. Dizkirici A., Guren H.E., Onde S., Temel F., Akar T., Budak H., Kaya Z. (2008) Microsatellite (SSR) variation in Barley germplasm and its potential use for marker assisted selection in scald resistance breeding. International journal of Integrative Biology, 4: 9-15.

65. Draye X., Lin Y., Qian X., Bowers J., Burow G.B., Morrell P.L., Peterson D.G., Presting G.G., Ren S., Wing R.A., Paterson A.H. (2001) Toward Integration of Comparative Genetic, Physical, Diversity, and Cytomolecular Maps for Grasses and Grains, Using the

Sorghum Genome as a Foundation1. Plant Physiology, 125: 1325–1341.

66. Echt C. S., Nelson C. D. (1997) Linkage mapping and genome length in eastern white pine (*Pinus strobes* L.). Theoretical and Applied Genetics, 94: 1031-1037.

67. Echt C., Knapp S., Liu BH. (1992) Genome mapping with non-inbred crosses using Gmendel 2.0. Maize Genet. Coop. Newslett, 66: 27-29.

68. Edwards K., Barker J., Daly A., Jones C., Krap A. (1996) Microsatellite libraries enriched for several microsatellite sequences in plants. Biotechniques, 205: 758-760.

69. El Hadrami A., Daayf F., El Hadrami I. (2011) Date Palm Genetics and Breeding. pp. 479-502. In: Jain, S.M., J.M. Al-Khayri and D.V. Johnson (Eds.) Date Palm Biotechnology, Springer, Netherlands.

70. El Hadrami I., El Hadrami A. (2009) Breeding date palm. In: Jain S.M., Priyadarshan P.M. (eds) Breeding plantation tree crops. Springer, New York, pp 191–216.

71. El-Kharbotly A., Al-Mamari A., Al-Lawatia A. (2006) Enrichment of the gene pool of date palm through breeding, in Proceedings of the International Conference on Date Palm Production and Processing Technology, Muscat, Oman.

72. Elhoumaizi M. A., Deanand P. S., Fang J., Chao C. T. (2006) Confirmation of 'Medjool' date as Landrace variety through genetic analysis of 'Medjool' accessions in Morocco. Journal of the American Society for Horticultural Science, 131: 403-407.

73. Elhoumaizi M.A., Saaidi M., Oihabi A., Cilas C. (2002) Phenotypic diversity of date-palm cultivars (*Phoenix dactylifera* L.) from Morocco. Genetic Resources and Crop Evolution, 49: 483–490.

74. El-Kharbotly A., El-Mardi M., Al-Saadi N., Al-Mahruki Y. (1998) 'Towards the construction of a genetic map of date palms using the amplified fragment length polymorphism technique (AFLP)', in Proceedings of the First International Conference on Date Palms, Al-Ain, United Arab Emirates, 194–207.

75. Elmeer K., Mattat I. (2012) Marker-assisted sex differentiation in date palm using simple sequence repeats. 3 Biotech, 2: 241–247.

76. Elmeer K., Sarwath H., Malek J., Baum M., Hamwieh A. (2011) New microsatellite markers for assessment of genetic diversity in date palm (*Phoenix dactylifera* L.). Biotechnology, 1: 91–97.

77. Elshibli S. (2009) Genetic Diversity and Adaptation of Date Palm (*Phoenix dactylifera* L.)', an MBA Dissertation for the University of Helsinki. ISBN 978-952-10-5937-7. Source: http://ethesis.helsinki.fi.

78. Elshibli S., Korpelainen H. (2008) 'Microsatellite markers reveal high genetic diversity in date palm (*Phoenix dactylifera* L.) germplasm from Sudan', in Journal of Genetics, Vol.134, 251-260.

79. Elshibli S., Korpelainen H. (2009) Biodiversity of date palms (*Phoenix dactylifera* L.) in Sudan: chemical, morphological and DNA polymorphisms of selected cultivars. Plant genetic resources: Characterization and Utilization, 1-10.

80. Elshibli S., Korpelainen H. (2010) Identity of date palm (*Phoenix dactylifera* L.) germplasm in Sudan: from the morphology and chemical characters to molecular markers, Acta Horticulturae 859: 143-153.

81. El-Tarras A., Al-Tawati N., Al-Malki S. (2007) Genetic fingerprint of some KSA date palm cultivars using modern biotechnological techniques. Biotechnology, 6: 263-267.

82. Erskine W., Moustafa A.T., Osman A.E., Lashine Z., Nejatian A., Badawi T., Ragy, S. (2003) 'Date Palm in the GCC countries of the Arabian Peninsula', in International Center for Agricultural Research in the Dry Areas (ICARDA). Source: http://www.icarda.org/APRP/Datepalm/introduction/intro-body.htm, Accessed on 8-12-11.

83. Erskine W., Moustafa A.T., Osman A.E., Lashine Z., Nejatian A., Badawi T., Ragy S.M. (2004) Date palm Development in the Gulf Coorperation Council countries of the Arabian Peninsula. Summary Proceedings of a Regional Workshop, 29-31 May 2004, Abu Dhabi, UAE.

84. Excoffier L., Smouse P.E., Quattro J.M. (1992) Analysis of Molecular Variance inferred from metric distances among DNA haplotypes: applications to human mitochondrial DNA restriction data. Genetics, 131: 474-491.

85. FAO (2005) 'Food and Agriculture Organization of the United Nations', Food & Agriculture Organization of the United Nation. (http://www.faostat.fao.org), Accessed on 8-12-11.

86. FAO Statistics (2010) www.fao.org.

87. Farbos I., Veuskens J., Vyskot B., Oliveira M., Hinnisdaels S., Aghmir A., Mouras A., Negrutiu I. (1999) Sexual dimorphism in white campion: deletion on the Y chromosome results in floral asexual phenotype. Genetics, 151: 1187–1196.

88. Faris J.D., Laddomada B., Gill B. S. (1998) Molecular Mapping of Segregation Distortion Loci in *Aegilops tauschii*. Genetics, 149: 319–327.

89. Farooq S., Azam F. (2002) Molecular Markers in Plant Breeding-I: Concepts and Characterization. Pakistan Journal of Biological Sciences 5: 1135-1140.

90. Felsenstein J. (1985) Confidence limits on phylogenies: An approach using the bootstrap. Evolution, 39: 783-791.

91. Ferreira A., da Silva M. F., da Costa e Silva L., Cruz C. D. (2006) Estimating the effects of population size and type on the accuracy of genetic maps. Genetics and Molecular Biology, 29: 187-192.

92. Feuillet C., Leach J.E., Rogers J., Schnable P.S., Eversole K. (2011) Crop genome sequencing: lessons and Rationales. Trends in Plant Science, 16: 77-88.

93. Frankel O.H., Brown A.H.D., Burdon J.J. (1995) The conservation of plant biodiversity. Cambridge University Press. Source: Amazon.co.uk, accessed on 9/1/12.

94. Fu S., Zhan Y., Zhi H., Gai J., Yu D. (2006) Mapping of SMV resistance gene Rsc-7 by SSR markers in soybean. Genetica, 128: 63-9.

95. Gaston K.J., Spicer J.I. (2004) Biodiversity: an introduction. 2nd Edition. Blackwell.

96. Gaur R., Azam S., Jeena G., Khan A.W., Choudhary S., Jain M., Yadav G., Tyagi A.K., Chattopadhyay D., Bhatia S. (2012) High-Throughput SNP Discovery and Genotyping for Constructing a

Saturated Linkage Map of Chickpea (*Cicer arietinum* L.). DNA Research, pp. 1–17.

97. Gebhardt C., Salamini F. (1992) Restriction Fragment Length Polymorphism analysis of plant genomes and its application to Plant Breeding. International Review of Cytology, 135: 201-237.

98. Ghaleb H. H.A. (2008) Date palm cultivars atlas of United Arab Emirates. In: Ghaleb H. H.A. (ed) Zayed Center for Heritage and History Puplications. Al-Ain, UAE.

99. Ghazanfar S.A. (2003) Flora of the Sultanate of Oman. Vol. 1: Piperaceae to Primulaceae. Scripta Botanica Belgica 25, 262 pp.

100. Ghislain M., Spooner D.M., Rodriguez F., Villamon F., Nunez J., Vasquez C., Waugh R., Bonierbale M. (2004) Selection of highly informative and user-friendly microsatellites (SSRs) for genotyping of cultivated potato. Theoretical and Applied Genetics, 108: 881-890.

101. Goloenko I. M., Davydenko O. G., Shimkevich A. M. (2002) Segregation distortion of marker nuclear genes in alloplasmic and isoplasmic lines of barley. Genetika, 38: 944–949.

102. Grant D., Shoemaker R. C. (2001) Plant gene mapping techniques. Encyclopedia of Life Sciences, John Wiley & Sons, Ltd. www.els.net.

103. Gros-Balthazard M., Newton N., Ivorra S., Tengberg M., Pintaud J.-C., Terral J.-F. (in press) Origines et domestication du palmier-dattier (*Phoenix dactylifera* L.): état de l'art et perspectives d'étude. Anthropobotanica.

104. Guimaraes E.P., Ruane J., Scherf B.D., Sonnino A., Dargie J.D. (2007) Marker-assisted selection: Current status and future perspectives in crops, livestock, forestry and fish. Food and Agriculture Organization of the United Nations, Rome.

105. Gupta P.K., Balyan H.S., Edwards K.J., Isaac P., Korzun V., Röder M., Gautier M.F., Joudrier P., Schlatter A.R., Dubcovsky J., De la Pena R.C., Khairallah M., Penner G., Hayden M.J., Sharp P., Keller B., Wang R.C.C., Hardouin J.P., Jack P., Leroy P. (2002) Genetic mapping of 66 new microsatellite (SSR) loci in bread wheat. Theoretical and Applied Genetics, 105: 413–422.

106. Gurevich V., Lavi U., Cohen Y. (2005) Genetic variation in date palms propagated from offshoots & tissue culture. Journal of the American Society for Horticultural Science, 130: 46-53.

107. Hackett CA., Broadfoot LB. (2003) Effects of genotyping errors, missing values and segregation distortion in molecular marker data on the construction of linkage maps. Heredity, 90: 33-38.

108. Haider N, Nabulsi I, MirAli N (2012) Phylogenetic relationships among date palm (*Phoenix dactylifera* L.) cultivars in Syria using RAPD and ISSR markers. Journal of Plant Biology Research, 1: 12-24.

109. Haldane JBS. (1931) The cytological basis of genetical interference. Cytologia 3: 54–65.

110. Hammadi H., Mokhtar R., Mokhtar E., Ali F. (2009) New approach for the morphological identification of date palm (*Phoenix dactylifera* L.) cultivars from Tunisia. Pakistan Journal of Botany, 41: 2671-2681.

111. Hammadi H., Vendramin G.G., Ali F. (2011) Microsatellite diversity among date palm (*Phoenix dactylifera* L.) subpopulations. Pakistan Journal of Botany, 43: 1257-1264.

112. Hamwieh A., Farah J., Moussally S., Al-Shamaa K., Almer K., Khierallah H., Udupa S., Lababidi S., Malek J., Aaouine M., Baum M. (2010) Development of 1000 microsatellite markers across the date palm (*Phoenix dactylifera* L.) genome. Acta Horticulturae 3: 26-32.

113. Hamza H., Benabderrahim M.A., Elbekkay M. (2012) Investigation of genetic variation in Tunisian date palm (*Phoenix dactylifera* L.) cultivars using ISSR marker systems and their relation with fruit characteristics. Turkish Journal of Biology, 36.

114. Hayden M.J., Sharp P.J. (2001) Sequence tagged microsatellite proling (STMP): a rapid technique for developing SSR markers. Nucleic Acids Res., 29, e43.

115. Herrán A., Estioko L., Becker D., Rodriguez M.J.B., Rohde W., Ritter E. (2000) Linkage mapping and QTL analysis in coconut (*Cocos nucifera* L.). Theoretical and Applied Genetics, 101: 292–300.

116. Ijaz S. (2011) Microsatellite markers: An important fingerprinting tool for characterization of crop plants. African Journal of Biotechnology, 10(40): 7723-7726.

117. Imelfort M., Edwards D. (2009) De novo sequencing of plant genomes using second-generation technologies. Briefings in Bioinformatics, 10(6): 609-618.

118. IPGRI (2005) Descripteurs du Palmier dattier (*Phoenix dactylifera* L.). Ipgri, Rome

119. Jaccoud D., Peng K., Feinstein D., Kilian A. (2001) Diversity arrays: a solid state technology for sequence information independent genotyping. Nucleic Acids Res. [online], 29: e25. Available from http://nar.oxfordjournals.org/cgi/reprint/29/4/e25.

120. Jain S.M., Al-Khayri J.M., Johnson D.V. (2011) 'Date Palm Biotechnology', ISBN 978-94-007-1317-8, 1-342, Springer Dordrecht Heidelberg London New York.

121. Jain S.M., Brar D.S., Ahloowalia B.S. (2002) Molecular techniques in crop improvement. Kluwer Academic Publishers, The Netherlands.

122. Jaradat A.A., Zaid A. (2004) Quality traits of date palm fruits in a center of origin and center of diversity. Food Agriculture and Environment, 2: 208-217.

123. Jenczewski E., Gherardi M., Bonnin I., Prosperi J.M., Olivieri I., Huguet T. (1997) Insight on segregation distortions in two intraspecific crosses between annual species of Medicago (Leguminosae). Theoretical and applied genetics, 94: 682-691.

124. Johnson C., Cullis T.A., Cullis M.A., Cullis C.A. (2009) DNA markers for variety identification in date palm (*Phoenix dactylifera* L.). Journal of Horticultural Science & Biotechnology, 84: 591–594.

125. Jonah P.M., Bello L.L., Lucky O., Midau A., Moruppa S.M. (2011) Review: The importance of molecular markers in plant breeding programmes. Global Journal of Science Frontier Research, 11(5).

126. Jones C.J., Edwards K.J., Castaglione S., Winfield M.O., Sala F., van deWiel C., Bredemeijer G., Vosman B., Matthes M., Daly A., Brettschneider R., Bettini P., Buiatti M., Maestri E., Malcevschi A., Marmiroli N., Aert R., Volckaert G., Rueda J., Linacero R., Vazquez A., Karp A. (1997) Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. Molecular Breeding, 3: 381–390.

127. Jubrael J., Udupa S., Baum M. (2005) Assessment of AFLP-based genetic relationships among date palm (*Phoenix dactylifera* L.) varieties of Iraq. Journal of the American Society for Horticultural Science, 130: 442-447.

128. Kaga A., Ohnishi M., Ishii T., Kamijima O. (1996) A genetic linkage map of azuki bean constructed with molecular and morphological markers using an interspecific population (Vigna angularis V nakashimaet). Theoretical and Applied Genetics, 93: 658-663.

129. Karim K., Chokri B., Amel S., Wafa H., Richid H., Nouredine D. (2010) Genetic diversity of Tunisian date palm germplasm using ISSR markers. International Journal of Botany, 6: 182-186.

130. Karp A., Kresovich S., Bhat K.V., Ayad W.G., Hodgkin T. (1997) Molecular tools in plant genetic resources conservation: a guide to the technologies. IPGRI Technical Bulletin No. 2. International Plant Genetic Resources Institute, Rome, Italy.

131. Kasha K. J., Kao K. N. (1970) High frequency haploid production in barley (*Hordeum vulgare* L.). Nature, 225: 874-876.

132. Kearsey M.J. (1998) The principles of QTL analysis (a minimal mathematics approach). Journal of Experimental Botany, 49: 1619–1623.

133. Kgazal M.A., Salbi M.I., Alsaadawi I.S., Fattah F.A., Al-Jibouri A.A.M. (1990) Bisexuality in date palm in Iraq. Journal of Islamic Academy of Sciences, 3: 131-133.

134. Khanam S., Sham A., Bennetzen J.L., Aly M.A.M. (2012) Analysis of molecular marker-based characterization and genetic variation in date palm (*Phoenix dactylifera* L.). Australian Journal of Crop Science, 6: 1236-1244.

135. Khierallah H., Bader S.M., Baum M., Hamwieh A. (2011a) Assessment of genetic diversity for some Iraqi date palms (*Phoenix*

dactylifera L.) using amplified fragment length polymorphisms (AFLP) markers. African Journal of Biotechnology, 10: 9570-9576.

136. Khierallah H.S.M., Bader S.M., Baum M., Hamwieh A. (2011b) Genetic diversity of Iraqi date palms revealed by microsatellite polymorphism. Journal of the American Society for Horticultural Science, 136: 282- 287.

137. Kloda J.M. (2004) The development and utilisation of molecular genetic markers to study genetic diversity in Ononis repenns, O. spinosa and related species. PhD Thesis, Robinson College, University of Cambridge, UK.

138. Knox M. R., Ellis T. H.N. (2002) Excess heterozygosity conttributes to genetic map espansmn in pea recombinant lnbred populations. Genetics, 162, 861–873.

139. Kosambi D.D. (1944) The estimation of map distances from recombination values. Ann. Eugen. 12:172-175.

140. Kotze A., Muller G.H. (1994) Genetic relationships between southern African cattlebreeds. Proc. 3rd World Congr. Genet. Appl. Livestock Prod., Guelph, Canada, 20, 201.

141. Kumar S., Gill B.S., Faris, J. D. (2007) Identification and characterization of segregation distortion loci along chromosome 5B in tetraploid wheat. Molecular Genetics and Genomics, 278, 187–196.

142. Kumar S., Tamura K., Nei M. (2004) MEGA3: an integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Brief Bioinform, 5:150–163.

143. Ky P. C.-L., Orieux B. M., Trouslot P., Akaffou S., Louarn J., Charrier A., Hamon S., Noirot M. (2000) Interspecific genetic linkage map, segregation distortion and genetic conversion in coffee (*Coffea* sp.) Theoretical and Applied Genetics, 101: 569-676.

144. Lambrides C.J., Godwin I.D., Lawn R.J., Imrie B.C. (2004) Segregation distortion for seed testa color in mungbean (*Vigna radiate* L. Wilcek). Journal of Heredity, 95: 532–535.

145. Lander ES., Green P., Abrahamson J., Barlow A., Daly MJ., Lincoln SE., Newburg L. (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics, 1: 174–181.

146. Lebrun P., Baudouin L., Bourdeix R., Konan J.L., Barker J.H., Aldam C., Herran A., Ritter E. (2001) Construction of a linkage map of the Rennell Island Tall coconut type (*Cocos nucifera* L.) and QTL analysis for yield characters. Genome, 44: 962-790.

147. Liu K., Muse S.V. (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics, 21: 2128–2129.

148. Loegering W.G., Sears E. R. (1963) Distorted inheritance of stem-rust resistance of timstein wheat caused by a pollen killing gene. Canadian Journal of Genetics and Cytology, 5, 67–72.

149. Lu H., Romero-Severson J., Bernardo R. (2002) Chromosomal regions associated with segregation distortion in maize. Theoretical and Applied Genetics, 105: 622–628.

150. Ma H. (2003) Molecular characterization of the sex determination locus in Carica papaya L. PhD Thesis, University of Hawaii.

151. MAF (Ministry of Agriculture and Fisheries) (2005) Oman date production (2003-04), Ministry of Agriculture and Fisheries, Sultanate of Oman.

152. Mahmoudi H., Hosseininia G., Azadi H., Fatemi M. (2008) 'Enhancing date palm processing, Marketing and Pest Control through organic culture'. Journal of Organic Systems, 3: 29-39.

153. Majourhat K., Bendiab K., Medraoui L., Baaziz M. (2002) Diversity of leaf peroxidases in date palm (*Phoenix dactylifera* L.) as revealed in an example of marginal (seedling derived) palm groves. Scientia Horticulturae, 95: 31–38.

154. Mangelsdorf P.C., Jones D. F. (1926) The expression of Mendelian factors in the gametophyte of maize. Genetics, 11, 423–455.

155. Manly K.F., Cudmore Jr R.H., Meer J.M. (2001) Map Manager QTX, crossplatform software for genetic mapping. Mammalian Genome, 12: 930–932.

156. Mantovani P., Maccaferri M., Sanguineti M.C., Tuberosa R., Catizone I., Wenzl P., Thomson B., Carling J., Huttner E., DeAmbrogio E., Kilian A. (2008) An integrated DArT-SSR linkage map of durum wheat. Molecular Breeding, 22: 629-648.

157. Marguerit E., Boury C., Manicki A., Donnart M., Butterlin G., Némorin A., Wiedemann-Merdinoglu S., Merdinoglu D., Ollat N.,

Decroocq S. (2009) Genetic dissection of sex determinism, inflorescence morphology and downy mildew resistance in grapevine. Theoretical and Applied Genetics, 118: 1261–1278.

158. Martins W.S., Lucas D.C.S., Neves K.F.S., Bertioli D.J. (2009) WebSat - A Web Software for MicroSatellite Marker Development, Bioinformation, 3: 282-283.

159. Masmoudi-Allouche F., Anissa Cha^ari-Rkhis A., Kriaa W., Gargouri-Bouzid R., Jain S., Drira N. (2008) In vitro hermaphrodism induction in date palm female flower. Plant Cell Reports, 28: 1–10.

160. Matsushita S., Iseki T., Fukuta Y., Araki E., Kobayashi S., Osaki M., Yamagishi M. (2003) Characterization of segregation distortion on chromosome 3 induced in wide hybridization between indica and japonica type rice varieties. Euphytica 134: 27-32.

161. Mayes S, Hafeez F, Price Z, MacDonald D, Billotte N, Roberts J (2008). Molecular research in Oil Palm, the key oil crop for the future. In: Genomic of Tropical Crop Plants, 1: 371-403.

162. Mayes S., Jack P.L., Marshall D.F., Corley R.H.V. (1997) Construction of a RFLP genetic linkage map for oil palm (*Elaeis guineensis* Jacq.). Genome, 40: 116-122.

163. McCouch S.R., Chen X., Panaud O., Temnykh S., Xu Y., Cho Y.G., Huang N., Ishii T., Blair M. (1997) Microsatellite marker development, mapping and applications in rice genetics and breeding. Plant Molecular Biology, 35: 89-99.

164. Millan T., Winter P., Jungling R., Gil J., Rubio J., Cho S., Cobos M.J., Iruela M., Rajesh P.N., Tekeoglu M., Kahl G., Muehlbauer F.J. (2010) A consensus genetic map of chickpea (*Cicer arietinum* L.) based on 10 mapping populations. Euphytica 175:175-189.

165. Ming R., Yu Q., Moore P.H. (2007) Sex determination in papaya. Seminars in Cell and Developmental Biology, 18: 401–408.

166. MoA (2009) Annual Agriculture Statistics. Directorate General of Planning and Investment Promotions. Department of Statistic and Information, Ministry of Agriculture, Oman.

167. MoA (2011) Agriculture & Livestock Research-Five-Year Research Strategy 2011-2015, Directorate General of Agriculture &

Livestock Research, Ministry of Agriculture, Sultanate of Oman. I + 52 pages.

168. Mohamed Ahmed M.V.O., Bouna Z.E.O., Mohamed Lemine F.M., Djeh T.K.O., Mokhtar T., Mohamed Salem A.O. (2011) Use of multivariate analysis to assess phenotypic diversity of date palm (*Phoenix dactylifera* L.) cultivars. Scientia Horticulturae 127:367-371.

169. Mohan M., Nair S., Bhagwat A., Krishna T.G., Yano M. (1997) Genome mapping, molecular markers and marker-assisted selection in crop plants. Molecular Breeding, 3: 87-103.

170. Mondini L., Noorani A., Pagnotta M.A. (2009) Assessing Plant Genetic Diversity by Molecular Tools. Diversity, 1: 19-35.

171. Moretzsohn M.C., Nunes C.D.M., Ferreira M.E., Grattapaglia D. (2000) RAPD linkage mapping of the shell thickness locus in oil palm (*Elaeis guineensis* Jacq.). Theoretical and Applied Genetics, 100: 63-70.

172. Muchugi A., Kadu C., Kindt R., Kipruto H., Lemurt S., Olale K., Nyadoi P., Dawson I., Jamnadass R. (2008) Molecular Markers for Tropical Trees, A Practical Guide to Principles and Procedures. ICRAF Technical Manual no. 9. Dawson I and Jamnadass R. eds. Nairobi: World Agroforestry Centre.

173. Mullis K.B., Faloona F., Scharf S., Saiki R., Horn G., Erlich H. (1986) Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. Cold Spring Harbor Symposia on Quantitative Biology, 51: 263-73.

174. Naturland (2002) 'Date Palm: Exemplary description of 20 Crops', in Organic Farming in the Tropics and Subtropics, First Edition, 1-13.

175. Neale D.B., Devey M.E., Jermstad IS.D., Ahuja M.R., Alosi M.C., Marshall K.A. (1992) Use of DNA markers in forest tree improvement research. New Forests, 6: 391-407.

176. Nei M. (1973) Analysis of gene diversity in subdivided populations. Proceedings of the National Academy of Sciences, USA., 70: 3321-3323.

177. Nikaido A., Yoshimaru H., Tsumura Y., Suyama Y., Murai M. (1999) Segregation distortion of AFLP markers in *Cryptomeria japonica*. *Genes & Genetic System*, 74: 55-59.

178. Nyrén P. (2007) The history of pyrosequencing. Methods in Molecular Biology, 373:1–14.

179. Osman A.M.A. (2001) Development of date palm culture in Republic of Sudan. Presented in Date Palm Research and Development Network, Khartoum, Sudan. 22-17 August.

180. Ould Mohamed Salem A., Rhouma S., Zehdi S., Marrakchi M., Trifi M. (2008) Morphological variability of Mauritanian date-palm (*Phoenix dactylifera* L.) cultivars as revealed by vegetative traits. Acta Botanica Croatica, 67: 81–90.

181. Ovesna J., Polakova K., Leisova L. (2002) DNA Analyses and their applications in plant breeding. Czech Journal of Genetics and Plant Breeding (CJGPB), 38: 29–40.

182. Palliyarakkal M.K., Ramaswamy M., Vadivel A. (2011) Microsatellites in palm (*Arecaceae*) sequences. Bioinformation, 7: 347–351.

183. Paran I., Goldman I., Tanksley S.D., Zamir D. (1995) Recombinant Inbred Lines for Genetic-Mapping in Tomato. Theoretical and Applied Genetics 90: 542-548.

184. Paterson A., Lander E., Hewitt J., Paterson S., Lincoln S., Tanksley S. (1988) Resolution of quantitative traits into Mendelian factors by using a complete map of restriction fragment length polymorphisms. Nature, 335: 721–726.

185. Paterson A.H. (1996) Making genetic maps. In: A.H. Paterson (Ed.), Genome mapping in plants, pp. 23–39. R. G. Landes Company, San Diego, California; Academic Press, Austin, Texas.

186. Pavez Wellmann A., Escobar Araya H., Johnson D.V. (2007) Date palm cultivation in Chile and Peru (South America): current status and future prospects for development. Acta Horticulturae, Issue 736: 71-85.

187. Peakall R., Smouse P.E. (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research, Molecular Ecology Notes, 6: 288-295.

188. Peng J., Korol A.B., Fahima T., Ro der M.S., Ronin Y.I., Li Y.C., Nevo E. (2000) Molecular Genetic Maps in Wild Emmer Wheat, Triticum dicoccoides: Genome-Wide Coverage, Massive Negative Interference, and Putative Quasi-Linkage. Genome Research, 10: 1509–1531.

189. Pereira M. G., Lee M., Bramel-Cox P., Woodman W., Doebley J., Whitkus R. (1994) Construction of an RFLP map in sorghum and comparative mapping in maize. Genome, 37: 236–243.

190. Perera L., Russell J.R., Provan J., Powell W. (2001) Levels and distribution of genetic diversity of coconut (*Cocos nucifera* L., var. *Typica* form *typica*) from Sri Lanka assessed by microsatellite markers. Euphytica, 122: 381–389.

191. Perrier X., Jacquemoud-Collet J.P. (2006) DARwin software http://darwin.cirad.fr/.

192. Peters J.L., Constandt H., Neyt P., Cnops G., Zethof J., Zabeau M., Gerats T. (2001) A physical amplified fragmentlength polymorphism map of Arabidopsis. Plant Physiology, 127: 1579–1589.

193. Pintaud J-C., Zehdi S., Couvreur T. (2010) Species delimitation in the genus *Phoenix* (Arecaceae) based on SSR markers, with emphasis on the identity of the date palm (*Phoenix dactylifera* L.). In: Seberg A., Peterson G., Barfod A., Davis J. (eds) Diversity, phylogeny, and evolution in the monocotoledons. Aarhus Univ Press, Denmark, pp 267-286.

194. Powell W., Machray G.C., Provan J. (1996) Polymorphism revealed by simple sequence repeats. Trends in Plant Science, 7: 215–222.

195. Qi-Lun Y., Ping F., Ke-Cheng K., Guang-Tang P. (2008) Genetic diversity based on SSR markers in maize (Zea mays L.) landraces from Wuling mountain region in China. Journal of Genetics, 87: 287-291.

196. Queller D.C., Strassman J.E., Hughes C.R. (1993) Microsatellites and Kinship. Trends in Ecology and Evolution, 8: 285-288.

197. Rhouma S., Dakhlaoui-Dkhil S., Ould Mohamed Salem A., Zehdi-Azouzi S., Rhouma A., Marrakchi M., Trifi M. (2008) Genetic diversity and phylogenic relation-ships in date-palms (*Phoenix dactylifera* L.) as assessed by random amplified microsatellite polymorphism markers (RAMPOs). Scientia Horticulturae, 117: 53–57.

198. Risterucci M.A., Hippolyte I., Perrier X., Xia L., Caig V., Evers M., Huttner E., Kilian A., Glaszmann J-C. (2009) Development and assessment of Diversity Arrays Technology for high-throughput DNA analyses in Musa. Theoretical and Applied Genetics, 119: 1093-1103.

199. Ritter E., Rodriguez M.J.B., Herrán A., Estioko L., Becker D., Rohde W. (2000) Analysis of quantitative trait loci (QTLs) based on linkage maps in coconut (*Cocos nucifera* L.). In: (Arencibia, A.D., ed.) Plant Genetic Engineering: Towards the Third Millenium. Elsevier Science B.V., Amsterdam, pp. 42-48.

200. Roder M. S., Korzun V., Wendehake K., Plaschke J., Tixer M. H., Leroy P., Ganal M. W. (1998) A microsatellite map of wheat. Genetics, 149: 2007–2023.

201. Roger T., Chetelat R.T., Meglic V., Cisneros P. (2000) A genetic map of tomato based on BC_1 lycopersicon esculentum X solanum lycopersicoides reveals overall synteny but suppressed recombination between these homeologous genomes. Genetics, 154: 857–867.

202. Rohde W., Becker D., Kullaya A., Rodriguez M.J.B., Herrán A., Ritter E. (1999) Analysis of coconut germplasm biodiversity by DNA marker technologies and construction of a first genetic linkage map. In: (Oropeza C, Verdeil JL, Ashburner GR, Cardeqa R and Santamaria JM, eds.) *Current Advances in Coconut Biotechnology*. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 99-120.

203. Rounsley S., Marri P.R., Yeisoo Yu Y., He R., Sisneros N., Goicoechea J.L., Lee S.J., Angelova A., Kudrna D., Luo M., Affourtit J., Desany B., Knight J., Niazi F., Egholm M., Wing R.A. (2009) De Novo Next Generation Sequencing of Plant Genomes. Rice, 2: 35–43.

204. Rozen S., Skaletsky H.J. (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) Bioinformatics Methods and Protocols: Methods in Molecular Biology. Humana Press, Totowa, NJ, pp 365-386. Source code available at http: fokker.wi.mit.edu/primer3/.

205. Saghai-Maroof M.A., Biyashev R.M., Yang G.P., Zhangf Q., Allard R.W. (1994) Extraordinarily polymorphic microsatellite DNA in barley: Species diversity, chromosomal locations, and population dynamics. Proceeding of the National Academy of Sciences of the United States of America, 91: 5466-5470.

206. Schuelke M. (2000) An economic method for the fluorescent labelling of PCR fragments. Nature biotechnology, 18:233–234.

207. Scotti I., Magni F., Fink R., Powell W., Binelli G., Hedley PE. (2000) Microsatellite repeats are not randomly distributed within Norway spruce (*Picea abies* K.) expressed sequences. Genome, 43: 41-46.

208. Sedra M.H., Lashermes P., Trouslot P., Combes M., Hamon S. (1998) 'Identification and genetic diversity analysis of date palm (*Phoenix dactylifera* L.) varieties from Morocco using RAPD markers', in Euphytica, 103: 75–82.

209. Semagn K., Bjørnstad Å., Ndjiondjop M. N. (2006) Principles, requirements and prospects of genetic mapping in plants. African Journal of Biotechnology, 5: 2569-2587.

210. Semagn K., Bjornstad A., Ndjiondjop M.N. (2006) An overview of molecular marker methods for plants. African Journal of Biotechnology, 5: 2540-2568.

211. Semagn K., Bjørnstad Å., Skinnes H., Guri Marøy A., Tarkegne Y., William M. (2006) Distribution of DArT, AFLP, and SSR markers in a genetic linkage map of a doubled-haploid hexaploid wheat population. Genome, 49: 545–555.

212. Sharopova N., McMullen M.D., Schultz L., Schroeder S., Sanchez-Villeda H., Gardiner J., Bergstrom D., Houchins K., Melia-Hancock S., Musket T., Duru N., Polacco M., Edwards K., Ruff T., Register J.C., Brouwer C., Thompson R., Velasco R., Chin E., Lee M., Woodman-Clikeman W., Long M.J., Liscum E., Cone K., Davis G., Coe E.H.J. (2002) Development and mapping of SSR markers for maize. Plant Molecular Biology, 48: 463-481.

213. Sibov ST., de Souza Jr CL., Garcia AAF., Garcia AF., Silva AR., Mangolin CA., Benchimol LL., de Souza AP. (2003) Molecular mapping in tropical maize (Zea mays L.) using microsatellite markers.

1. Map construction and localization of loci showing distorted segregation. Hereditas (Lund), 139: 96-106.

214. Soliman S. S., Ali B. A., Ahmed M. M. M. (2003) Genetic comparisons of Egyptian date palm cultivars (*Phoenix dactylifera* L.) by RAPD-PCR. African Journal of Biotechnology, 2: 86-87.

215. Somers D.J., Isaac P., Edwards K. (2004) A high-density microsatellite consensus map for bread wheat (Triticum aestivum L.). Theoretical and Applied Genetics, 109: 1105–1114.

216. Soranzo N., Provan J., Powell W. (1999) An example of microsatellite length variation in the mitochondrial genome of conifers. Genome, 42: 158-161.

217. Sousa-Correia C.M. (2008) The ecosystem approach to sustainable management of the agroforestry systems of Q. ilex subsp. Rotundifolia in the south of Portugal. PhD Thesis, School of Biosciences, University of Birmingham, UK.

218. Southern E. (1975) Detection of specific sequences among DNA fragments separated by gel-electrophoresis. Journal of Molecular Biology, 98: 503-508.

219. Stam P. (1993a) Construction of integrated genetic linkage maps by means of a new computer package: JoinMap. Plant Journal, 3: 739-744.

220. Stam P. (1993b) JoinMap Version 1.4: A computer program to generate genetic linkage maps. CPRO-DLO, Wageningen. Centre for Plant.

221. Stewart S., Wickramasinghe D., Dorrance A.E., Robertson A.E. (2011) Comparison of three microsatellite analysis methods for detecting genetic diversity in *Phytophthora* sojae (Stramenopila: Oomycete). Biotechnology Letter, 33: 2217–2223.

222. Stodart B.J., Mackay M.C., Raman H. (2007) Assessment of molecular diversity in landraces of bread wheat (Triticumaestivum L.) held in an ex situ collection with Diversity Arrays Technology (DArT) Australian Journal of Agricultural Research, 58: 1174–1182.

223. Suiter K.A., Wendel J.F., Case J.S. (1983) Linkage-I: A pascal computer program for the detection and analysis of genetic linkage. Journal of Heredity, 74: 203-204.

224. Tang S., Yu J.K., Slabaugh M.B., Shintani D.K., Knapp S.J. (2002) Simple sequence repeat map of the sunflower genome. Theoretical and Applied Genetics, 105: 1124–1130.

225. Tanurdzic M., Banks J.A. (2004) Sex determining mechanisms in land plants. The Plant Cell, 16: 61–71.

226. Terral J-F., Newton C., Ivorra S., Gros-Balthazard M., de Morais C.T., Picq S., Tengberg M., Pintaud J-C. (2011) Insights into the historical biogeography of the date palm (*Phoenix dactylifera* L.) using geometric morphometry of modern and ancient seeds. Journal of Biogeography, 39: 929–941.

227. Tinker N.A., Kilian A., Wight C.P., Heller-Uszynska K., Wenzl P., Rines H.W., Bjornstad A., Howarth C.J., Jannink J-L., Anderson J.M., Rossnagel B.G., Stuthman D.D., Sorrells M.E., Jackson E.W., Tuvesson S., Kolb F.L., Olsson O., Federizzi L.C., Carson M.L., Ohm H.H., Molnar S.J., Scoles G.J., Eckstein P.E., Bonman J.M., Ceplitis A., Langdon T. (2009) New DArT markers for oat provide enhanced map coverage and global germplasm characterization. BMC Genomics, 10: 39.

228. Tisserat B., DeMason D.A. (1982) A scanning electron microscope study of pollen of *Phoenix* (Arecaceae). Journal of the American Society for Horticultural Science, 107: 883-887.

229. Van Ooijen J.W. (2006) JoinMap 4. Software for the calculation of genetic linkage maps in experimental populations. Kyazma BV: Wageningen, Netherlands.

230. Vos P., Hogers R., Bleekar M., Reijans M., Vandele T., Hornes M., Fritjers A., Pot J., Pelemman J., Kuiper M., Zabeau M. (1995) AFLP: A new technique for DNA finger printing. Nucleic Acid Research. 23: 4407-4414.

231. Wahid M. T., Kamruzzaman A.Z.M., Mohamed S., Abdul R. (2004) Comparative study of plant biodiversity data modeling in research perspective. Jurnal Teknologi Maklumat, 16: 157-171.

232. Wang C., Zhu C., Zhai H., Wan J. (2005) Mapping segregation distortion loci and quantitative trait loci for spikelet sterility in rice (*Oryza sativa* L.). Genetic Research, 86: 97–106.

233. Wang M., Li R-Z., Yang W-M., Du W-J. (2010) Assessing the genetic diversity of cultivars and wild soybeans using SSR markers. African Journal of Biotechnology, 9: 4857-4866.

234. Wenzl P., Carling J., Kudrna D., Jaccoud D., Huttner E., Kleinhofs A., Kilian A. (2004) Diversity arrays technology (DArT) for whole-genome profiling of barley. Proceedings of the National Academy of Sciences of the USA 101: 9915-9920.

235. Werner K., Pellio B., Ordon F., Friedt W. (2000) Development of an STS marker and SSRs suitable for marker- assisted selection for the BaMMV resistance gene rym9 in barley. Plant Breeding. 119: 517–519.

236. Wittenberg A.H..J, van der Lee T., Cayla C., Kilian A., Visser R.G.F., Schouten H.J. (2005) Validation of the high-throughput marker technology DArT using the model plant Arabidopsis thaliana. Molecular Genetics and Genomics. 274: 30-39.

237. Worldstat (2007) http://en.worldstat.info/Asia/Oman/Land

238. Wright S. (1965) The interpretation of population structure by F-statistics with special regard to systems of mating, Evolution 19: 395- 420.

239. Wu K.S., Tanksley S.D. (1993) Abundance, polymorphism and genetic mapping of microsatellites in rice. Molecular and general Genetics, 241: 225–235.

240. Xia L., Peng K., Yang S., Wenzl P., de Vicente M.C., Fregene M., Kilian A. (2005) DArT for high-throughput genotyping of cassava (*Manihot esculenta*) and its wild relatives. Theoretical and Applied Genetics. 110: 1092–1098.

241. Xie Y., McNally K., Li C.Y., Leung H., Zhu Y.Y. (2006) A high-throughput genomic tool: diversity array technology complementary for rice genotyping. Journal of Integrative Plant Biology 48: 1069-1076.

242. Yamagishi M., Takeuchi Y., Tanaka I., Kono I., Murai K., Yano M. (2010) Segregation distortion in F2 and doubled haploid populations of temperate japonica rice. Journal of Genetic, 89, 237–241.

243. Yang M., Zhang X., Liu G., Yin Y., Chen K. (2010) The Complete Chloroplast Genome Sequence of Date Palm (*Phoenix dactylifera* L.). PLoS ONE 5.

244. Yin T.M., DiFazio S.P., Gunter L.E., Riemenschneider D., Tuskan G.A. (2004) Large-scale heterospeci segregation distortion in *Populus* revealed by dense genetic map. Theoretical and Applied Genetics, 109: 451–463.

245. Yin X., Stam P., Dourleijn C.J., Kropff M.J. (1999) AFLP mapping of quantitative trait loci for yield-determining physiological characters in spring barley. Theoretical and Applied Genetics, 99: 244–253.

246. Yin Y., Zhang X., Fang Y., Pan L., Sun G., Xin C., Ba Abdullah M., Yu X., Hu S., Al-Mssallem I., Yu J. (2012) High-throughput sequencing-based gene profiling on multi-staged fruit development of date palm (*Phoenix dactylifera* L.), Plant Molecular Biology. 78: 617–626.

247. Young N.D. (1994) Constructing a plant genetic linkage map with DNA markers. In: Phillips RL, Vasil IK, editors. DNA Based Markers in Plants. Kluwer Academic Publishers, The Netherlands, pp 32-47.

248. Younis R.A.A., Ismail O.M., Soliman S.S. (2008) Identification of sex-specific DNA for date palm (*Phoenix dactylifera* L.) using RAPD and ISSR techniques. Research Journal of Agriculture and Biological Sciences, 4: 278-284.

249. Zaid A., (2002) 'Date palm cultivation.' FAO Plant Production and Protection Paper no. 156, Food and Agricuture Organisation of the United Nations, Rome, pp 73-105.

250. Zaid A., de Wet P.F., Djerbi M., Oihabi A. (2002) 'Date palm propagation' In: Zaid, A. (2002) 'Date palm cultivation.' FAO Plant Production and Protection Paper no. 156, Food and Agricuture Organisation of the United Nations, Rome, pp 73-105.

251. Zaki N. M., Singh R., Rosli R., Ismail I. (2012) *Elaeis oleifera* Genomic-SSR Markers: Exploitation in Oil Palm Germplasm Diversity and Cross-Amplification in Arecaceae. International Journal of Molecular Sciences, 13: 4069-4088.

252. Zalapa J.E., Cuevas H., Zhu H., Steffan S., Senalik D., Zeldin E., Mccown B., Harbut R., Simon P. (2012) Using next generation

sequencing approaches to isolate simple sequence repeat (SSR) loci in the plant sciences. American Journal of Botany, 99: 193-208.

253. Zamir D., Tadmor Y. (1986) Unequal segregation of nuclear genes in plants. *Botanical Gazette*, 147: 355-358.

254. Zehdi S., Cherif E., Rhouma S., Santoni S., Hannachi A.S., Pintaud J.C. (2012) Molecular polymorphism and genetic relationships in date palm (*Phoenix dactylifera* L.): The utility of nuclear microsatellite markers. Scientia Horticulturae, 148: 255-263.

255. Zehdi S., Sakka H., Rhouma A., Ould Mohamed Salem A., Marrakchi M., Trifi M. (2004) Analysis of Tunisian date palm germplasm using simple sequence repeat primer. African Journal of Biotechnology, 3: 215-219.

256. Zhang L. Y., Wang S. Q., Li H. H., Deng Q. M., Zheng A. P., Li S. C., Li P., Li Z. L., Wang J. K. (2010) Effects of missing marker and segregation distortion on QTL mapping in F2 populations. Theoretical and Applied Genetics, 121: 1071–1082.

Appendices

Appendix 1: Lists of 171 date palm microsatellite primers used in this study, their name, sequences and their source. *CACGACGTTGTAAAACGAC is M13-extension.

Oligo's Lab Code	Oligo Name	Sequences (5 '-3')	Source
1	mPdCIR010F	***CACGACGTTGTAAAACGAC**ACC CCG GAC GTG AGG TG	Billotte *et al.*, 2004
	mPdCIR010R	CGT CGA TCT CCT CCT TTG TCT C	
2	mPdCIR015F	**CACGACGTTGTAAAACGAC**AGC TGG CTC CTC CCT TCT TA	Billotte *et al.*, 2004
	mPdCIR015R	GCT CGG TTG GAC TTG TTC T	
3	mPdCIR016F	**CACGACGTTGTAAAACGAC**AGC GGG AAA TGA AAA GGT AT	Billotte *et al.*, 2004
	mPdCIR016R	ATG AAA ACG TGC CAA ATG TC	
4	mPdCIR025F	**CACGACGTTGTAAAACGAC**GCA CGA GAA GGC TTA TAG T	Billotte *et al.*, 2004
	mPdCIR025R	CCC CTC ATT AGG ATT CTA C	
5	mPdCIR044F	**CACGACGTTGTAAAACGAC**ATG CGG ACT ACA CTA TTC TAC	Billotte *et al.*, 2004
	mPdCIR044R	GGT GAT TGA CTT TCT TTG AG	
6	mPdCIR048F	**CACGACGTTGTAAAACGAC**CGA GAC CTA CCT TCA ACA AA	Billotte *et al.*, 2004
	mPdCIR048R	CCA CCA ACC AAA TCA AAA AC	
7	mPdCIR050F	**CACGACGTTGTAAAACGAC**CTG CCA TTT CTT CTG AC	Billotte *et al.*, 2004
	mPdCIR050R	CAC CAT GCA CAA AAA TG	
8	mPdCIR057F	**CACGACGTTGTAAAACGAC**AAG CAG CAG CCC TTC CGT AG	Billotte *et al.*, 2004
	mPdCIR057R	GTT CTC ACT CGC CCA AAA ATA C	
9	mPdCIR070F	**CACGACGTTGTAAAACGAC**CAA GAC CCA AGG CTA AC	Billotte *et al.*, 2004
	mPdCIR070R	GGA GGT GGC TTT GTA T	

10	mPdCIR078F	**CACGACGTTGTAAAACGAC**TGG ATT TCC ATT GTG AG	Billotte *et al.*, 2004
	mPdCIR078R	CCC GAA GAG ACG CTA TT	
11	mPdCIR085F	**CACGACGTTGTAAAACGAC**GAG AGA GGG TGG TGT TAT T	Billotte *et al.*, 2004
	mPdCIR085R	TTC ATC CAG AAC CAC AGT A	
12	mPdCIR090F	**CACGACGTTGTAAAACGAC**GCA GTC AGT CCC TCA TA	Billotte *et al.*, 2004
	mPdCIR090R	TGC TTG TAG CCC TTC AG	
13	mPdCIR093F	**CACGACGTTGTAAAACGAC**CCA TTT ATC ATT CCC TCT CTT G	Billotte *et al.*, 2004
	mPdCIR093R	CTT GGT AGC TGC GTT TCT TG	
14	PDCAT1F	**CACGACGTTGTAAAACGAC**CTGAAATCTCTGTTCAAATCCA	Akkak *et al.*, 2009
	PDCAT1R	GTTTGGATCTATTTGTGAGTTATTTTCTTT	
15	PDCAT2F	**CACGACGTTGTAAAACGAC**GGCCTTCTCTTCCCTAATGGGA	Akkak *et al.*, 2009
	PDCAT2R	GTTTCTTGCCCCTGTTCTTTCCCTC	
16	PDCAT3F	**CACGACGTTGTAAAACGAC**CAAGGATAGGTGTGATGACCACC	Akkak *et al.*, 2009
	PDCAT3R	GTTTGTCCTTTTAACTTCTTGCTGGAATT	
17	PDCAT4F	**CACGACGTTGTAAAACGAC**TAACGAGTCCACACAC	Akkak *et al.*, 2009
	PDCAT4R	CTGGGTAAAGCTTATAAG	
18	PDCAT5F	**CACGACGTTGTAAAACGAC**GGCCCGTCCTTGGATTAGAG	Akkak *et al.*, 2009
	PDCAT5R	CTACGTTGTCCCGTCAATTGG	
19	PDCAT6F	**CACGACGTTGTAAAACGAC**AATCAGGGAAACCACAGCCA	Akkak *et al.*, 2009
	PDCAT6R	GTTTAAAGCCTTCTCAAGATAGCCTCAG	
20	PDCAT8F	**CACGACGTTGTAAAACGAC**GCTTAAGTGGTTAGTTGCCAA	Akkak *et al.*, 2009
	PDCAT8R	GTTTGGCAGAAGTATTGAAAAGTTGA	
21	PDCAT10F	**CACGACGTTGTAAAACGAC**CACTGCTCCTGTTGCCCTGT	Akkak *et al.*, 2009
	PDCAT10R	TGTAGAAGGGCAGAGGACGG	

22	PDCAT11F	**CACGACGTTGTAAAACGAC**TTAGTAGACTCCCCACCGTCCT	Akkak *et al.*, 2009
	PDCAT11R	GTTTCATGGTGCTGGAGAATGAA	
23	PDCAT12F	**CACGACGTTGTAAAACGAC**CATCGTTGATTCCTAACCCCTC	Akkak *et al.*, 2009
	PDCAT12R	GTTTAGATCTTGCATGGCAACGC	
24	PDCAT13F	**CACGACGTTGTAAAACGAC**TGTTGCCATTCACATGCTGC	Akkak *et al.*, 2009
	PDCAT13R	GTTTGGACTAGTCCCTCCCTCCC	
25	PDCAT14F	**CACGACGTTGTAAAACGAC**TGCTGCAAATCTAGGTCACGA	Akkak *et al.*, 2009
	PDCAT14R	GTTTACCCCTCGGCCAAATGTAA	
26	PDCAT15F	**CACGACGTTGTAAAACGAC**ACAGAGAGGTGGAGTTTTCGGATT	Akkak *et al.*, 2009
	PDCAT15R	TCTTCCTTTCAAACCAGCAAGCT	
27	PDCAT17F	**CACGACGTTGTAAAACGAC**CAGCGGAGGGTGGGCCTC	Akkak *et al.*, 2009
	PDCAT17R	GTTTCTCCATCTCCCTTTTTCTTCTGCTACTC	
28	PDCAT18F	**CACGACGTTGTAAAACGAC**CCTAAACCTGAATGAATCAAAGCA	Akkak *et al.*, 2009
	PDCAT18R	ACTAACATAAGGACAGTGCTATGTGATTG	
29	PDCAT20F	**CACGACGTTGTAAAACGAC**TTTCAGACACATCAAGTAACGATGA	Akkak *et al.*, 2009
	PDCAT20R	GTTTACGTCCACCCCAAGTTACGA	
30	PDCAT21F	**CACGACGTTGTAAAACGAC**GTGTTTGAAGATTGATTTTGTGTTATGAG	Akkak *et al.*, 2009
	PDCAT21R	GTTTCGAACTATAGGCATGCACAATAGTATATTG	
31	DateS1F	**CACGACGTTGTAAAACGAC**GGGCCTCTTGTTCTGCCTCTTTAT	New Primer
	DateS1R	CTCGCCAACGATATCTAACGGCTA	
32	DateS2F	**CACGACGTTGTAAAACGAC**ACCCGATCTAGGGACTCGAAGAAG	New Primer
	DateS2R	CGGAAGCAGCAGCAGAAGAAC	
33	DateS3F	**CACGACGTTGTAAAACGAC**TTGGAAGGGAAACACACACA	New Primer
	DateS3R	CCCTAACTTAATTTTCTTGTTTCTTG	

34	DateS4F	**CACGACGTTGTAAAACGAC**GGAATTCCCAAAACCAAAA	New Primer
	DateS4R	CTAACGATCCCGAAAACGAA	
35	DateS8F	**CACGACGTTGTAAAACGAC**TTTCCCTTGATAGGGGATAAGC	New Primer
	DateS8R	TCGGCATGTCCCATACTTCTATCC	
36	DateS9F	**CACGACGTTGTAAAACGACGAC**ACTTATTTGGGCCGCTACATC	New Primer
	DateS9R	ACATACAGACGCAAGCAGACGAGA	
37	DateS10F	**CACGACGTTGTAAAACGAC**AACTGGAGAAGGAAACAGCAGTGG	New Primer
	DateS10R	TCATCCCTTGTCACACTCTTGACC	
38	DateS11F	**CACGACGTTGTAAAACGAC**CGATCGCGAAGAGAGGAGAAAAA	New Primer
	DateS11R	AAACAATGATCTCTCCGGATCCAA	
39	DateS12F	**CACGACGTTGTAAAACGAC**CCACACCCTTGAGAAATCTGAACA	New Primer
	DateS12R	TTCTCCTTCTATTCTCCCCGGTTC	
40	DateS13F	**CACGACGTTGTAAAACGAC**ATGGTGGCGATGGCATCTGT	New Primer
	DateS13R	CCCAACCACTACGTACGACCTACC	
41	DateS14F	**CACGACGTTGTAAAACGAC**ACCACACATGCTCTCAGAATTGA	New Primer
	DateS14R	GGAAGTAGACGAACGTAGGGAACG	
42	DateS15F	**CACGACGTTGTAAAACGAC**ACGAAAGGGGAATAACAAGGAGAA	New Primer
	DateS15R	CCGCCAATCCCTACTCCACCT	
43	DateS16F	**CACGACGTTGTAAAACGAC**CTCCTTGCTCGAAACCCTAATCCT	New Primer
	DateS16R	ACTTACCTAGGCCTGGGGACCTC	
44	DateS17F	**CACGACGTTGTAAAACGAC**TGATCCCAGCCTGATCTTCTCTTC	New Primer
	DateS17R	CCTCACTCTCCTCCCCTCAGGTAT	
45	DateS18F	**CACGACGTTGTAAAACGAC**CCGACTTCCTTCCTCGGCTTATAC	New Primer
	DateS18R	TAGTGTTAAAGGCCCAGCTTGATG	
46	DateS19F	**CACGACGTTGTAAAACGAC**AAGCCTATTGGAGATCTCTCTCTCTC	New Primer
	DateS19R	AACCCGGTTTACCCGTTTTA	

47	DateS21F	**CACGACGTTGTAAAACGAC**TGGTCTAGGCCACTGATTGCTACTC	New Primer
	DateS21R	GTACCCTAACCCAAACCCAACCAC	
48	DateS22F	**CACGACGTTGTAAAACGAC**GGGAAGTTTTTGGCACCGTGAT	New Primer
	DateS22R	AGGCGTTTAGGTACGGGTAGGTCT	
49	DateS23F	**CACGACGTTGTAAAACGAC**CCGGGCCTGTACCGACATAGTA	New Primer
	DateS23R	GCGGGGTAGGAGGAAGGAAG	
50	DateS24F	**CACGACGTTGTAAAACGAC**CCGTACGACGAACACCTACCTACC	New Primer
	DateS24R	CGTTTTACGTTACCCTAACCAACGA	
51	DateS25F	**CACGACGTTGTAAAACGAC**GAGAGAGGAGTGGGAGAGGAGTTG	New Primer
	DateS25R	ATACGTAGTACGACGCCGTTCCTT	
52	DateS26F	**CACGACGTTGTAAAACGAC**GTCTACCACCGCAGGCTTGG	New Primer
	DateS26R	TTACGGGTTTACGGGTTACGGTTA	
53	DateS41F	**CACGACGTTGTAAAACGAC**GAGAGACCAGGCAGACATTCAACC	New Primer
	DateS41R	TGGTCTGTCCTTCTTATTCATCTTAAA	
54	DateS57F	**CACGACGTTGTAAAACGAC**TGTGGTTCAAAGGTGCCAGCTTAC	New Primer
	DateS57R	GTGTGTGTGTGTGTGTTTGGGTTG	
55	DateS60F	**CACGACGTTGTAAAACGAC**TCGGGAAAATAAGGGAAAGAAAGA	New Primer
	DateS60R	AGATGGTCTACAGTGCGGGTAAGG	
56	DateS78F	**CACGACGTTGTAAAACGAC**CATGGATTGGTAGCTTTTGTCCTCA	New Primer
	DateS78R	CGCTCCTTCATACTATGCCCCTAC	
57	DateS84F	**CACGACGTTGTAAAACGAC**GACCGGGACGATTCCAACAAC	New Primer
	DateS84R	GGCCTCCTTCCTTCCTCTTCCT	
58	DateS88F	**CACGACGTTGTAAAACGAC**ACTGTAGAATAACCCTTCTTTGATTT	New Primer
	DateS88R	TGCAATGGGGCATAGATATTG	
59	DateS90F	**CACGACGTTGTAAAACGAC**TCACCTATACTCTCTGCAAAACCA	New Primer
	DateS90R	GTCGGTCCATCGCCTACTT	
60	DateS100F	**CACGACGTTGTAAAACGAC**GGTGCCAGGTGTGACGTACT	New Primer
	DateS100R	AAGGGACAAGCCCACCTC	

61	DateS103F	**CACGACGTTGTAAAACGAC**GATGGTGATGGGAAGAGGAG	New Primer
	DateS103R	TCTCACGCTTCCATTGTTTG	
62	DateS110F	**CACGACGTTGTAAAACGAC**AAAGTGAGAACCTGAGGTGAGAG	New Primer
	DateS110R	CATGCAATGCACTGACAAAG	
63	DateS111F	**CACGACGTTGTAAAACGAC**CAAAGGACCTTAGCATATTCTTCTT	New Primer
	DateS111R	CTCTCGCTCGCTCGCTCT	
64	DateS116F	**CACGACGTTGTAAAACGAC**ACACCGAGTTCTTCCCAATG	New Primer
	DateS116R	CCTTGTAAACACCCAGCAAAA	
65	DateS120F	**CACGACGTTGTAAAACGAC**CTGGTGGCAGGGAGGATT	New Primer
	DateS120R	TCCTATCCTCGGTTTTGCAG	
66	DateS130F	**CACGACGTTGTAAAACGAC**TCATGGAAAACCCACCTCAT	New Primer
	DateS130R	AGTCGGTCAACTTGGATTGG	
67	DateS131F	**CACGACGTTGTAAAACGAC**TTTCCCTTGATAGGGATAAGC	New Primer
	DateS131R	ATCGGCATGTCCCATACTTC	
68	DateS137F	**CACGACGTTGTAAAACGAC**TACGGAATTCAGACCCCTCA	New Primer
	DateS137R	GCTTAGCCGAGGACTACTGC	
69	DateS138F	**CACGACGTTGTAAAACGAC**ACGGGAGATCCCTGATGC	New Primer
	DateS138R	GATCGGAGAAACGACCTCAC	
70	DateS176F	**CACGACGTTGTAAAACGAC**CCATCACTATCTCCACTATTGCTTT	New Primer
	DateS176R	AGATGCACTTAAGTCAGCCAAG	
71	DateS185F	**CACGACGTTGTAAAACGAC**CGAAGTTGAGCTCGTGAGAG	New Primer
	DateS185R	GCGAAGCACAACACCAGTAA	
72	DPALM301F	**CACGACGTTGTAAAACGAC**TTTTCCATCCCTTGACTTGG	Hamwieh *et al.*, (2010)
	DPALM301R	ACAGGCTGACGCACTTCTCT	
73	DPALM302F	**CACGACGTTGTAAAACGAC**AGCCAGATCATGGGAATGAG	Hamwieh *et al.*, (2010)
	DPALM302R	TGGATTCGTGCAAAGAATTG	
74	DPALM303F	**CACGACGTTGTAAAACGAC**TCCCCACTATGAGAAAGAACAAA	Hamwieh *et al.*, (2010)

	DPALM303R	CCCAACTACAAGCATCAGCA	
75	DPALM305F	**CACGACGTTGTAAAACGAC**TCCGAACTTGAATTCCCTCA	Hamwieh *et al.*, (2010)
	DPALM305R	GACCACCATCGTCATCATCA	
76	DPALM306F	**CACGACGTTGTAAAACGAC**ACACCCATTTCTGGAACAGC	Hamwieh *et al.*, (2010)
	DPALM306R	CAACACATTTTGGCAGCATT	
77	DPALM307F	**CACGACGTTGTAAAACGAC**CTACTTGAGCTGGGGTGGTC	Hamwieh *et al.*, (2010)
	DPALM307R	AACCTACGTGCCAATGGAAG	
78	DPALM308F	**CACGACGTTGTAAAACGAC**TTGGGTAAGATGGATGGTGAG	Hamwieh *et al.*, (2010)
	DPALM308R	GGTACATTGATTGGCAGCAA	
79	DPALM309F	**CACGACGTTGTAAAACGAC**AGCAAATGCTACTCGGGAAC	Hamwieh *et al.*, (2010)
	DPALM309R	TGGATCCATGGGGAGTGTAG	
80	DPALM310F	**CACGACGTTGTAAAACGAC**TGAAACTGCCAAACGATAAAGA	Hamwieh *et al.*, (2010)
	DPALM310R	AACCTCTCCGAGAAAACCAG	
81	DPALM311F	**CACGACGTTGTAAAACGAC**GAAAGGCTAGCCCCCATTAT	Hamwieh *et al.*, (2010)
	DPALM311R	CAATTGCGTGGAATCGACTA	
82	DPALM312F	**CACGACGTTGTAAAACGAC**CTGCGGATAAGGAATCTCCA	Hamwieh *et al.*, (2010)
	DPALM312R	GGGAGGCCTACCTCTAGCTC	
83	DPALM315F	**CACGACGTTGTAAAACGAC**TTGTTGTTGATGCTGCTGCT	Hamwieh *et al.*, (2010)
	DPALM315R	TGTTATCGGCAATTTGAAACC	
84	DPALM316F	**CACGACGTTGTAAAACGAC**GGTGTGATTCTCTCTCTTATTTGC	Hamwieh *et al.*, (2010)
	DPALM316R	TTGCAAGTTGAACAACACGA	
85	DPALM317F	**CACGACGTTGTAAAACGAC**CCTTCCTCAGTGATGGGCTA	Hamwieh *et al.*, (2010)
	DPALM317R	CCAAGAGGGAGAACTTGCAG	
86	DPALM318F	**CACGACGTTGTAAAACGAC**TTTGATGCAAGTAGGAAACCA	Hamwieh *et al.*, (2010)
	DPALM318R	CAAAAGTTAGTACTGCTGTTGTTGC	
87	DPALM319F	**CACGACGTTGTAAAACGAC**TCAGTCTAGCTGCTGACCTGTT	Hamwieh *et al.*, (2010)
	DPALM319R	TGCCTCGGGCTTCACTATAA	

88	DPALM320F	**CACGACGTTGTAAAACGAC**GTCTAGGGTGGCAAAACCAA	Hamwich *et al.*, (2010)
	DPALM320R	CAAATGGCTTCAATGCTCCT	
89	DPALM321F	**CACGACGTTGTAAAACGAC**GCATGTCCTGGGATCAGATT	Hamwich *et al.*, (2010)
	DPALM321R	GGCCACCTAATCATTTTTGG	
90	DPALM322F	**CACGACGTTGTAAAACGAC**TGGCATGGCAACTAATCAAA	Hamwich *et al.*, (2010)
	DPALM322R	TCATCTCGAATGCATCTGCT	
91	DPALM323F	**CACGACGTTGTAAAACGAC**CGACATTCCTGAAATTTGGA	Hamwich *et al.*, (2010)
	DPALM323R	CGCATTTAGTTGTCAAATCCTTC	
92	DPALM324F	**CACGACGTTGTAAAACGAC**TTCCTTCCCTCTCCAAACCT	Hamwich *et al.*, (2010)
	DPALM324R	AGGATCACTGCAACAATCACC	
93	DPALM325F	**CACGACGTTGTAAAACGAC**CAAGGGTCTTTGTGCAACCT	Hamwich *et al.*, (2010)
	DPALM325R	TTCCCAACCAGGGTAGTTCA	
94	DPALM326F	**CACGACGTTGTAAAACGAC**AAAAAGGATGGAGGCGAAAG	Hamwich *et al.*, (2010)
	DPALM326R	TGGCAAACAACTCTCCAAGA	
95	DPALM327F	**CACGACGTTGTAAAACGAC**TCCAACCTAGGCATGCAGAC	Hamwich *et al.*, (2010)
	DPALM327R	TGCGTTTACCATTTTGCTTG	
96	DPALM328F	**CACGACGTTGTAAAACGAC**TTCCATATGGTACATAGAGGACCTAA	Hamwich *et al.*, (2010)
	DPALM328R	TGGGCTAGTCCAGTAAAGCCTA	
97	DPALM329F	**CACGACGTTGTAAAACGAC**AAATGAGCCGCTTCTTTCAG	Hamwich *et al.*, (2010)
	DPALM329R	CAACCAAAGGATTGAATGGTG	
98	DPALM331F	**CACGACGTTGTAAAACGAC**GCGCAAGGCACAATTAAGAT	Hamwich *et al.*, (2010)
	DPALM331R	GAAGAAATGGAAACCCCAGA	
99	DPALM332F	**CACGACGTTGTAAAACGAC**ATTTGCTCCTGTCCTGCATC	Hamwich *et al.*, (2010)
	DPALM332R	AAACTCGAAGGCTTTGGTGA	
100	DPALM333F	**CACGACGTTGTAAAACGAC**TGGACAAAAGCAAAAGCCTAA	Hamwich *et al.*, (2010)
	DPALM333R	ATGAAACCAGTTGCCCAGTT	
101	DPALM335F	**CACGACGTTGTAAAACGAC**TGCTGAAACAAATTGATTTTGAC	Hamwich *et al.*, (2010)
	DPALM335R	TTAGGCAGGCAGCTGTTTTT	

102	DPALM336F	**CACGACGTTGTAAAACGAC**GTCAAAGATGGGCCAGAAAA	Hamwieh *et al.*, (2010)
	DPALM336R	TGCTGTGTTACAGTTGGAATCAT	
103	DPALM337F	**CACGACGTTGTAAAACGAC**GCCGCATACCCTTTTGTTAG	Hamwieh *et al.*, (2010)
	DPALM337R	TACATTGATTGGCAGCAACC	
104	DPALM338F	**CACGACGTTGTAAAACGAC**GCTGATAAAACAAGCTGGCAAT	Hamwieh *et al.*, (2010)
	DPALM338R	CAGAGAGAAAGCGTATTGGAGA	
105	DPALM339F	**CACGACGTTGTAAAACGAC**TGGTGGAATTAGCTCAAAGC	Hamwieh *et al.*, (2010)
	DPALM339R	CAAATCGATGATCCACACCA	
106	DPALM340F	**CACGACGTTGTAAAACGAC**CCCAAGCCTAACCTATCAGC	Hamwieh *et al.*, (2010)
	DPALM340R	ATTGCCTGCCACCAAGTATC	
107	DPALM341F	**CACGACGTTGTAAAACGAC**CCTCCTTCAAGTCACCATTCA	Hamwieh *et al.*, (2010)
	DPALM341R	TTGTTTTGCTGCTTCCATGT	
108	DPALM342F	**CACGACGTTGTAAAACGAC**GCAGTGCAACCCATTATCAA	Hamwieh *et al.*, (2010)
	DPALM342R	ATGGCATATGGTCCGAGTGT	
109	DPALM343F	**CACGACGTTGTAAAACGAC**TTCCGTGGCCTGTAATTTTT	Hamwieh *et al.*, (2010)
	DPALM343R	GCTGATTGTTGTTGTGATGAGC	
110	DPALM344F	**CACGACGTTGTAAAACGAC**GGCCTTTTGATGGTTGTGAG	Hamwieh *et al.*, (2010)
	DPALM344R	GATGATATGGGCTTGGCAAC	
111	DPALM345F	**CACGACGTTGTAAAACGAC**TGAATTTTGACCCCATGAAA	Hamwieh *et al.*, (2010)
	DPALM345R	TGACTGCAACCCAACATGTAA	
112	DPALM346F	**CACGACGTTGTAAAACGAC**ATGTGTGAGCCCAAACACTG	Hamwieh *et al.*, (2010)
	DPALM346R	TGGCTACTTTGATCCCATCC	
113	DPALM347F	**CACGACGTTGTAAAACGAC**GTTGCAGGGAGGTTTGTCAC	Hamwieh *et al.*, (2010)
	DPALM347R	TGTAGGCTTATTCCCATCCAA	
114	DPALM348F	**CACGACGTTGTAAAACGAC**TTCTTCCCATCTCCGAGAAA	Hamwieh *et al.*, (2010)
	DPALM348R	TTTGAGGGATTCTAAAAGGTGTTT	
115	DPALM349F	**CACGACGTTGTAAAACGAC**CTGATTGCCAGTCCAAGACA	Hamwieh *et al.*, (2010)
	DPALM349R	TGGGCAAACCTACTAAAATTGTG	

116	DPALM350F	**CACGACGTTGTAAAACGAC**AAGACCTCTTCGCAACTGGA	Hamwieh *et al.*, (2010)
	DPALM350R	TTCTCATGGAGTAGGATGGTCA	
117	DPALM351F	**CACGACGTTGTAAAACGAC**TGAGGATGTGATCCACATGAA	Hamwieh *et al.*, (2010)
	DPALM351R	TGAACGCACACAAGAATGAA	
118	DPALM352F	**CACGACGTTGTAAAACGAC**CCACCCCCATTAATTCCTCT	Hamwieh *et al.*, (2010)
	DPALM352R	TTCATATGGGATTGCGTGTG	
119	DPALM353F	**CACGACGTTGTAAAACGAC**TGGTTATGGTGGTGGTGATG	Hamwieh *et al.*, (2010)
	DPALM353R	TGTGATTTGCTTGCAATCCT	
120	DPALM354F	**CACGACGTTGTAAAACGAC**TGGTTCCGACCTGTTTCTTT	Hamwieh *et al.*, (2010)
	DPALM354R	CTTAACGCTCACCGCTCATT	
121	DPALM355F	**CACGACGTTGTAAAACGAC**TTTCGCTGCCTTAAAACCAT	Hamwieh *et al.*, (2010)
	DPALM355R	ACTTGCGCTGTTTGTTTCCT	
122	DPALM356F	**CACGACGTTGTAAAACGAC**AGTTTGTCAGGCCATTTGGT	Hamwieh *et al.*, (2010)
	DPALM356R	TACATGTGCGTATCGGGAGA	
123	DPALM357F	**CACGACGTTGTAAAACGAC**CGAATCCAACGAAGAGGAGT	Hamwieh *et al.*, (2010)
	DPALM357R	ATCATATTTGGCGCACCACT	
124	DPALM358F	**CACGACGTTGTAAAACGAC**CATCCGATGCTTGTAGCTGT	Hamwieh *et al.*, (2010)
	DPALM358R	TTGTTCCAGCTAGGCGGTAT	
125	DPALM359F	**CACGACGTTGTAAAACGAC**TCAATGCAGTATGCCTTCCA	Hamwieh *et al.*, (2010)
	DPALM359R	TCTGCTGCTCTTCCTCTCCT	
126	DPALM360F	**CACGACGTTGTAAAACGAC**GATTGGAGAGCGAGAACAGC	Hamwieh *et al.*, (2010)
	DPALM360R	GGTCGAGCTGTGGAAAGAGA	
127	DPALM361F	**CACGACGTTGTAAAACGAC**AACTGCAGTGAAGGCAACAA	Hamwieh *et al.*, (2010)
	DPALM361R	CGCCGTAATCCAGGTAAGG	
128	DPALM362F	**CACGACGTTGTAAAACGAC**CGACTTTGGTGGTCCTTGTT	Hamwieh *et al.*, (2010)
	DPALM362R	CAAGAGAGCGAGAGCGAGAG	
129	DPALM363F	**CACGACGTTGTAAAACGAC**GGGTGGGATCCCTTCTCTT	Hamwieh *et al.*, (2010)
	DPALM363R	TGTTACAAGGCCTGATGCAA	

130	DPALM364F	**CACGACGTTGTAAAACGAC**TTGCTCGTTTAGGTGATCCA	Hamwieh *et al.*, (2010)
	DPALM364R	GCATCACACCAAGGATGTTG	
131	DPALM365F	**CACGACGTTGTAAAACGAC**GCAATCAAGAACAAGGGTGAG	Hamwieh *et al.*, (2010)
	DPALM365R	CGAGAATTTTCGTTCCCAAA	
132	DPALM366F	**CACGACGTTGTAAAACGAC**CCAAGTGGTGAATGGAGAGC	Hamwieh *et al.*, (2010)
	DPALM366R	GACGCCCATATTGATGATGA	
133	DPALM367F	**CACGACGTTGTAAAACGAC**CAAAGGTGTGGGTTAGTAGGTTG	Hamwieh *et al.*, (2010)
	DPALM367R	GGTACATTGATTGGCAGCAA	
134	DPALM368F	**CACGACGTTGTAAAACGAC**TCAGCACCAAATAGCTGCAC	Hamwieh *et al.*, (2010)
	DPALM368R	TCCTATCCGTGGTGATGTGA	
135	DPALM369F	**CACGACGTTGTAAAACGAC**TGGTAGCTGTTGTGGCAAAG	Hamwieh *et al.*, (2010)
	DPALM369R	CAACCCGTCAAATCGTAAGG	
136	DPALM370F	**CACGACGTTGTAAAACGAC**CGGATGGTTCGTGAACTTTT	Hamwieh *et al.*, (2010)
	DPALM370R	TCGAGCGAGCCATCTAAAAT	
137	DPALM371F	**CACGACGTTGTAAAACGAC**CTTGATGATCGAAGGTGCAA	Hamwieh *et al.*, (2010)
	DPALM371R	TGAGGAACAAGAGCAAAAATTG	
138	DPALM372F	**CACGACGTTGTAAAACGAC**CGTCCTTGAAACTGTGACCA	Hamwieh *et al.*, (2010)
	DPALM372R	TCGGATGGCTTCTTTTTACC	
139	DPALM373F	**CACGACGTTGTAAAACGAC**AGGAAAGAGCAGACCAACCA	Hamwieh *et al.*, (2010)
	DPALM373R	CACCTCTCCGAGAAAACCAG	
140	DPALM374F	**CACGACGTTGTAAAACGAC**TAATGCAAGCGTCAGCTCCT	Hamwieh *et al.*, (2010)
	DPALM374R	GCCCATGAGCACAGAGATTT	
141	DPALM375F	**CACGACGTTGTAAAACGAC**TCCTCCCTGACTTGACCAAC	Hamwieh *et al.*, (2010)
	DPALM375R	TCGCAAGGTTTTCTTTCCTC	
142	DPALM376F	**CACGACGTTGTAAAACGAC**AAAAAGGCTGAAGGGGAAAG	Hamwieh *et al.*, (2010)
	DPALM376R	TGCAAATCTTGTCTGTTTCCA	
143	DPALM377F	**CACGACGTTGTAAAACGAC**GGAGGAGGTGGAAAAGGAAG	Hamwieh *et al.*, (2010)
	DPALM377R	CTGTGTGAAACAGGGGACCT	

144	DPALM378F	**CACGACGTTGTAAAACGAC**GGGGGCATTTTCAAAGAACT	Hamwieh *et al.*, (2010)
	DPALM378R	TCATGTTTAGGCCCTCCTTG	
145	DPALM379F	**CACGACGTTGTAAAACGAC**GGAACCCTGGGATAGCTGTT	Hamwieh *et al.*, (2010)
	DPALM379R	GAGGCCTACCCGTTTCCTAT	
146	DPALM380F	**CACGACGTTGTAAAACGAC**TGCATGATGGATGTCTTTGG	Hamwieh *et al.*, (2010)
	DPALM380R	TGTTTTCTTGGTTGCCCTTC	
147	DPALM381F	**CACGACGTTGTAAAACGAC**GCTTGCCTGCATCTCTTCTC	Hamwieh *et al.*, (2010)
	DPALM381R	TCCAGCAATCAGGAATGACA	
148	DPALM383F	**CACGACGTTGTAAAACGAC**ATCCGTCTCCTTCCCTTTTT	Hamwieh *et al.*, (2010)
	DPALM383R	TAGGCATAGGCGCCAGGT	
149	DPALM385F	**CACGACGTTGTAAAACGAC**GGTCCTCGGCAACTCAATTA	Hamwieh *et al.*, (2010)
	DPALM385R	TCTCAAGCCAAAGCAGGATT	
150	DPALM386F	**CACGACGTTGTAAAACGAC**ATCAATTTACCGACGGCATT	Hamwieh *et al.*, (2010)
	DPALM386R	CACCTCTCCGAGAAAACCAG	
151	DPALM387F	**CACGACGTTGTAAAACGAC**ACCGGAGCATAAAAGATCCA	Hamwieh *et al.*, (2010)
	DPALM387R	CAAGTGCTCACACTGGCAAT	
152	DPALM388F	**CACGACGTTGTAAAACGAC**AAAAAGGGGACCCACAAAAG	Hamwieh *et al.*, (2010)
	DPALM388R	GCAGGTTGCCGTTTTTGTAT	
153	DPALM389F	**CACGACGTTGTAAAACGAC**GCCTCATGCTCAAAAACCTC	Hamwieh *et al.*, (2010)
	DPALM389R	AGGTGGCTGCTGATCAAAAA	
154	DPALM390F	**CACGACGTTGTAAAACGAC**TTTCTGCTGATTGCTGTGTTG	Hamwieh *et al.*, (2010)
	DPALM390R	GGTACATTGATTGGCAGCAA	
155	DPALM391F	**CACGACGTTGTAAAACGAC**ATCGATGGATGGATGGATG	Hamwieh *et al.*, (2010)
	DPALM391R	TGGAGACCTATGCCTTATGC	
156	DPALM392F	**CACGACGTTGTAAAACGAC**CAAAAACCCGCTCCAATAAG	Hamwieh *et al.*, (2010)
	DPALM392R	ATCATCGGGATCCATTGAAG	
157	DPALM393F	**CACGACGTTGTAAAACGAC**CCCAAAGCAAGGATGAGGTA	Hamwieh *et al.*, (2010)
	DPALM393R	ATGCCATCTCCGTATTGAGG	

158	DPALM394F	**CACGACGTTGTAAAACGAC**TGCGTTATTGGTTCCTTTTCA	Hamwich *et al.*, (2010)
	DPALM394R	GTGCTCGTCCAATCCTAAGC	
159	DPALM395F	**CACGACGTTGTAAAACGAC**GGATGAAGGCAATCGAAAAA	Hamwich *et al.*, (2010)
	DPALM395R	CACCTCTCCGAGAAAACCAG	
160	DPALM397F	**CACGACGTTGTAAAACGAC**AATCCAAGGCTCAAAAGCAA	Hamwich *et al.*, (2010)
	DPALM397R	AAATTGGATTGGCTGCAGAG	
161	DPALM398F	**CACGACGTTGTAAAACGAC**TTCATCCTTCCTTCCTCTGTG	Hamwich *et al.*, (2010)
	DPALM398R	GCATTGATCGCATGAAAGAG	
162	DPALM399F	**CACGACGTTGTAAAACGAC**TGATTTGCTCCCTCTGTTCC	Hamwich *et al.*, (2010)
	DPALM399R	TCCCATCATGTTCGAAATCT	
163	DPALM400F	**CACGACGTTGTAAAACGAC**TGGGTATGGTAAGTGGAAGTCA	Hamwich *et al.*, (2010)
	DPALM400R	CCACCCTTCCATGCCTAAAT	
164	DPALM401F	**CACGACGTTGTAAAACGAC**CGTACCCCATAATGTGACACC	Hamwich *et al.*, (2010)
	DPALM401R	TCAAACGCATATGCTCGATT	
165	DPALM402F	**CACGACGTTGTAAAACGAC**GGAAATGATTTTGCGAGGAA	Hamwich *et al.*, (2010)
	DPALM402R	TACCGCACCATTTTTGAGTG	
166	DPALM403F	**CACGACGTTGTAAAACGAC**TGCCCAAGGACAAAGATTTC	Hamwich *et al.*, (2010)
	DPALM403R	TCTGCTGCTCTTCCTCTCCT	
167	DPALM404F	**CACGACGTTGTAAAACGAC**GACACGTTGACGATGTGGAA	Hamwich *et al.*, (2010)
	DPALM404R	CCATTGCTGTTGAGGAGGAG	
168	DPALM405F	**CACGACGTTGTAAAACGAC**GCATATCTTGCAGCTGAGCA	Hamwich *et al.*, (2010)
	DPALM405R	ATACCGCAAAAGCCAAAGAA	
169	DPALM407F	**CACGACGTTGTAAAACGAC**TTGCTGCATCATTGTCTGAA	Hamwich *et al.*, (2010)
	DPALM407R	GCAGCAACCCAACATGTAAA	
170	DPALM408F	**CACGACGTTGTAAAACGAC**CCATTAACAACCAGGCATCA	Hamwich *et al.*, (2010)
	DPALM408R	CGTGTGTGCAATGAGCGTAT	
171	DPALM410F	**CACGACGTTGTAAAACGAC**AGCCATGACAGCCAAAGAAG	Hamwich *et al.*, (2010)
	DPALM410R	GGTTTCTGCCACTTGGTGAG	

1. **5 *X* TBE / Tris Borate EDTA stock solution (5 litres)**

   Dissolve 54.0 g of Tris base [Tris (hydroxymethyl) aminomethane] in 4000 ml of distilled water. Add 27.5 g of boric acid and stir using a magnetic stirrer to dissolve. Add 20 mL of 0.5 M EDTA (Ethylenediamine tetra acetic acid) (pH 8.0) stock solution and adjust the volume to 5000 ml with distilled water. Store at room temperature.

   Dilutions needed for electrophoresis **0.5 *X* or 1.0 *X* TBE**

   0.5 *X* TBE Buffer: dilute 100 mL of 5 *X* TBE in 900mL Distilled water

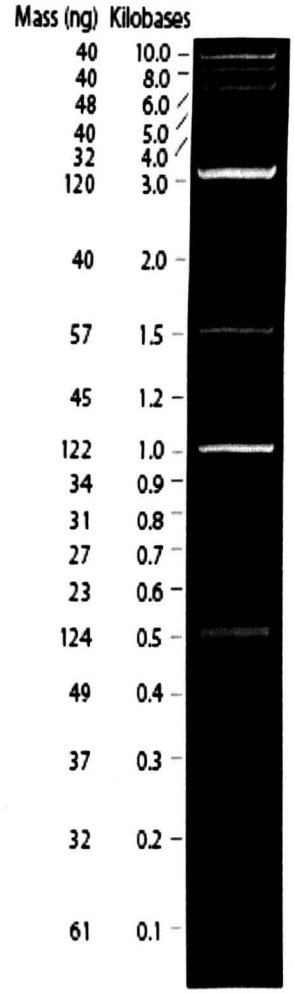
   1.0 *X* TBE Buffer: dilute 200 mL of 5 *X* TBE in 800mL Distilled water

2. **6 *X* agarose gel loading buffer (30 ml)**

   Pour ~ 9 mL of glycerol into a screw top bottle and add a pinch of bromophenol blue and a pinch of xylene cyanol FF. Make up to 30 ml with distilled water. Mix. Store at room temperature. The loading buffer was aliquoted into 1.5 mL Eppendroph tubes.

Appendix 3: Table of 2-log DNA Ladder, their fragment and size in (bp)

Fragment	Size (bp)	Mass(ng)
1	10,002	**40**
2	8,001	40
3	6,001	48
4	5,001	40
5	4,001	32
6	**3,001**	**120**
7	2,017	40
8	1,517	57
9	1,200	45
10	**1,000**	**122**
11	900	34
12	800	31
13	700	27
14	600	23
15a	**517**	**124**
15b	**500**	**124**
16	400	49
17	300	37
18	200	32
19	100	61



Appendix 4: Major allele frequency, number of genotypes showing polymorphism, number of allele generated, heterozygosity and polymorphism information content (PIC) for eight parents of Omani date palm estimated by 72 SSR markers produced by Billotte *et al.* (2004), Akkak *et al.* (2009), ourselves and Hamwieh *et al.* (2010)

Marker	Major allele frequency	Genotype no.	Allele no.	Gene diversity	Heterozygosity	PIC
mPdCIR010	0.31	7	9	0.83	0.75	0.81
mPdCIR015	0.38	6	7	0.78	1.00	0.75
mPdCIR016	0.31	5	6	0.79	0.88	0.76
mPdCIR025	0.44	5	5	0.70	0.50	0.66
mPdCIR050	0.44	7	7	0.73	0.63	0.69
mPdCIR057	0.50	4	5	0.68	0.50	0.64
mPdCIR078	0.25	7	9	0.84	0.63	0.83
mPdCIR085	0.19	8	9	0.87	0.50	0.85
mPdCIR093	0.31	6	6	0.78	0.25	0.75
PDCAT2	0.38	7	10	0.81	0.75	0.80
PDCAT5	0.25	7	8	0.83	0.38	0.81
PDCAT10	0.63	2	2	0.47	0.00	0.36
PDCAT11	0.19	7	10	0.88	1.00	0.87
PDCAT12	0.50	5	5	0.66	0.25	0.62
PDCAT14	0.19	8	9	0.87	0.75	0.85
PDCAT17	0.31	7	7	0.80	0.63	0.78
PDCAT18	0.19	7	10	0.88	0.75	0.86
PDCAT20	0.38	7	7	0.78	0.63	0.75
PDCAT21	0.25	6	6	0.80	0.63	0.77
DateS1	0.75	3	2	0.38	0.25	0.30
DateS8	0.75	4	3	0.40	0.25	0.35
DateS9	0.81	3	4	0.33	0.25	0.31
DateS12	0.75	4	4	0.41	0.25	0.39
DateS16	0.81	4	4	0.33	0.38	0.31
DateS17	0.63	5	4	0.54	0.38	0.48
DateS103	0.50	3	3	0.59	0.25	0.51
DateS110	0.38	6	5	0.75	0.88	0.71
DateS111	0.38	5	6	0.73	0.25	0.69
DateS130	0.69	2	2	0.43	0.63	0.34
DateS131	0.63	5	5	0.57	0.38	0.54
DPALM302	0.63	3	3	0.53	0.00	0.47
DPALM303	0.38	6	5	0.74	0.75	0.70
DPALM305	0.56	3	3	0.54	0.13	0.45
DPALM307	0.63	3	3	0.53	0.00	0.47
DPALM309	0.69	3	3	0.46	0.13	0.40
DPALM311	0.50	5	4	0.66	0.13	0.62
DPALM312	0.25	5	5	0.78	0.75	0.75
DPALM315	0.63	4	3	0.53	0.25	0.47
DPALM319	0.63	4	4	0.55	0.38	0.51
DPALM325	0.50	5	5	0.66	0.25	0.62
DPALM327	0.38	6	6	0.75	0.88	0.71
DPALM328	0.44	7	8	0.75	0.88	0.73
DPALM332	0.50	4	4	0.66	0.50	0.60
DPALM333	0.25	8	7	0.83	0.63	0.81
DPALM336	0.31	5	6	0.78	0.88	0.75
DPALM340	0.63	3	3	0.53	0.50	0.47
DPALM341	0.31	7	8	0.80	0.88	0.78
DPALM342	0.81	2	2	0.30	0.38	0.26
DPALM343	0.50	3	3	0.59	0.25	0.51

DPALM344	0.88	2	2	0.22	0.25	0.19
DPALM348	0.56	5	3	0.59	0.38	0.52
DPALM349	0.50	5	5	0.68	0.50	0.64
DPALM350	0.31	7	7	0.81	0.75	0.79
DPALM352	0.50	3	2	0.50	0.75	0.38
DPALM357	0.50	4	4	0.66	0.25	0.60
DPALM361	0.31	7	7	0.77	0.75	0.74
DPALM362	0.25	6	7	0.82	0.38	0.80
DPALM363	0.38	4	4	0.73	0.63	0.68
DPALM366	0.31	6	5	0.75	0.25	0.71
DPALM369	0.44	7	8	0.76	0.75	0.74
DPALM374	0.50	4	3	0.55	0.75	0.46
DPALM377	0.25	5	5	0.78	0.00	0.75
DPALM378	0.31	7	6	0.78	0.50	0.75
DPALM379	0.38	5	5	0.73	0.25	0.68
DPALM380	0.69	4	4	0.49	0.13	0.46
DPALM388	0.38	4	3	0.66	0.38	0.59
DPALM398	0.50	4	5	0.64	0.25	0.58
DPALM402	0.38	6	5	0.74	0.50	0.70
DPALM404	0.50	4	4	0.60	0.25	0.53
DPALM405	0.63	3	3	0.53	0.25	0.47
DPALM408	0.56	3	3	0.54	0.13	0.45
DPALM410	0.38	5	5	0.73	0.13	0.68
Mean	0.46	4.97	5.13	0.66	0.46	0.61

Appendix 5: The locus genotype frequency for BC_1 map of phase determination

S/n	Locus	Segregation	ac	ad	bc	bd	ee	ef	eg	fg	hh	hk	kk	h-	k-	ll	lm	nn	np	--	X2	Df	Signif.	Classification
1	Locus001	<efxeg>	0	0	0	0	17	10	12	14	0	0	0	0	0	0	0	0	0	0	2.02	3	-	[ee:ef:eg:fg]
2	Locus002	<hkxhk>	0	0	0	0	0	0	0	0	13	25	15	0	0	0	0	0	0	0	0.32	2	-	[hh:hk:kk]
3	Locus003	<hkxhk>	0	0	0	0	0	0	0	0	11	30	12	0	0	0	0	0	0	0	0.96	2	-	[hh:hk:kk]
4	Locus007	<lmxll>	0	0	0	0	0	0	0	0	0	0	0	0	0	48	5	0	0	0	34.9	1	*******	[ll:lm]
5	Locus008	<hkxhk>	0	0	0	0	0	0	0	0	12	26	12	0	0	0	0	0	0	3	0.08	2	-	[hh:hk:kk]
6	Locus010	<efxeg>	0	0	0	0	5	19	14	13	0	0	0	0	0	0	0	0	0	2	7.9	3	**	[ee:ef:eg:fg]
7	Locus015	<efxeg>	0	0	0	0	14	10	16	10	0	0	0	0	0	0	0	0	0	3	2.16	3	-	[ee:ef:eg:fg]
8	Locus022	<efxeg>	0	0	0	0	8	9	13	23	0	0	0	0	0	0	0	0	0	0	10.6	3	**	[ee:ef:eg:fg]
9	Locus025	<efxeg>	0	0	0	0	10	8	20	14	0	0	0	0	0	0	0	0	0	1	6.46	3	*	[ee:ef:eg:fg]
10	Locus027	<nnxnp>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	33	0	3.19	1	*	[nn:np]
11	Locus028	<efxeg>	0	0	0	0	12	8	25	8	0	0	0	0	0	0	0	0	0	0	14.7	3	****	[ee:ef:eg:fg]
12	Locus029	<lmxll>	0	0	0	0	0	0	0	0	0	0	0	0	0	24	26	0	0	3	0.08	1	-	[ll:lm]
13	Locus030	<nnxnp>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	25	26	2	0.02	1	-	[nn:np]
14	Locus031	<nnxnp>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	28	24	1	0.31	1	-	[nn:np]
15	Locus035	<efxeg>	0	0	0	0	10	17	13	12	0	0	0	0	0	0	0	0	0	1	2	3	-	[ee:ef:eg:fg]
16	Locus039	<nnxnp>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	28	25	0	0.17	1	-	[nn:np]
17	Locus042	<hkxhk>	0	0	0	0	0	0	0	0	12	25	16	0	0	0	0	0	0	0	0.77	2	-	[hh:hk:kk]
18	Locus047	<hkxhk>	0	0	0	0	0	0	0	0	15	21	16	0	0	0	0	0	0	1	1.96	2	-	[hh:hk:kk]
19	Locus054	<nnxnp>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	23	30	0	0.92	1	-	[nn:np]
20	Locus055	<nnxnp>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	22	31	0	1.53	1	-	[nn:np]
21	Locus056	<hkxhk>	0	0	0	0	0	0	0	0	14	22	17	0	0	0	0	0	0	0	1.87	2	-	[hh:hk:kk]
22	Locus068	<efxeg>	0	0	0	0	13	12	17	7	0	0	0	0	0	0	0	0	0	4	4.14	3	-	[ee:ef:eg:fg]
23	Locus069	<efxeg>	0	0	0	0	18	8	15	11	0	0	0	0	0	0	0	0	0	1	4.46	3	-	[ee:ef:eg:fg]
24	Locus072	<hkxhk>	0	0	0	0	0	0	0	0	11	29	12	0	0	0	0	0	0	1	0.73	2	-	[hh:hk:kk]
25	Locus073	<lmxll>	0	0	0	0	0	0	0	0	0	0	0	0	0	21	31	0	0	1	1.92	1	-	[ll:lm]
26	Locus079	<hkxhk>	0	0	0	0	0	0	0	0	14	26	13	0	0	0	0	0	0	0	0.06	2	-	[hh:hk:kk]

27	Locus080	<nnxnp>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	25	23	5	0.08	1	-	[nn:np]
28	Locus081	<hkxhk>	0	0	0	0	0	0	0	0	0	53	0	0	0	0	0	0	0	0	53	2	*******	[hh:hk:kk]
29	Locus087	<hkxhk>	0	0	0	0	0	0	0	0	13	30	10	0	0	0	0	0	0	0	1.26	2	-	[hh:hk:kk]
30	Locus088	<hkxhk>	0	0	0	0	0	0	0	0	14	30	9	0	0	0	0	0	0	0	1.87	2	-	[hh:hk:kk]
31	Locus089	<hkxhk>	0	0	0	0	0	0	0	0	13	29	11	0	0	0	0	0	0	0	0.62	2	-	[hh:hk:kk]
32	Locus091	<nnxnp>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	33	0	3.19	1	*	[nn:np]
33	Locus096	<nnxnp>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	31	2	2.37	1	-	[nn:np]
34	Locus100	<efxeg>	0	0	0	0	11	12	12	13	0	0	0	0	0	0	0	0	0	5	0.17	3	-	[ee:ef:eg:fg]
35	Locus101	<hkxhk>	0	0	0	0	0	0	0	0	12	26	12	0	0	0	0	0	0	3	0.08	2	-	[hh:hk:kk]
36	Locus102	<hkxhk>	0	0	0	0	0	0	0	0	0	53	0	0	0	0	0	0	0	0	53	2	*******	[hh:hk:kk]
37	Locus108	<nnxnp>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	29	23	1	0.69	1	-	[nn:np]
38	Locus113	<hkxhk>	0	0	0	0	0	0	0	0	15	22	16	0	0	0	0	0	0	0	1.57	2	-	[hh:hk:kk]
39	Locus118	<nnxnp>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	23	27	3	0.32	1	-	[nn:np]
40	Locus138	<hkxhk>	0	0	0	0	0	0	0	0	12	30	10	0	0	0	0	0	0	1	1.38	2	-	[hh:hk:kk]
41	Locus141	<hkxhk>	0	0	0	0	0	0	0	0	14	28	11	0	0	0	0	0	0	0	0.51	2	-	[hh:hk:kk]
42	Locus145	<nnxnp>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	26	27	0	0.02	1	-	[nn:np]
43	100003487	<hkxhk>	0	0	0	0	0	0	0	0	13	27	13	0	0	0	0	0	0	0	0.02	2	-	[hh:hk:kk]
44	100006658	<hkxhk>	0	0	0	0	0	0	0	0	12	28	13	0	0	0	0	0	0	0	0.21	2	-	[hh:hk:kk]
45	100006666	<hkxhk>	0	0	0	0	0	0	0	0	11	30	12	0	0	0	0	0	0	0	0.96	2	-	[hh:hk:kk]
46	100006795	<hkxhk>	0	0	0	0	0	0	0	0	11	31	11	0	0	0	0	0	0	0	1.53	2	-	[hh:hk:kk]
47	100007066	<hkxhk>	0	0	0	0	0	0	0	0	18	11	24	0	0	0	0	0	0	0	19.5	2	*******	[hh:hk:kk]
48	100007118	<hkxhk>	0	0	0	0	0	0	0	0	15	27	11	0	0	0	0	0	0	0	0.62	2	-	[hh:hk:kk]
49	100007234	<hkxhk>	0	0	0	0	0	0	0	0	11	30	12	0	0	0	0	0	0	0	0.96	2	-	[hh:hk:kk]
50	100007720	<hkxhk>	0	0	0	0	0	0	0	0	11	30	12	0	0	0	0	0	0	0	0.96	2	-	[hh:hk:kk]
51	100007831	<hkxhk>	0	0	0	0	0	0	0	0	12	20	21	0	0	0	0	0	0	0	6.25	2	**	[hh:hk:kk]
52	100008469	<hkxhk>	0	0	0	0	0	0	0	0	14	25	14	0	0	0	0	0	0	0	0.17	2	-	[hh:hk:kk]
53	100008834	<hkxhk>	0	0	0	0	0	0	0	0	13	11	29	0	0	0	0	0	0	0	27.8	2	*******	[hh:hk:kk]
54	100009021	<hkxhk>	0	0	0	0	0	0	0	0	16	23	14	0	0	0	0	0	0	0	1.08	2	-	[hh:hk:kk]

55	100009038	<hkxhk>	0	0	0	0	0	0	0	0	11	31	11	0	0	0	0	0	0	0	1.53	2	-	[hh:hk:kk]
56	100009167	<hkxhk>	0	0	0	0	0	0	0	0	12	18	23	0	0	0	0	0	0	0	10	2	***	[hh:hk:kk]
57	100009495	<hkxhk>	0	0	0	0	0	0	0	0	16	15	22	0	0	0	0	0	0	0	11.3	2	****	[hh:hk:kk]
58	100009499	<hkxhk>	0	0	0	0	0	0	0	0	11	28	14	0	0	0	0	0	0	0	0.51	2	-	[hh:hk:kk]
59	100009512	<hkxhk>	0	0	0	0	0	0	0	0	15	21	17	0	0	0	0	0	0	0	2.43	2	-	[hh:hk:kk]
60	100009895	<hkxhk>	0	0	0	0	0	0	0	0	14	26	13	0	0	0	0	0	0	0	0.06	2	-	[hh:hk:kk]
61	100010290	<hkxhk>	0	0	0	0	0	0	0	0	11	19	23	0	0	0	0	0	0	0	9.68	2	***	[hh:hk:kk]
62	100010318	<hkxhk>	0	0	0	0	0	0	0	0	12	26	15	0	0	0	0	0	0	0	0.36	2	-	[hh:hk:kk]
63	100011533	<hkxhk>	0	0	0	0	0	0	0	0	12	15	26	0	0	0	0	0	0	0	17.4	2	******	[hh:hk:kk]
64	100011759	<hkxhk>	0	0	0	0	0	0	0	0	13	28	12	0	0	0	0	0	0	0	0.21	2	-	[hh:hk:kk]
65	100012164	<hkxhk>	0	0	0	0	0	0	0	0	14	23	16	0	0	0	0	0	0	0	1.08	2	-	[hh:hk:kk]
66	100012300	<hkxhk>	0	0	0	0	0	0	0	0	12	27	14	0	0	0	0	0	0	0	0.17	2	-	[hh:hk:kk]
67	100012305	<hkxhk>	0	0	0	0	0	0	0	0	12	25	16	0	0	0	0	0	0	0	0.77	2	-	[hh:hk:kk]
68	100012437	<hkxhk>	0	0	0	0	0	0	0	0	11	29	13	0	0	0	0	0	0	0	0.62	2	-	[hh:hk:kk]
69	100012477	<hkxhk>	0	0	0	0	0	0	0	0	12	13	28	0	0	0	0	0	0	0	23.4	2	*******	[hh:hk:kk]
70	100012497	<hkxhk>	0	0	0	0	0	0	0	0	13	21	19	0	0	0	0	0	0	0	3.64	2	-	[hh:hk:kk]
71	100012971	<hkxhk>	0	0	0	0	0	0	0	0	18	13	22	0	0	0	0	0	0	0	14.4	2	*****	[hh:hk:kk]
72	100013194	<hkxhk>	0	0	0	0	0	0	0	0	20	22	11	0	0	0	0	0	0	0	4.58	2	-	[hh:hk:kk]
73	100013212	<hkxhk>	0	0	0	0	0	0	0	0	11	22	20	0	0	0	0	0	0	0	4.58	2	-	[hh:hk:kk]
74	100013314	<hkxhk>	0	0	0	0	0	0	0	0	12	27	14	0	0	0	0	0	0	0	0.17	2	-	[hh:hk:kk]
75	100013350	<hkxhk>	0	0	0	0	0	0	0	0	17	17	19	0	0	0	0	0	0	0	6.96	2	**	[hh:hk:kk]
76	100013446	<hkxhk>	0	0	0	0	0	0	0	0	14	8	31	0	0	0	0	0	0	0	36.7	2	*******	[hh:hk:kk]
77	100013515	<hkxhk>	0	0	0	0	0	0	0	0	16	12	25	0	0	0	0	0	0	0	18.9	2	*******	[hh:hk:kk]
78	100013538	<hkxhk>	0	0	0	0	0	0	0	0	13	28	12	0	0	0	0	0	0	0	0.21	2	-	[hh:hk:kk]
79	100013911	<hkxhk>	0	0	0	0	0	0	0	0	13	24	16	0	0	0	0	0	0	0	0.81	2	-	[hh:hk:kk]
80	100014713	<hkxhk>	0	0	0	0	0	0	0	0	14	26	13	0	0	0	0	0	0	0	0.06	2	-	[hh:hk:kk]
81	100014845	<hkxhk>	0	0	0	0	0	0	0	0	12	7	34	0	0	0	0	0	0	0	47	2	*******	[hh:hk:kk]
82	100014927	<hkxhk>	0	0	0	0	0	0	0	0	15	22	16	0	0	0	0	0	0	0	1.57	2	-	[hh:hk:kk]

83	100015250	<hkxhk>	0	0	0	0	0	0	0	0	14	24	15	0	0	0	0	0	0	0	0.51	2	-	[hh:hk:kk]
84	100015461	<hkxhk>	0	0	0	0	0	0	0	0	15	18	20	0	0	0	0	0	0	0	6.4	2	**	[hh:hk:kk]
85	100015635	<hkxhk>	0	0	0	0	0	0	0	0	12	23	18	0	0	0	0	0	0	0	2.28	2	-	[hh:hk:kk]
86	100015698	<hkxhk>	0	0	0	0	0	0	0	0	16	21	16	0	0	0	0	0	0	0	2.28	2	-	[hh:hk:kk]
87	100015788	<hkxhk>	0	0	0	0	0	0	0	0	13	15	25	0	0	0	0	0	0	0	15.4	2	******	[hh:hk:kk]
88	100015926	<hkxhk>	0	0	0	0	0	0	0	0	11	30	12	0	0	0	0	0	0	0	0.96	2	-	[hh:hk:kk]
89	100016039	<hkxhk>	0	0	0	0	0	0	0	0	12	20	21	0	0	0	0	0	0	0	6.25	2	**	[hh:hk:kk]
90	100016643	<hkxhk>	0	0	0	0	0	0	0	0	25	12	16	0	0	0	0	0	0	0	18.9	2	*******	[hh:hk:kk]
91	100016828	<hkxhk>	0	0	0	0	0	0	0	0	11	15	27	0	0	0	0	0	0	0	19.6	2	*******	[hh:hk:kk]
92	100016873	<hkxhk>	0	0	0	0	0	0	0	0	11	29	13	0	0	0	0	0	0	0	0.62	2	-	[hh:hk:kk]
93	100017033	<hkxhk>	0	0	0	0	0	0	0	0	11	29	13	0	0	0	0	0	0	0	0.62	2	-	[hh:hk:kk]
94	100017048	<hkxhk>	0	0	0	0	0	0	0	0	11	10	32	0	0	0	0	0	0	0	37.2	2	*******	[hh:hk:kk]
95	100017066	<hkxhk>	0	0	0	0	0	0	0	0	16	24	13	0	0	0	0	0	0	0	0.81	2	-	[hh:hk:kk]
96	100017275	<hkxhk>	0	0	0	0	0	0	0	0	11	30	12	0	0	0	0	0	0	0	0.96	2	-	[hh:hk:kk]
97	100017550	<hkxhk>	0	0	0	0	0	0	0	0	15	18	20	0	0	0	0	0	0	0	6.4	2	**	[hh:hk:kk]
98	100017722	<hkxhk>	0	0	0	0	0	0	0	0	16	15	22	0	0	0	0	0	0	0	11.3	2	****	[hh:hk:kk]
99	100017754	<hkxhk>	0	0	0	0	0	0	0	0	23	16	14	0	0	0	0	0	0	0	11.4	2	****	[hh:hk:kk]
100	100018039	<hkxhk>	0	0	0	0	0	0	0	0	11	28	14	0	0	0	0	0	0	0	0.51	2	-	[hh:hk:kk]
101	100018090	<hkxhk>	0	0	0	0	0	0	0	0	14	24	15	0	0	0	0	0	0	0	0.51	2	-	[hh:hk:kk]
102	100018230	<hkxhk>	0	0	0	0	0	0	0	0	14	27	12	0	0	0	0	0	0	0	0.17	2	-	[hh:hk:kk]
103	100018282	<hkxhk>	0	0	0	0	0	0	0	0	14	21	18	0	0	0	0	0	0	0	2.89	2	-	[hh:hk:kk]
104	100018479	<hkxhk>	0	0	0	0	0	0	0	0	11	14	28	0	0	0	0	0	0	0	22.7	2	*******	[hh:hk:kk]
105	100018480	<hkxhk>	0	0	0	0	0	0	0	0	12	30	11	0	0	0	0	0	0	0	0.96	2	-	[hh:hk:kk]
106	100018657	<hkxhk>	0	0	0	0	0	0	0	0	14	27	12	0	0	0	0	0	0	0	0.17	2	-	[hh:hk:kk]
107	100018989	<hkxhk>	0	0	0	0	0	0	0	0	15	21	17	0	0	0	0	0	0	0	2.43	2	-	[hh:hk:kk]
108	100019268	<hkxhk>	0	0	0	0	0	0	0	0	15	27	11	0	0	0	0	0	0	0	0.62	2	-	[hh:hk:kk]
109	100019408	<hkxhk>	0	0	0	0	0	0	0	0	11	31	11	0	0	0	0	0	0	0	1.53	2	-	[hh:hk:kk]
110	100019455	<hkxhk>	0	0	0	0	0	0	0	0	17	12	24	0	0	0	0	0	0	0	17.7	2	******	[hh:hk:kk]

111	100019480	<hkxhk>	0	0	0	0	0	0	0	0	16	24	13	0	0	0	0	0	0	0	0.81	2	-	[hh:hk:kk]
112	100019500	<hkxhk>	0	0	0	0	0	0	0	0	13	25	15	0	0	0	0	0	0	0	0.32	2	-	[hh:hk:kk]
113	100019701	<hkxhk>	0	0	0	0	0	0	0	0	12	13	28	0	0	0	0	0	0	0	23.4	2	*******	[hh:hk:kk]
114	100019743	<hkxhk>	0	0	0	0	0	0	0	0	17	21	15	0	0	0	0	0	0	0	2.43	2	-	[hh:hk:kk]
115	100019834	<hkxhk>	0	0	0	0	0	0	0	0	11	29	13	0	0	0	0	0	0	0	0.62	2	-	[hh:hk:kk]
116	100019906	<hkxhk>	0	0	0	0	0	0	0	0	12	27	14	0	0	0	0	0	0	0	0.17	2	-	[hh:hk:kk]
117	100020136	<hkxhk>	0	0	0	0	0	0	0	0	16	23	14	0	0	0	0	0	0	0	1.08	2	-	[hh:hk:kk]
118	100020254	<hkxhk>	0	0	0	0	0	0	0	0	15	13	25	0	0	0	0	0	0	0	17.5	2	******	[hh:hk:kk]
119	100020443	<hkxhk>	0	0	0	0	0	0	0	0	11	14	28	0	0	0	0	0	0	0	22.7	2	*******	[hh:hk:kk]
120	100020510	<hkxhk>	0	0	0	0	0	0	0	0	20	7	26	0	0	0	0	0	0	0	30.1	2	*******	[hh:hk:kk]
121	100020651	<hkxhk>	0	0	0	0	0	0	0	0	12	25	16	0	0	0	0	0	0	0	0.77	2	-	[hh:hk:kk]
122	100020768	<hkxhk>	0	0	0	0	0	0	0	0	12	28	13	0	0	0	0	0	0	0	0.21	2	-	[hh:hk:kk]
123	100021135	<hkxhk>	0	0	0	0	0	0	0	0	11	28	14	0	0	0	0	0	0	0	0.51	2	-	[hh:hk:kk]
124	100021260	<hkxhk>	0	0	0	0	0	0	0	0	12	28	13	0	0	0	0	0	0	0	0.21	2	-	[hh:hk:kk]
125	100021433	<hkxhk>	0	0	0	0	0	0	0	0	13	27	13	0	0	0	0	0	0	0	0.02	2	-	[hh:hk:kk]
126	100021788	<hkxhk>	0	0	0	0	0	0	0	0	11	29	13	0	0	0	0	0	0	0	0.62	2	-	[hh:hk:kk]
127	100022371	<hkxhk>	0	0	0	0	0	0	0	0	18	11	24	0	0	0	0	0	0	0	19.5	2	*******	[hh:hk:kk]
128	100022382	<hkxhk>	0	0	0	0	0	0	0	0	11	29	13	0	0	0	0	0	0	0	0.62	2	-	[hh:hk:kk]
129	100022444	<hkxhk>	0	0	0	0	0	0	0	0	12	30	11	0	0	0	0	0	0	0	0.96	2	-	[hh:hk:kk]
130	100022789	<hkxhk>	0	0	0	0	0	0	0	0	13	7	13	0	0	0	0	0	0	20	10.9	2	****	[hh:hk:kk]
131	100023222	<hkxhk>	0	0	0	0	0	0	0	0	28	13	12	0	0	0	0	0	0	0	23.4	2	*******	[hh:hk:kk]
132	100023274	<hkxhk>	0	0	0	0	0	0	0	0	30	9	12	0	0	0	0	0	0	2	34.1	2	*******	[hh:hk:kk]
133	100023720	<hkxhk>	0	0	0	0	0	0	0	0	12	22	19	0	0	0	0	0	0	0	3.38	2	-	[hh:hk:kk]
134	100023875	<hkxhk>	0	0	0	0	0	0	0	0	11	28	14	0	0	0	0	0	0	0	0.51	2	-	[hh:hk:kk]
135	100024229	<hkxhk>	0	0	0	0	0	0	0	0	15	15	23	0	0	0	0	0	0	0	12.4	2	****	[hh:hk:kk]
136	100024243	<hkxhk>	0	0	0	0	0	0	0	0	20	20	13	0	0	0	0	0	0	0	5.04	2	*	[hh:hk:kk]
137	100024252	<hkxhk>	0	0	0	0	0	0	0	0	14	16	23	0	0	0	0	0	0	0	11.4	2	****	[hh:hk:kk]
138	100024282	<hkxhk>	0	0	0	0	0	0	0	0	14	26	13	0	0	0	0	0	0	0	0.06	2	-	[hh:hk:kk]

139	1000024426	<hkxhk>	0	0	0	0	0	0	0	0	14	20	19	0	0	0	0	0	0	0	4.13	2	-	[hh:hk:kk]
140	1000024509	<hkxhk>	0	0	0	0	0	0	0	0	18	13	22	0	0	0	0	0	0	0	14.4	2	*****	[hh:hk:kk]
141	1000024877	<hkxhk>	0	0	0	0	0	0	0	0	11	15	27	0	0	0	0	0	0	0	19.6	2	*******	[hh:hk:kk]
142	1000025098	<hkxhk>	0	0	0	0	0	0	0	0	8	11	14	0	0	0	0	0	0	20	5.88	2	*	[hh:hk:kk]
143	1000025241	<hkxhk>	0	0	0	0	0	0	0	0	14	20	19	0	0	0	0	0	0	0	4.13	2	-	[hh:hk:kk]
144	1000025359	<hkxhk>	0	0	0	0	0	0	0	0	13	12	28	0	0	0	0	0	0	0	24.4	2	*******	[hh:hk:kk]
145	1000025439	<hkxhk>	0	0	0	0	0	0	0	0	17	12	24	0	0	0	0	0	0	0	17.7	2	******	[hh:hk:kk]
146	1000025464	<hkxhk>	0	0	0	0	0	0	0	0	19	13	21	0	0	0	0	0	0	0	13.9	2	*****	[hh:hk:kk]
147	1000025632	<hkxhk>	0	0	0	0	0	0	0	0	12	14	27	0	0	0	0	0	0	0	20.3	2	*******	[hh:hk:kk]
148	1000025769	<hkxhk>	0	0	0	0	0	0	0	0	18	22	13	0	0	0	0	0	0	0	2.47	2	-	[hh:hk:kk]
149	1000025859	<hkxhk>	0	0	0	0	0	0	0	0	19	7	27	0	0	0	0	0	0	0	31.1	2	*******	[hh:hk:kk]
150	1000025975	<hkxhk>	0	0	0	0	0	0	0	0	11	28	14	0	0	0	0	0	0	0	0.51	2	-	[hh:hk:kk]
151	1000026244	<hkxhk>	0	0	0	0	0	0	0	0	13	27	13	0	0	0	0	0	0	0	0.02	2	-	[hh:hk:kk]
152	1000026315	<hkxhk>	0	0	0	0	0	0	0	0	13	13	27	0	0	0	0	0	0	0	21.2	2	*******	[hh:hk:kk]
153	1000026827	<hkxhk>	0	0	0	0	0	0	0	0	12	22	19	0	0	0	0	0	0	0	3.38	2	-	[hh:hk:kk]
154	1000026920	<hkxhk>	0	0	0	0	0	0	0	0	14	24	15	0	0	0	0	0	0	0	0.51	2	-	[hh:hk:kk]
155	1000026967	<hkxhk>	0	0	0	0	0	0	0	0	13	24	16	0	0	0	0	0	0	0	0.81	2	-	[hh:hk:kk]
156	1000026982	<hkxhk>	0	0	0	0	0	0	0	0	11	20	22	0	0	0	0	0	0	0	7.75	2	**	[hh:hk:kk]
157	1000027048	<hkxhk>	0	0	0	0	0	0	0	0	13	17	23	0	0	0	0	0	0	0	10.6	2	***	[hh:hk:kk]
158	1000027130	<hkxhk>	0	0	0	0	0	0	0	0	14	14	25	0	0	0	0	0	0	0	16.4	2	******	[hh:hk:kk]
159	1000027270	<hkxhk>	0	0	0	0	0	0	0	0	18	23	12	0	0	0	0	0	0	0	2.28	2	-	[hh:hk:kk]
160	1000027553	<hkxhk>	0	0	0	0	0	0	0	0	11	22	20	0	0	0	0	0	0	0	4.58	2	-	[hh:hk:kk]
161	1000027989	<hkxhk>	0	0	0	0	0	0	0	0	11	28	14	0	0	0	0	0	0	0	0.51	2	-	[hh:hk:kk]
162	1000028019	<hkxhk>	0	0	0	0	0	0	0	0	11	30	12	0	0	0	0	0	0	0	0.96	2	-	[hh:hk:kk]
163	1000028105	<hkxhk>	0	0	0	0	0	0	0	0	16	23	14	0	0	0	0	0	0	0	1.08	2	-	[hh:hk:kk]
164	1000028259	<hkxhk>	0	0	0	0	0	0	0	0	20	7	26	0	0	0	0	0	0	0	30.1	2	*******	[hh:hk:kk]
165	1000028285	<hkxhk>	0	0	0	0	0	0	0	0	16	24	13	0	0	0	0	0	0	0	0.81	2	-	[hh:hk:kk]
166	1000028289	<hkxhk>	0	0	0	0	0	0	0	0	25	15	13	0	0	0	0	0	0	0	15.4	2	******	[hh:hk:kk]

167	100028385	· hkxhk ·	0	0	0	0	0	0	0	0	11	12	30	0	0	0	0	0	0	0	29.5	2	*******	[hh:hk:kk]
168	100028514	· hkxhk ·	0	0	0	0	0	0	0	0	13	20	20	0	0	0	0	0	0	0	5.04	2	*	[hh:hk:kk]
169	100028591	· hkxhk ·	0	0	0	0	0	0	0	0	15	12	26	0	0	0	0	0	0	0	20.4	2	*******	[hh:hk:kk]
170	100028599	· hkxhk ·	0	0	0	0	0	0	0	0	12	21	20	0	0	0	0	0	0	0	4.7	2	*	[hh:hk:kk]
171	100028735	· hkxhk ·	0	0	0	0	0	0	0	0	12	15	26	0	0	0	0	0	0	0	17.4	2	******	[hh:hk:kk]
172	100029160	· hkxhk ·	0	0	0	0	0	0	0	0	13	26	14	0	0	0	0	0	0	0	0.06	2	-	[hh:hk:kk]
173	100029255	· hkxhk ·	0	0	0	0	0	0	0	0	11	27	15	0	0	0	0	0	0	0	0.62	2	-	[hh:hk:kk]
174	100029481	· hkxhk ·	0	0	0	0	0	0	0	0	13	21	19	0	0	0	0	0	0	0	3.64	2	-	[hh:hk:kk]
175	100029779	· hkxhk ·	0	0	0	0	0	0	0	0	12	25	16	0	0	0	0	0	0	0	0.77	2	-	[hh:hk:kk]
176	100029849	· hkxhk ·	0	0	0	0	0	0	0	0	11	19	23	0	0	0	0	0	0	0	9.68	2	***	[hh:hk:kk]
177	100029986	· hkxhk ·	0	0	0	0	0	0	0	0	13	28	12	0	0	0	0	0	0	0	0.21	2	-	[hh:hk:kk]
178	100030282	· hkxhk ·	0	0	0	0	0	0	0	0	15	25	13	0	0	0	0	0	0	0	0.32	2	-	[hh:hk:kk]
179	100030449	<hkxhk>	0	0	0	0	0	0	0	0	13	27	13	0	0	0	0	0	0	0	0.02	2	-	[hh:hk:kk]
180	100030862	· hkxhk>	0	0	0	0	0	0	0	0	12	23	18	0	0	0	0	0	0	0	2.28	2	-	[hh:hk:kk]
181	100031005	<hkxhk>	0	0	0	0	0	0	0	0	14	26	13	0	0	0	0	0	0	0	0.06	2	-	[hh:hk:kk]
182	100031189	<hkxhk>	0	0	0	0	0	0	0	0	18	13	22	0	0	0	0	0	0	0	14.4	2	*****	[hh:hk:kk]
183	100031281	<hkxhk>	0	0	0	0	0	0	0	0	11	27	15	0	0	0	0	0	0	0	0.62	2	-	[hh:hk:kk]
184	100031369	<hkxhk>	0	0	0	0	0	0	0	0	11	28	14	0	0	0	0	0	0	0	0.51	2	-	[hh:hk:kk]
185	100031678	<hkxhk>	0	0	0	0	0	0	0	0	12	27	14	0	0	0	0	0	0	0	0.17	2	-	[hh:hk:kk]
186	100031764	<hkxhk>	0	0	0	0	0	0	0	0	13	13	27	0	0	0	0	0	0	0	21.2	2	*******	[hh:hk:kk]
187	100031871	<hkxhk>	0	0	0	0	0	0	0	0	11	28	14	0	0	0	0	0	0	0	0.51	2	-	[hh:hk:kk]
188	100032017	<hkxhk>	0	0	0	0	0	0	0	0	11	18	24	0	0	0	0	0	0	0	11.8	2	****	[hh:hk:kk]
189	100032166	<hkxhk>	0	0	0	0	0	0	0	0	16	16	21	0	0	0	0	0	0	0	9.26	2	***	[hh:hk:kk]
190	100032688	<hkxhk>	0	0	0	0	0	0	0	0	23	11	19	0	0	0	0	0	0	0	18.7	2	*******	[hh:hk:kk]
191	100032723	<hkxhk>	0	0	0	0	0	0	0	0	15	22	16	0	0	0	0	0	0	0	1.57	2	-	[hh:hk:kk]
192	100032762	<hkxhk>	0	0	0	0	0	0	0	0	15	14	24	0	0	0	0	0	0	0	14.9	2	*****	[hh:hk:kk]
193	100032803	<hkxhk>	0	0	0	0	0	0	0	0	15	25	13	0	0	0	0	0	0	0	0.32	2	-	[hh:hk:kk]
194	100032926	<hkxhk>	0	0	0	0	0	0	0	0	21	13	19	0	0	0	0	0	0	0	13.9	2	*****	[hh:hk:kk]

195	100033508	· hkxhk ·	0	0	0	0	0	0	0	0	11	21	21	0	0	0	0	0	0	0	6.06	2	**	[hh:hk:kk]
196	100033701	· hkxhk ·	0	0	0	0	0	0	0	0	31	10	12	0	0	0	0	0	0	0	34.2	2	*******	[hh:hk:kk]
197	100033709	· hkxhk ·	0	0	0	0	0	0	0	0	11	14	28	0	0	0	0	0	0	0	22.7	2	*******	[hh:hk:kk]
198	100033733	· hkxhk ·	0	0	0	0	0	0	0	0	13	16	24	0	0	0	0	0	0	0	12.9	2	****	[hh:hk:kk]
199	100034000	· hkxhk ·	0	0	0	0	0	0	0	0	11	30	12	0	0	0	0	0	0	0	0.96	2	-	[hh:hk:kk]
200	100034043	· hkxhk ·	0	0	0	0	0	0	0	0	15	23	15	0	0	0	0	0	0	0	0.92	2	-	[hh:hk:kk]
201	100034443	· hkxhk ·	0	0	0	0	0	0	0	0	8	14	14	0	0	0	0	0	0	17	3.78	2	-	[hh:hk:kk]
202	100034525	· hkxhk ·	0	0	0	0	0	0	0	0	12	26	15	0	0	0	0	0	0	0	0.36	2	-	[hh:hk:kk]
203	100034555	· hkxhk ·	0	0	0	0	0	0	0	0	11	29	13	0	0	0	0	0	0	0	0.62	2	-	[hh:hk:kk]
204	100034614	· hkxhk ·	0	0	0	0	0	0	0	0	17	23	13	0	0	0	0	0	0	0	1.53	2	-	[hh:hk:kk]
205	100034632	<hkxhk>	0	0	0	0	0	0	0	0	11	21	21	0	0	0	0	0	0	0	6.06	2	**	[hh:hk:kk]
206	100034760	<hkxhk>	0	0	0	0	0	0	0	0	14	28	11	0	0	0	0	0	0	0	0.51	2	-	[hh:hk:kk]
207	100035096	<hkxhk>	0	0	0	0	0	0	0	0	13	28	12	0	0	0	0	0	0	0	0.21	2	-	[hh:hk:kk]
208	100035308	<hkxhk>	0	0	0	0	0	0	0	0	13	27	13	0	0	0	0	0	0	0	0.02	2	-	[hh:hk:kk]
209	100035332	<hkxhk>	0	0	0	0	0	0	0	0	17	12	24	0	0	0	0	0	0	0	17.7	2	******	[hh:hk:kk]
210	100035723	<hkxhk>	0	0	0	0	0	0	0	0	16	22	15	0	0	0	0	0	0	0	1.57	2	-	[hh:hk:kk]
211	100036121	<hkxhk>	0	0	0	0	0	0	0	0	12	30	11	0	0	0	0	0	0	0	0.96	2	-	[hh:hk:kk]
212	100036165	<hkxhk>	0	0	0	0	0	0	0	0	13	27	13	0	0	0	0	0	0	0	0.02	2	-	[hh:hk:kk]
213	100036533	<hkxhk>	0	0	0	0	0	0	0	0	21	16	16	0	0	0	0	0	0	0	9.26	2	***	[hh:hk:kk]
214	100036544	<hkxhk>	0	0	0	0	0	0	0	0	16	15	22	0	0	0	0	0	0	0	11.3	2	****	[hh:hk:kk]
215	100036641	<hkxhk>	0	0	0	0	0	0	0	0	16	16	21	0	0	0	0	0	0	0	9.26	2	***	[hh:hk:kk]
216	100036775	<hkxhk>	0	0	0	0	0	0	0	0	17	25	11	0	0	0	0	0	0	0	1.53	2	-	[hh:hk:kk]
217	100036930	<hkxhk>	0	0	0	0	0	0	0	0	14	20	19	0	0	0	0	0	0	0	4.13	2	-	[hh:hk:kk]
218	100036947	<hkxhk>	0	0	0	0	0	0	0	0	11	23	19	0	0	0	0	0	0	0	3.34	2	-	[hh:hk:kk]
219	100037219	<hkxhk>	0	0	0	0	0	0	0	0	13	29	11	0	0	0	0	0	0	0	0.62	2	-	[hh:hk:kk]
220	100037623	<hkxhk>	0	0	0	0	0	0	0	0	13	26	14	0	0	0	0	0	0	0	0.06	2	-	[hh:hk:kk]
221	100037727	<hkxhk>	0	0	0	0	0	0	0	0	12	29	12	0	0	0	0	0	0	0	0.47	2	-	[hh:hk:kk]
222	100037899	<hkxhk>	0	0	0	0	0	0	0	0	12	30	11	0	0	0	0	0	0	0	0.96	2	-	[hh:hk:kk]

223	100037906	<hkxhk>	0	0	0	0	0	0	0	0	15	27	11	0	0	0	0	0	0	0	0.62	2	-	[hh:hk:kk]
224	100037955	<hkxhk>	0	0	0	0	0	0	0	0	13	15	25	0	0	0	0	0	0	0	15.4	2	******	[hh:hk:kk]
225	100038458	<hkxhk>	0	0	0	0	0	0	0	0	16	25	12	0	0	0	0	0	0	0	0.77	2	-	[hh:hk:kk]
226	100038509	<hkxhk>	0	0	0	0	0	0	0	0	13	27	13	0	0	0	0	0	0	0	0.02	2	-	[hh:hk:kk]
227	100038767	<hkxhk>	0	0	0	0	0	0	0	0	11	17	25	0	0	0	0	0	0	0	14.2	2	*****	[hh:hk:kk]
228	100038815	<hkxhk>	0	0	0	0	0	0	0	0	21	18	14	0	0	0	0	0	0	0	7.3	2	**	[hh:hk:kk]
229	100038973	<hkxhk>	0	0	0	0	0	0	0	0	19	21	13	0	0	0	0	0	0	0	3.64	2	-	[hh:hk:kk]
230	100039014	<hkxhk>	0	0	0	0	0	0	0	0	17	20	16	0	0	0	0	0	0	0	3.23	2	-	[hh:hk:kk]
231	100039438	<hkxhk>	0	0	0	0	0	0	0	0	14	12	27	0	0	0	0	0	0	0	22.3	2	*******	[hh:hk:kk]
232	100039533	<hkxhk>	0	0	0	0	0	0	0	0	14	16	23	0	0	0	0	0	0	0	11.4	2	****	[hh:hk:kk]
233	100039707	<hkxhk>	0	0	0	0	0	0	0	0	12	21	20	0	0	0	0	0	0	0	4.7	2	*	[hh:hk:kk]
234	100039783	<hkxhk>	0	0	0	0	0	0	0	0	14	14	25	0	0	0	0	0	0	0	16.4	2	******	[hh:hk:kk]
235	100040007	<hkxhk>	0	0	0	0	0	0	0	0	11	14	28	0	0	0	0	0	0	0	22.7	2	*******	[hh:hk:kk]
236	100040023	<hkxhk>	0	0	0	0	0	0	0	0	19	23	11	0	0	0	0	0	0	0	3.34	2	-	[hh:hk:kk]
237	100040112	<hkxhk>	0	0	0	0	0	0	0	0	12	15	26	0	0	0	0	0	0	0	17.4	2	******	[hh:hk:kk]
238	100040285	<hkxhk>	0	0	0	0	0	0	0	0	12	27	14	0	0	0	0	0	0	0	0.17	2	-	[hh:hk:kk]
239	100040322	<hkxhk>	0	0	0	0	0	0	0	0	27	14	12	0	0	0	0	0	0	0	20.3	2	*******	[hh:hk:kk]
240	100040393	<hkxhk>	0	0	0	0	0	0	0	0	13	28	12	0	0	0	0	0	0	0	0.21	2	-	[hh:hk:kk]
241	100040545	<hkxhk>	0	0	0	0	0	0	0	0	15	15	23	0	0	0	0	0	0	0	12.4	2	****	[hh:hk:kk]
242	100040593	<hkxhk>	0	0	0	0	0	0	0	0	11	28	14	0	0	0	0	0	0	0	0.51	2	-	[hh:hk:kk]
243	100040641	<hkxhk>	0	0	0	0	0	0	0	0	12	26	15	0	0	0	0	0	0	0	0.36	2	-	[hh:hk:kk]
244	100040999	<hkxhk>	0	0	0	0	0	0	0	0	19	13	21	0	0	0	0	0	0	0	13.9	2	*****	[hh:hk:kk]
245	100041012	<hkxhk>	0	0	0	0	0	0	0	0	11	26	16	0	0	0	0	0	0	0	0.96	2	-	[hh:hk:kk]
246	100041349	<hkxhk>	0	0	0	0	0	0	0	0	21	20	12	0	0	0	0	0	0	0	6.25	2	**	[hh:hk:kk]
247	100041525	<hkxhk>	0	0	0	0	0	0	0	0	12	25	16	0	0	0	0	0	0	0	0.77	2	-	[hh:hk:kk]
248	100041534	<hkxhk>	0	0	0	0	0	0	0	0	12	20	21	0	0	0	0	0	0	0	6.25	2	**	[hh:hk:kk]
249	100041599	<hkxhk>	0	0	0	0	0	0	0	0	13	12	28	0	0	0	0	0	0	0	24.4	2	*******	[hh:hk:kk]
250	100041763	<hkxhk>	0	0	0	0	0	0	0	0	11	17	25	0	0	0	0	0	0	0	14.2	2	*****	[hh:hk:kk]

251	100042316	<hkxhk>	0	0	0	0	0	0	0	0	11	27	15	0	0	0	0	0	0	0	0.62	2	-	[hh:hk:kk]
252	100042431	<hkxhk>	0	0	0	0	0	0	0	0	15	22	16	0	0	0	0	0	0	0	1.57	2	-	[hh:hk:kk]
253	100043095	<hkxhk>	0	0	0	0	0	0	0	0	14	13	26	0	0	0	0	0	0	0	19.2	2	*******	[hh:hk:kk]
254	100043280	<hkxhk>	0	0	0	0	0	0	0	0	24	13	16	0	0	0	0	0	0	0	16.2	2	******	[hh:hk:kk]
255	100043408	<hkxhk>	0	0	0	0	0	0	0	0	13	25	15	0	0	0	0	0	0	0	0.32	2	-	[hh:hk:kk]
256	100043628	<hkxhk>	0	0	0	0	0	0	0	0	13	18	22	0	0	0	0	0	0	0	8.51	2	**	[hh:hk:kk]
257	100043754	<hkxhk>	0	0	0	0	0	0	0	0	14	19	20	0	0	0	0	0	0	0	5.6	2	*	[hh:hk:kk]
258	100043908	<hkxhk>	0	0	0	0	0	0	0	0	12	25	16	0	0	0	0	0	0	0	0.77	2	-	[hh:hk:kk]
259	100044415	<hkxhk>	0	0	0	0	0	0	0	0	17	12	24	0	0	0	0	0	0	0	17.7	2	******	[hh:hk:kk]
260	100045296	<hkxhk>	0	0	0	0	0	0	0	0	12	27	14	0	0	0	0	0	0	0	0.17	2	-	[hh:hk:kk]
261	100045448	<hkxhk>	0	0	0	0	0	0	0	0	16	24	13	0	0	0	0	0	0	0	0.81	2	-	[hh:hk:kk]
262	100045489	<hkxhk>	0	0	0	0	0	0	0	0	12	14	26	0	0	0	0	0	0	1	18.6	2	*******	[hh:hk:kk]
263	100045613	<hkxhk>	0	0	0	0	0	0	0	0	15	25	13	0	0	0	0	0	0	0	0.32	2	-	[hh:hk:kk]
264	100045682	<hkxhk>	0	0	0	0	0	0	0	0	13	12	28	0	0	0	0	0	0	0	24.4	2	*******	[hh:hk:kk]
265	100045831	<hkxhk>	0	0	0	0	0	0	0	0	15	27	11	0	0	0	0	0	0	0	0.62	2	-	[hh:hk:kk]
266	100046025	<hkxhk>	0	0	0	0	0	0	0	0	13	28	12	0	0	0	0	0	0	0	0.21	2	-	[hh:hk:kk]
267	100046280	<hkxhk>	0	0	0	0	0	0	0	0	13	22	18	0	0	0	0	0	0	0	2.47	2	-	[hh:hk:kk]
268	100046469	<hkxhk>	0	0	0	0	0	0	0	0	15	25	13	0	0	0	0	0	0	0	0.32	2	-	[hh:hk:kk]
269	100046903	<hkxhk>	0	0	0	0	0	0	0	0	13	27	13	0	0	0	0	0	0	0	0.02	2	-	[hh:hk:kk]
270	100047150	<hkxhk>	0	0	0	0	0	0	0	0	15	25	13	0	0	0	0	0	0	0	0.32	2	-	[hh:hk:kk]
271	100047308	<hkxhk>	0	0	0	0	0	0	0	0	13	28	12	0	0	0	0	0	0	0	0.21	2	-	[hh:hk:kk]
272	100048016	<hkxhk>	0	0	0	0	0	0	0	0	12	28	13	0	0	0	0	0	0	0	0.21	2	-	[hh:hk:kk]
273	100048019	<hkxhk>	0	0	0	0	0	0	0	0	15	27	11	0	0	0	0	0	0	0	0.62	2	-	[hh:hk:kk]
274	100048078	<hkxhk>	0	0	0	0	0	0	0	0	12	28	13	0	0	0	0	0	0	0	0.21	2	-	[hh:hk:kk]
275	100048112	<hkxhk>	0	0	0	0	0	0	0	0	20	21	12	0	0	0	0	0	0	0	4.7	2	*	[hh:hk:kk]
276	100048168	<hkxhk>	0	0	0	0	0	0	0	0	11	26	16	0	0	0	0	0	0	0	0.96	2	-	[hh:hk:kk]
277	100048615	<hkxhk>	0	0	0	0	0	0	0	0	13	28	12	0	0	0	0	0	0	0	0.21	2	-	[hh:hk:kk]
278	100048632	<hkxhk>	0	0	0	0	0	0	0	0	12	28	13	0	0	0	0	0	0	0	0.21	2	-	[hh:hk:kk]

279	100049015	hkxhk	0	0	0	0	0	0	0	0	11	28	14	0	0	0	0	0	0	0	0.51	2	-	[hh:hk:kk]
280	100049194	hkxhk	0	0	0	0	0	0	0	0	11	28	14	0	0	0	0	0	0	0	0.51	2	-	[hh:hk:kk]
281	100050419	hkxhk	0	0	0	0	0	0	0	0	13	23	17	0	0	0	0	0	0	0	1.83	2	-	[hh:hk:kk]
282	100050450	hkxhk	0	0	0	0	0	0	0	0	20	9	10	0	0	0	0	0	0	14	16.4	2	******	[hh:hk:kk]
283	100050487	hkxhk	0	0	0	0	0	0	0	0	21	19	13	0	0	0	0	0	0	0	6.66	2	**	[hh:hk:kk]
284	100050857	hkxhk	0	0	0	0	0	0	0	0	28	13	12	0	0	0	0	0	0	0	23.4	2	*******	[hh:hk:kk]
285	100050970	hkxhk	0	0	0	0	0	0	0	0	19	22	12	0	0	0	0	0	0	0	3.38	2	-	[hh:hk:kk]
286	100051207	hkxhk	0	0	0	0	0	0	0	0	13	26	14	0	0	0	0	0	0	0	0.06	2	-	[hh:hk:kk]
287	100051908	hkxhk	0	0	0	0	0	0	0	0	19	20	14	0	0	0	0	0	0	0	4.13	2	-	[hh:hk:kk]
288	100052401	hkxhk	0	0	0	0	0	0	0	0	22	14	17	0	0	0	0	0	0	0	12.7	2	****	[hh:hk:kk]
289	100052570	hkxhk	0	0	0	0	0	0	0	0	12	11	30	0	0	0	0	0	0	0	30.4	2	*******	[hh:hk:kk]
290	100052860	hkxhk	0	0	0	0	0	0	0	0	14	26	13	0	0	0	0	0	0	0	0.06	2	-	[hh:hk:kk]
291	100053128	hkxhk	0	0	0	0	0	0	0	0	13	26	14	0	0	0	0	0	0	0	0.06	2	-	[hh:hk:kk]
292	100053687	hkxhk	0	0	0	0	0	0	0	0	13	29	11	0	0	0	0	0	0	0	0.62	2	-	[hh:hk:kk]
293	100054433	hkxhk	0	0	0	0	0	0	0	0	29	11	13	0	0	0	0	0	0	0	27.8	2	*******	[hh:hk:kk]
294	100054662	hkxhk	0	0	0	0	0	0	0	0	16	26	11	0	0	0	0	0	0	0	0.96	2	-	[hh:hk:kk]
295	100055758	hkxhk	0	0	0	0	0	0	0	0	14	24	15	0	0	0	0	0	0	0	0.51	2	-	[hh:hk:kk]
296	100056579	hkxhk	0	0	0	0	0	0	0	0	20	17	16	0	0	0	0	0	0	0	7.42	2	**	[hh:hk:kk]
297	100057062	hkxhk	0	0	0	0	0	0	0	0	20	11	22	0	0	0	0	0	0	0	18.3	2	******	[hh:hk:kk]
298	100057439	hkxhk	0	0	0	0	0	0	0	0	14	25	14	0	0	0	0	0	0	0	0.17	2	-	[hh:hk:kk]
299	100057592	hkxhk	0	0	0	0	0	0	0	0	14	28	11	0	0	0	0	0	0	0	0.51	2	-	[hh:hk:kk]
300	100057910	hkxhk	0	0	0	0	0	0	0	0	17	21	15	0	0	0	0	0	0	0	2.43	2	-	[hh:hk:kk]
301	100059495	hkxhk	0	0	0	0	0	0	0	0	12	29	12	0	0	0	0	0	0	0	0.47	2	-	[hh:hk:kk]
302	100059518	hkxhk	0	0	0	0	0	0	0	0	12	26	15	0	0	0	0	0	0	0	0.36	2	-	[hh:hk:kk]
303	100059912	hkxhk	0	0	0	0	0	0	0	0	12	26	15	0	0	0	0	0	0	0	0.36	2	-	[hh:hk:kk]
304	100060003	hkxhk	0	0	0	0	0	0	0	0	21	16	16	0	0	0	0	0	0	0	9.26	2	***	[hh:hk:kk]
305	100062795	hkxhk	0	0	0	0	0	0	0	0	12	24	17	0	0	0	0	0	0	0	1.42	2	-	[hh:hk:kk]
306	100063034	hkxhk	0	0	0	0	0	0	0	0	11	28	14	0	0	0	0	0	0	0	0.51	2	-	[hh:hk:kk]

307	1000064190	hkxhk	0	0	0	0	0	0	0	0	12	11	30	0	0	0	0	0	0	0	30.4	2	*******	[hh:hk:kk]
308	1000065594	hkxhk	0	0	0	0	0	0	0	0	16	25	12	0	0	0	0	0	0	0	0.77	2		[hh:hk:kk]
309	1000065700	hkxhk	0	0	0	0	0	0	0	0	13	10	30	0	0	0	0	0	0	0	31.5	2	*******	[hh:hk:kk]
310	1000066627	hkxhk	0	0	0	0	0	0	0	0	12	18	23	0	0	0	0	0	0	0	10	2	***	[hh:hk:kk]
311	1000067119	hkxhk	0	0	0	0	0	0	0	0	17	22	14	0	0	0	0	0	0	0	1.87	2		[hh:hk:kk]
312	1000069544	hkxhk	0	0	0	0	0	0	0	0	27	8	18	0	0	0	0	0	0	0	28.9	2	*******	[hh:hk:kk]
313	1000069637	hkxhk	0	0	0	0	0	0	0	0	13	14	26	0	0	0	0	0	0	0	18.2	2	******	[hh:hk:kk]
314	1000070876	hkxhk	0	0	0	0	0	0	0	0	27	14	12	0	0	0	0	0	0	0	20.3	2	*******	[hh:hk:kk]
315	1000070986	hkxhk	0	0	0	0	0	0	0	0	15	24	14	0	0	0	0	0	0	0	0.51	2	-	[hh:hk:kk]
316	1000071689	hkxhk	0	0	0	0	0	0	0	0	12	17	24	0	0	0	0	0	0	0	12.3	2	****	[hh:hk:kk]
317	1000071907	hkxhk	0	0	0	0	0	0	0	0	13	29	11	0	0	0	0	0	0	0	0.62	2	-	[hh:hk:kk]
318	1000072097	hkxhk	0	0	0	0	0	0	0	0	13	24	16	0	0	0	0	0	0	0	0.81	2	-	[hh:hk:kk]
319	1000072406	hkxhk	0	0	0	0	0	0	0	0	19	23	11	0	0	0	0	0	0	0	3.34	2	-	[hh:hk:kk]
320	1000072715	hkxhk	0	0	0	0	0	0	0	0	19	10	24	0	0	0	0	0	0	0	21.5	2	*******	[hh:hk:kk]
321	1000073360	hkxhk	0	0	0	0	0	0	0	0	11	21	21	0	0	0	0	0	0	0	6.06	2	**	[hh:hk:kk]
322	1000075009	hkxhk	0	0	0	0	0	0	0	0	12	30	11	0	0	0	0	0	0	0	0.96	2	-	[hh:hk:kk]
323	1000076502	hkxhk	0	0	0	0	0	0	0	0	12	28	13	0	0	0	0	0	0	0	0.21	2	-	[hh:hk:kk]
324	1000077687	hkxhk	0	0	0	0	0	0	0	0	11	25	17	0	0	0	0	0	0	0	1.53	2	-	[hh:hk:kk]
325	1000078595	hkxhk	0	0	0	0	0	0	0	0	24	14	15	0	0	0	0	0	0	0	14.9	2	*****	[hh:hk:kk]
326	1000079612	hkxhk	0	0	0	0	0	0	0	0	32	8	13	0	0	0	0	0	0	0	39.5	2	*******	[hh:hk:kk]
327	1000086424	hkxhk	0	0	0	0	0	0	0	0	17	20	16	0	0	0	0	0	0	0	3.23	2	-	[hh:hk:kk]

Appendix 6: The locus genotype frequency for F_1 map of phase determination

S n	Locus	Segregation	ac	ad	bc	bd	ee	e f	eg	fg	hh	hk	kk	h-	k-	ll	lm	nn	np	--	X2	DI	Signif	Classification
1	100008469	<hkxhk>	0	0	0	0	0	0	0	0	12	10	8	0	0	0	0	0	0	0	4.4	2	-	[hh:hk:kk]
2	100016058	<hkxhk>	0	0	0	0	0	0	0	0	9	11	10	0	0	0	0	0	0	0	2.2	2		[hh:hk:kk]
3	100018090	<hkxhk>	0	0	0	0	0	0	0	0	10	10	10	0	0	0	0	0	0	0	3.33	2	-	[hh:hk:kk]
4	100019201	<hkxhk>	0	0	0	0	0	0	0	0	10	11	9	0	0	0	0	0	0	0	2.2	2	-	[hh:hk:kk]
5	100023574	<hkxhk>	0	0	0	0	0	0	0	0	13	7	10	0	0	0	0	0	0	0	9.13	2	**	[hh:hk:kk]
6	100023875	<hkxhk>	0	0	0	0	0	0	0	0	12	8	10	0	0	0	0	0	0	0	6.8	2	**	[hh:hk:kk]
7	100024282	<hkxhk>	0	0	0	0	0	0	0	0	8	11	11	0	0	0	0	0	0	0	2.73	2	-	[hh:hk:kk]
8	100025590	<hkxhk>	0	0	0	0	0	0	0	0	10	8	12	0	0	0	0	0	0	0	6.8	2	**	[hh:hk:kk]
9	100025881	<hkxhk>	0	0	0	0	0	0	0	0	10	9	11	0	0	0	0	0	0	0	4.87	2	*	[hh:hk:kk]
10	100026967	<hkxhk>	0	0	0	0	0	0	0	0	11	9	10	0	0	0	0	0	0	0	4.87	2	*	[hh:hk:kk]
11	100028105	<hkxhk>	0	0	0	0	0	0	0	0	7	10	13	0	0	0	0	0	0	0	5.73	2	*	[hh:hk:kk]
12	100032632	<hkxhk>	0	0	0	0	0	0	0	0	11	9	10	0	0	0	0	0	0	0	4.87	2	*	[hh:hk:kk]
13	100046469	<hkxhk>	0	0	0	0	0	0	0	0	9	11	10	0	0	0	0	0	0	0	2.2	2	-	[hh:hk:kk]
14	100047308	<hkxhk>	0	0	0	0	0	0	0	0	8	11	11	0	0	0	0	0	0	0	2.73	2	-	[hh:hk:kk]
15	100048016	<hkxhk>	0	0	0	0	0	0	0	0	8	12	10	0	0	0	0	0	0	0	1.47	2	-	[hh:hk:kk]
16	100051207	<hkxhk>	0	0	0	0	0	0	0	0	8	12	10	0	0	0	0	0	0	0	1.47	2	-	[hh:hk:kk]
17	100057439	<hkxhk>	0	0	0	0	0	0	0	0	8	12	10	0	0	0	0	0	0	0	1.47	2	-	[hh:hk:kk]
18	100070986	<hkxhk>	0	0	0	0	0	0	0	0	9	12	9	0	0	0	0	0	0	0	1.2	2	-	[hh:hk:kk]
19	100006384	<hkxhk>	0	0	0	0	0	0	0	0	6	12	12	0	0	0	0	0	0	0	3.6	2	-	[hh:hk:kk]
20	100012437	<hkxhk>	0	0	0	0	0	0	0	0	7	15	8	0	0	0	0	0	0	0	0.07	2	-	[hh:hk:kk]
21	100013479	<hkxhk>	0	0	0	0	0	0	0	0	12	6	12	0	0	0	0	0	0	0	10.8	2	****	[hh:hk:kk]
22	100014104	<hkxhk>	0	0	0	0	0	0	0	0	18	1	11	0	0	0	0	0	0	0	29.4	2	*******	[hh:hk:kk]
23	100014804	<hkxhk>	0	0	0	0	0	0	0	0	13	0	17	0	0	0	0	0	0	0	31.07	2	*******	[hh:hk:kk]
24	100017178	<hkxhk>	0	0	0	0	0	0	0	0	13	5	12	0	0	0	0	0	0	0	13.4	2	****	[hh:hk:kk]
25	100017471	<hkxhk>	0	0	0	0	0	0	0	0	17	2	11	0	0	0	0	0	0	0	24.93	2	*******	[hh:hk:kk]

26	100018499	· hkxhk ·	0	0	0	0	0	0	0	0	17	2	11	0	0	0	0	0	0	0	24.93	2	*******	[hh:hk:kk]
27	100018674	· hkxhk ·	0	0	0	0	0	0	0	0	19	5	6	0	0	0	0	0	0	0	24.6	2	*******	[hh:hk:kk]
28	100018959	· hkxhk ·	0	0	0	0	0	0	0	0	16	2	12	0	0	0	0	0	0	0	23.6	2	*******	[hh:hk:kk]
29	100018973	· hkxhk ·	0	0	0	0	0	0	0	0	17	5	8	0	0	0	0	0	0	0	18.73	2	*******	[hh:hk:kk]
30	100019111	· hkxhk ·	0	0	0	0	0	0	0	0	16	2	12	0	0	0	0	0	0	0	23.6	2	*******	[hh:hk:kk]
31	100022045	· hkxhk ·	0	0	0	0	0	0	0	0	7	15	8	0	0	0	0	0	0	0	0.07	2	-	[hh:hk:kk]
32	100022478	· hkxhk ·	0	0	0	0	0	0	0	0	7	16	7	0	0	0	0	0	0	0	0.13	2	-	[hh:hk:kk]
33	100023297	· hkxhk ·	0	0	0	0	0	0	0	0	9	15	6	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
34	100024112	· hkxhk ·	0	0	0	0	0	0	0	0	13	9	8	0	0	0	0	0	0	0	6.47	2	**	[hh:hk:kk]
35	100024784	· hkxhk ·	0	0	0	0	0	0	0	0	12	12	6	0	0	0	0	0	0	0	3.6	2	-	[hh:hk:kk]
36	100027144	· hkxhk ·	0	0	0	0	0	0	0	0	7	15	8	0	0	0	0	0	0	0	0.07	2	-	[hh:hk:kk]
37	100028514	· hkxhk ·	0	0	0	0	0	0	0	0	10	12	8	0	0	0	0	0	0	0	1.47	2	-	[hh:hk:kk]
38	100029037	· hkxhk ·	0	0	0	0	0	0	0	0	12	11	7	0	0	0	0	0	0	0	3.8	2	-	[hh:hk:kk]
39	100030913	· hkxhk ·	0	0	0	0	0	0	0	0	6	16	8	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
40	100031369	· hkxhk ·	0	0	0	0	0	0	0	0	9	14	7	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
41	100031859	<hkxhk>	0	0	0	0	0	0	0	0	13	9	8	0	0	0	0	0	0	0	6.47	2	**	[hh:hk:kk]
42	100032710	<hkxhk>	0	0	0	0	0	0	0	0	7	12	11	0	0	0	0	0	0	0	2.27	2	-	[hh:hk:kk]
43	100048078	<hkxhk>	0	0	0	0	0	0	0	0	15	4	11	0	0	0	0	0	0	0	17.2	2	******	[hh:hk:kk]
44	100005111	<hkxhk>	0	0	0	0	0	0	0	0	8	16	6	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
45	100009010	<hkxhk>	0	0	0	0	0	0	0	0	7	17	6	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
46	100021788	<hkxhk>	0	0	0	0	0	0	0	0	7	16	7	0	0	0	0	0	0	0	0.13	2	-	[hh:hk:kk]
47	100026552	<hkxhk>	0	0	0	0	0	0	0	0	9	15	6	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
48	100031150	<hkxhk>	0	0	0	0	0	0	0	0	7	17	6	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
49	100037219	<hkxhk>	0	0	0	0	0	0	0	0	7	14	9	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
50	100046025	<hkxhk>	0	0	0	0	0	0	0	0	8	16	6	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
51	100006527	<hkxhk>	0	0	0	0	0	0	0	0	7	6	17	0	0	0	0	0	0	0	17.47	2	******	[hh:hk:kk]
52	100014292	<hkxhk>	0	0	0	0	0	0	0	0	13	10	7	0	0	0	0	0	0	0	5.73	2	*	[hh:hk:kk]
53	100020111	<hkxhk>	0	0	0	0	0	0	0	0	11	12	7	0	0	0	0	0	0	0	2.27	2	-	[hh:hk:kk]

54	100029652	-hkxhk-	0	0	0	0	0	0	0	0	11	12	7	0	0	0	0	0	0	0	2.27	2	-	[hh:hk:kk]
55	Locus138	-nnxnp-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	16	12	2	0.57	1		[nn:np]
56	100005427	-hkxhk-	0	0	0	0	0	0	0	0	11	9	10	0	0	0	0	0	0	0	4.87	2	*	[hh:hk:kk]
57	100005890	-hkxhk-	0	0	0	0	0	0	0	0	7	12	11	0	0	0	0	0	0	0	2.27	2	-	[hh:hk:kk]
58	100029233	-hkxhk-	0	0	0	0	0	0	0	0	6	16	8	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
59	Locus027	-efxeg-	0	0	0	0	5	6	10	9	0	0	0	0	0	0	0	0	0	0	2.27	3	-	[ee:ef:eg:fg]
60	100002653	-hkxhk-	0	0	0	0	0	0	0	0	10	13	7	0	0	0	0	0	0	0	1.13	2	-	[hh:hk:kk]
61	100004291	-hkxhk-	0	0	0	0	0	0	0	0	7	14	9	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
62	100006206	-hkxhk-	0	0	0	0	0	0	0	0	9	12	9	0	0	0	0	0	0	0	1.2	2	-	[hh:hk:kk]
63	100006963	-hkxhk-	0	0	0	0	0	0	0	0	8	15	7	0	0	0	0	0	0	0	0.07	2	-	[hh:hk:kk]
64	100007343	-hkxhk-	0	0	0	0	0	0	0	0	16	0	14	0	0	0	0	0	0	0	30.27	2	*******	[hh:hk:kk]
65	100008320	-hkxhk-	0	0	0	0	0	0	0	0	11	10	9	0	0	0	0	0	0	0	3.6	2	-	[hh:hk:kk]
66	100008345	-hkxhk-	0	0	0	0	0	0	0	0	9	14	7	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
67	100009512	-hkxhk-	0	0	0	0	0	0	0	0	8	16	6	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
68	100011705	-hkxhk-	0	0	0	0	0	0	0	0	14	7	9	0	0	0	0	0	0	0	10.2	2	***	[hh:hk:kk]
69	100012265	<hkxhk>	0	0	0	0	0	0	0	0	10	12	8	0	0	0	0	0	0	0	1.47	2	-	[hh:hk:kk]
70	100013570	<hkxhk>	0	0	0	0	0	0	0	0	7	13	10	0	0	0	0	0	0	0	1.13	2	-	[hh:hk:kk]
71	100013913	<hkxhk>	0	0	0	0	0	0	0	0	9	13	8	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
72	100014551	<hkxhk>	0	0	0	0	0	0	0	0	8	9	13	0	0	0	0	0	0	0	6.47	2	**	[hh:hk:kk]
73	100015250	<hkxhk>	0	0	0	0	0	0	0	0	6	13	11	0	0	0	0	0	0	0	2.2	2	-	[hh:hk:kk]
74	100017059	<hkxhk>	0	0	0	0	0	0	0	0	8	14	8	0	0	0	0	0	0	0	0.13	2	-	[hh:hk:kk]
75	100017459	<hkxhk>	0	0	0	0	0	0	0	0	13	2	15	0	0	0	0	0	0	0	22.8	2	*******	[hh:hk:kk]
76	100018291	<hkxhk>	0	0	0	0	0	0	0	0	9	13	8	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
77	100019602	<hkxhk>	0	0	0	0	0	0	0	0	14	8	8	0	0	0	0	0	0	0	8.93	2	**	[hh:hk:kk]
78	100021053	<hkxhk>	0	0	0	0	0	0	0	0	8	13	9	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
79	100021238	<hkxhk>	0	0	0	0	0	0	0	0	11	11	8	0	0	0	0	0	0	0	2.73	2	-	[hh:hk:kk]
80	100022146	<hkxhk>	0	0	0	0	0	0	0	0	8	14	8	0	0	0	0	0	0	0	0.13	2	-	[hh:hk:kk]
81	100022796	<hkxhk>	0	0	0	0	0	0	0	0	9	14	7	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]

82	100022821	<hkxhk>	0	0	0	0	0	0	0	0	13	12	5	0	0	0	0	0	0	0	5.47	2	*	[hh:hk:kk]
83	100023339	<hkxhk>	0	0	0	0	0	0	0	0	10	14	6	0	0	0	0	0	0	0	1.2	2	-	[hh:hk:kk]
84	100024471	<hkxhk>	0	0	0	0	0	0	0	0	9	15	6	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
85	100025976	<hkxhk>	0	0	0	0	0	0	0	0	9	13	8	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
86	100028540	<hkxhk>	0	0	0	0	0	0	0	0	9	13	8	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
87	100028735	<hkxhk>	0	0	0	0	0	0	0	0	15	1	14	0	0	0	0	0	0	0	26.2	2	*******	[hh:hk:kk]
88	100029752	<hkxhk>	0	0	0	0	0	0	0	0	9	13	8	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
89	100032017	<hkxhk>	0	0	0	0	0	0	0	0	6	15	9	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
90	100032251	<hkxhk>	0	0	0	0	0	0	0	0	6	13	11	0	0	0	0	0	0	0	2.2	2	-	[hh:hk:kk]
91	100033709	<hkxhk>	0	0	0	0	0	0	0	0	8	12	10	0	0	0	0	0	0	0	1.47	2	-	[hh:hk:kk]
92	100033824	<hkxhk>	0	0	0	0	0	0	0	0	6	10	14	0	0	0	0	0	0	0	7.6	2	**	[hh:hk:kk]
93	100038973	<hkxhk>	0	0	0	0	0	0	0	0	6	14	10	0	0	0	0	0	0	0	1.2	2	-	[hh:hk:kk]
94	100069637	<hkxhk>	0	0	0	0	0	0	0	0	9	14	7	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
95	Locus047	<efxeg>	0	0	0	0	7	7	10	6	0	0	0	0	0	0	0	0	0	0	1.2	3	-	[ee:ef:eg:fg]
96	100000976	<hkxhk>	0	0	0	0	0	0	0	0	7	15	8	0	0	0	0	0	0	0	0.07	2	-	[hh:hk:kk]
97	100008738	<hkxhk>	0	0	0	0	0	0	0	0	8	15	7	0	0	0	0	0	0	0	0.07	2	-	[hh:hk:kk]
98	100009895	<hkxhk>	0	0	0	0	0	0	0	0	7	13	10	0	0	0	0	0	0	0	1.13	2	-	[hh:hk:kk]
99	100012230	<hkxhk>	0	0	0	0	0	0	0	0	7	14	9	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
100	100012237	<hkxhk>	0	0	0	0	0	0	0	0	6	16	8	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
101	100013350	<hkxhk>	0	0	0	0	0	0	0	0	7	14	9	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
102	100015698	<hkxhk>	0	0	0	0	0	0	0	0	10	12	8	0	0	0	0	0	0	0	1.47	2	-	[hh:hk:kk]
103	100017309	<hkxhk>	0	0	0	0	0	0	0	0	8	15	7	0	0	0	0	0	0	0	0.07	2	-	[hh:hk:kk]
104	100017791	<hkxhk>	0	0	0	0	0	0	0	0	13	10	7	0	0	0	0	0	0	0	5.73	2	*	[hh:hk:kk]
105	100018989	<hkxhk>	0	0	0	0	0	0	0	0	7	13	10	0	0	0	0	0	0	0	1.13	2	-	[hh:hk:kk]
106	100019455	<hkxhk>	0	0	0	0	0	0	0	0	9	11	10	0	0	0	0	0	0	0	2.2	2	-	[hh:hk:kk]
107	100021760	<hkxhk>	0	0	0	0	0	0	0	0	8	14	8	0	0	0	0	0	0	0	0.13	2	-	[hh:hk:kk]
108	100023677	<hkxhk>	0	0	0	0	0	0	0	0	7	11	12	0	0	0	0	0	0	0	3.8	2	-	[hh:hk:kk]
109	100024426	<hkxhk>	0	0	0	0	0	0	0	0	8	13	9	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]

110	100024466	<hkxhk>	0	0	0	0	0	0	0	0	7	14	9	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
111	100024540	<hkxhk>	0	0	0	0	0	0	0	0	9	13	8	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
112	100026767	<hkxhk>	0	0	0	0	0	0	0	0	7	16	7	0	0	0	0	0	0	0	0.13	2	-	[hh:hk:kk]
113	100030840	<hkxhk>	0	0	0	0	0	0	0	0	8	14	8	0	0	0	0	0	0	0	0.13	2	-	[hh:hk:kk]
114	100032723	<hkxhk>	0	0	0	0	0	0	0	0	8	14	8	0	0	0	0	0	0	0	0.13	2	-	[hh:hk:kk]
115	100035332	<hkxhk>	0	0	0	0	0	0	0	0	9	14	7	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
116	100035723	<hkxhk>	0	0	0	0	0	0	0	0	8	15	7	0	0	0	0	0	0	0	0.07	2	-	[hh:hk:kk]
117	100040545	<hkxhk>	0	0	0	0	0	0	0	0	7	13	10	0	0	0	0	0	0	0	1.13	2	-	[hh:hk:kk]
118	100043280	<hkxhk>	0	0	0	0	0	0	0	0	11	11	8	0	0	0	0	0	0	0	2.73	2	-	[hh:hk:kk]
119	100086424	<hkxhk>	0	0	0	0	0	0	0	0	12	12	6	0	0	0	0	0	0	0	3.6	2	-	[hh:hk:kk]
120	Locus056	<nnxnp>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	14	16	0	0.13	1	-	[nn:np]
121	100001627	<hkxhk>	0	0	0	0	0	0	0	0	6	19	5	0	0	0	0	0	0	0	2.2	2	-	[hh:hk:kk]
122	100003487	<hkxhk>	0	0	0	0	0	0	0	0	6	16	8	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
123	100005987	<hkxhk>	0	0	0	0	0	0	0	0	8	16	6	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
124	100006028	<hkxhk>	0	0	0	0	0	0	0	0	9	11	10	0	0	0	0	0	0	0	2.2	2	-	[hh:hk:kk]
125	100007865	<hkxhk>	0	0	0	0	0	0	0	0	14	0	16	0	0	0	0	0	0	0	30.27	2	*******	[hh:hk:kk]
126	100008854	<hkxhk>	0	0	0	0	0	0	0	0	7	7	16	0	0	0	0	0	0	0	13.93	2	*****	[hh:hk:kk]
127	100009138	<hkxhk>	0	0	0	0	0	0	0	0	7	8	15	0	0	0	0	0	0	0	10.8	2	****	[hh:hk:kk]
128	100009167	<hkxhk>	0	0	0	0	0	0	0	0	8	7	15	0	0	0	0	0	0	0	11.8	2	****	[hh:hk:kk]
129	100010289	<hkxhk>	0	0	0	0	0	0	0	0	8	16	6	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
130	100010318	<hkxhk>	0	0	0	0	0	0	0	0	9	13	8	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
131	100012941	<hkxhk>	0	0	0	0	0	0	0	0	7	9	14	0	0	0	0	0	0	0	8.07	2	**	[hh:hk:kk]
132	100013212	<hkxhk>	0	0	0	0	0	0	0	0	7	10	13	0	0	0	0	0	0	0	5.73	2	*	[hh:hk:kk]
133	100017133	<hkxhk>	0	0	0	0	0	0	0	0	8	9	13	0	0	0	0	0	0	0	6.47	2	**	[hh:hk:kk]
134	100017754	<hkxhk>	0	0	0	0	0	0	0	0	14	1	15	0	0	0	0	0	0	0	26.2	2	*******	[hh:hk:kk]
135	100017829	<hkxhk>	0	0	0	0	0	0	0	0	8	9	13	0	0	0	0	0	0	0	6.47	2	**	[hh:hk:kk]
136	100018657	<hkxhk>	0	0	0	0	0	0	0	0	6	18	6	0	0	0	0	0	0	0	1.2	2	-	[hh:hk:kk]
137	100018991	<hkxhk>	0	0	0	0	0	0	0	0	9	14	7	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]

138	100019268	<hkxhk>	0	0	0	0	0	0	0	0	7	16	7	0	0	0	0	0	0	0	0.13	2	-	[hh:hk:kk]
139	100019701	<hkxhk>	0	0	0	0	0	0	0	0	14	10	6	0	0	0	0	0	0	0	7.6	2	**	[hh:hk:kk]
140	100019906	<hkxhk>	0	0	0	0	0	0	0	0	7	15	8	0	0	0	0	0	0	0	0.07	2	-	[hh:hk:kk]
141	100020085	<hkxhk>	0	0	0	0	0	0	0	0	9	14	7	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
142	100021259	<hkxhk>	0	0	0	0	0	0	0	0	7	17	6	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
143	100021260	<hkxhk>	0	0	0	0	0	0	0	0	8	15	7	0	0	0	0	0	0	0	0.07	2	-	[hh:hk:kk]
144	100021651	<hkxhk>	0	0	0	0	0	0	0	0	8	9	13	0	0	0	0	0	0	0	6.47	2	**	[hh:hk:kk]
145	100022390	<hkxhk>	0	0	0	0	0	0	0	0	9	7	14	0	0	0	0	0	0	0	10.2	2	***	[hh:hk:kk]
146	100022419	<hkxhk>	0	0	0	0	0	0	0	0	9	15	6	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
147	100024730	<hkxhk>	0	0	0	0	0	0	0	0	7	13	10	0	0	0	0	0	0	0	1.13	2	-	[hh:hk:kk]
148	100024767	<hkxhk>	0	0	0	0	0	0	0	0	10	13	7	0	0	0	0	0	0	0	1.13	2	-	[hh:hk:kk]
149	100025514	<hkxhk>	0	0	0	0	0	0	0	0	14	5	11	0	0	0	0	0	0	0	13.93	2	*****	[hh:hk:kk]
150	100025746	<hkxhk>	0	0	0	0	0	0	0	0	7	17	6	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
151	100025975	<hkxhk>	0	0	0	0	0	0	0	0	6	16	8	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
152	100026233	<hkxhk>	0	0	0	0	0	0	0	0	9	11	10	0	0	0	0	0	0	0	2.2	2	-	[hh:hk:kk]
153	100026258	<hkxhk>	0	0	0	0	0	0	0	0	8	15	7	0	0	0	0	0	0	0	0.07	2	-	[hh:hk:kk]
154	100027191	<hkxhk>	0	0	0	0	0	0	0	0	10	14	6	0	0	0	0	0	0	0	1.2	2	-	[hh:hk:kk]
155	100027545	<hkxhk>	0	0	0	0	0	0	0	0	9	5	16	0	0	0	0	0	0	0	16.6	2	******	[hh:hk:kk]
156	100027913	<hkxhk>	0	0	0	0	0	0	0	0	7	18	5	0	0	0	0	0	0	0	1.47	2	-	[hh:hk:kk]
157	100029779	<hkxhk>	0	0	0	0	0	0	0	0	7	15	8	0	0	0	0	0	0	0	0.07	2	-	[hh:hk:kk]
158	100030554	<hkxhk>	0	0	0	0	0	0	0	0	8	16	6	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
159	100030945	<hkxhk>	0	0	0	0	0	0	0	0	8	16	6	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
160	100032830	<hkxhk>	0	0	0	0	0	0	0	0	10	14	6	0	0	0	0	0	0	0	1.2	2	-	[hh:hk:kk]
161	100041012	<hkxhk>	0	0	0	0	0	0	0	0	6	17	7	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
162	100041525	<hkxhk>	0	0	0	0	0	0	0	0	9	8	13	0	0	0	0	0	0	0	7.6	2	**	[hh:hk:kk]
163	100048019	<hkxhk>	0	0	0	0	0	0	0	0	7	14	9	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
164	100048615	<hkxhk>	0	0	0	0	0	0	0	0	6	16	8	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
165	100049015	<hkxhk>	0	0	0	0	0	0	0	0	7	14	9	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]

166	100050857	<hkxhk>	0	0	0	0	0	0	0	0	15	9	6	0	0	0	0	0	0	0	10.2	2	***	[hh:hk:kk]
167	100053687	<hkxhk>	0	0	0	0	0	0	0	0	7	16	7	0	0	0	0	0	0	0	0.13	2	-	[hh:hk:kk]
168	100059495	<hkxhk>	0	0	0	0	0	0	0	0	7	8	15	0	0	0	0	0	0	0	10.8	2	****	[hh:hk:kk]
169	100059518	<hkxhk>	0	0	0	0	0	0	0	0	6	15	9	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
170	100063034	<hkxhk>	0	0	0	0	0	0	0	0	15	2	13	0	0	0	0	0	0	0	22.8	2	*******	[hh:hk:kk]
171	100064190	<hkxhk>	0	0	0	0	0	0	0	0	9	15	6	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
172	100070876	<hkxhk>	0	0	0	0	0	0	0	0	6	16	8	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
173	100076502	<hkxhk>	0	0	0	0	0	0	0	0	6	17	7	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
174	Locus031	<nnxnp>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	13	16	1	0.31	1	-	[nn:np]
175	Locus073	<lmxll>	0	0	0	0	0	0	0	0	0	0	0	0	0	13	15	0	0	2	0.14	1	-	[ll:lm]
176	Locus087	<nnxnp>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	15	15	0	0	1	-	[nn:np]
177	100005112	<hkxhk>	0	0	0	0	0	0	0	0	11	5	14	0	0	0	0	0	0	0	13.93	2	*****	[hh:hk:kk]
178	100007751	<hkxhk>	0	0	0	0	0	0	0	0	9	7	14	0	0	0	0	0	0	0	10.2	2	***	[hh:hk:kk]
179	100008743	<hkxhk>	0	0	0	0	0	0	0	0	7	9	14	0	0	0	0	0	0	0	8.07	2	**	[hh:hk:kk]
180	100011352	<hkxhk>	0	0	0	0	0	0	0	0	9	10	11	0	0	0	0	0	0	0	3.6	2	-	[hh:hk:kk]
181	100013898	<hkxhk>	0	0	0	0	0	0	0	0	8	9	13	0	0	0	0	0	0	0	6.47	2	**	[hh:hk:kk]
182	100014155	<hkxhk>	0	0	0	0	0	0	0	0	12	12	6	0	0	0	0	0	0	0	3.6	2	-	[hh:hk:kk]
183	100014240	<hkxhk>	0	0	0	0	0	0	0	0	12	12	6	0	0	0	0	0	0	0	3.6	2	-	[hh:hk:kk]
184	100014765	<hkxhk>	0	0	0	0	0	0	0	0	10	11	9	0	0	0	0	0	0	0	2.2	2	-	[hh:hk:kk]
185	100015054	<hkxhk>	0	0	0	0	0	0	0	0	13	6	11	0	0	0	0	0	0	0	11.07	2	****	[hh:hk:kk]
186	100019563	<hkxhk>	0	0	0	0	0	0	0	0	12	12	6	0	0	0	0	0	0	0	3.6	2	-	[hh:hk:kk]
187	100020964	<hkxhk>	0	0	0	0	0	0	0	0	6	7	17	0	0	0	0	0	0	0	16.6	2	******	[hh:hk:kk]
188	100022472	<hkxhk>	0	0	0	0	0	0	0	0	9	6	15	0	0	0	0	0	0	0	13.2	2	****	[hh:hk:kk]
189	100023317	<hkxhk>	0	0	0	0	0	0	0	0	11	13	6	0	0	0	0	0	0	0	2.2	2	-	[hh:hk:kk]
190	100024254	<hkxhk>	0	0	0	0	0	0	0	0	7	15	8	0	0	0	0	0	0	0	0.07	2	-	[hh:hk:kk]
191	100024268	<hkxhk>	0	0	0	0	0	0	0	0	14	8	8	0	0	0	0	0	0	0	8.93	2	**	[hh:hk:kk]
192	100026969	<hkxhk>	0	0	0	0	0	0	0	0	9	10	11	0	0	0	0	0	0	0	3.6	2	-	[hh:hk:kk]
193	100029237	<hkxhk>	0	0	0	0	0	0	0	0	15	6	9	0	0	0	0	0	0	0	13.2	2	****	[hh:hk:kk]

194	100029757	<hkxhk>	0	0	0	0	0	0	0	0	13	6	11	0	0	0	0	0	0	0	11.07	2	****	[hh:hk:kk]
195	100030191	<hkxhk>	0	0	0	0	0	0	0	0	13	11	6	0	0	0	0	0	0	0	5.4	2	*	[hh:hk:kk]
196	100030699	<hkxhk>	0	0	0	0	0	0	0	0	7	10	13	0	0	0	0	0	0	0	5.73	2	*	[hh:hk:kk]
197	100030709	<hkxhk>	0	0	0	0	0	0	0	0	9	11	10	0	0	0	0	0	0	0	2.2	2	-	[hh:hk:kk]
198	100030798	<hkxhk>	0	0	0	0	0	0	0	0	11	5	14	0	0	0	0	0	0	0	13.93	2	*****	[hh:hk:kk]
199	100030819	<hkxhk>	0	0	0	0	0	0	0	0	12	9	9	0	0	0	0	0	0	0	5.4	2	*	[hh:hk:kk]
200	100031710	<hkxhk>	0	0	0	0	0	0	0	0	11	12	7	0	0	0	0	0	0	0	2.27	2	-	[hh:hk:kk]
201	100031897	<hkxhk>	0	0	0	0	0	0	0	0	9	10	11	0	0	0	0	0	0	0	3.6	2	-	[hh:hk:kk]
202	100032021	<hkxhk>	0	0	0	0	0	0	0	0	18	4	8	0	0	0	0	0	0	0	22.8	2	*******	[hh:hk:kk]
203	100032169	<hkxhk>	0	0	0	0	0	0	0	0	8	6	16	0	0	0	0	0	0	0	15.07	2	*****	[hh:hk:kk]
204	100033874	<hkxhk>	0	0	0	0	0	0	0	0	12	11	7	0	0	0	0	0	0	0	3.8	2	-	[hh:hk:kk]
205	100057062	<hkxhk>	0	0	0	0	0	0	0	0	10	8	12	0	0	0	0	0	0	0	6.8	2	**	[hh:hk:kk]
206	100072715	<hkxhk>	0	0	0	0	0	0	0	0	7	8	15	0	0	0	0	0	0	0	10.8	2	****	[hh:hk:kk]
207	100003779	<hkxhk>	0	0	0	0	0	0	0	0	7	13	10	0	0	0	0	0	0	0	1.13	2	-	[hh:hk:kk]
208	100004962	<hkxhk>	0	0	0	0	0	0	0	0	7	11	12	0	0	0	0	0	0	0	3.8	2	-	[hh:hk:kk]
209	100004975	<hkxhk>	0	0	0	0	0	0	0	0	10	11	9	0	0	0	0	0	0	0	2.2	2	-	[hh:hk:kk]
210	100006013	<hkxhk>	0	0	0	0	0	0	0	0	8	13	9	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
211	100006666	<hkxhk>	0	0	0	0	0	0	0	0	10	10	10	0	0	0	0	0	0	0	3.33	2	-	[hh:hk:kk]
212	100009038	<hkxhk>	0	0	0	0	0	0	0	0	8	12	10	0	0	0	0	0	0	0	1.47	2	-	[hh:hk:kk]
213	100009080	<hkxhk>	0	0	0	0	0	0	0	0	9	10	11	0	0	0	0	0	0	0	3.6	2	-	[hh:hk:kk]
214	100009384	<hkxhk>	0	0	0	0	0	0	0	0	7	11	12	0	0	0	0	0	0	0	3.8	2	-	[hh:hk:kk]
215	100009592	<hkxhk>	0	0	0	0	0	0	0	0	20	0	10	0	0	0	0	0	0	0	36.67	2	*******	[hh:hk:kk]
216	100010160	<hkxhk>	0	0	0	0	0	0	0	0	10	9	11	0	0	0	0	0	0	0	4.87	2	*	[hh:hk:kk]
217	100010338	<hkxhk>	0	0	0	0	0	0	0	0	10	11	9	0	0	0	0	0	0	0	2.2	2	-	[hh:hk:kk]
218	100010873	<hkxhk>	0	0	0	0	0	0	0	0	16	0	14	0	0	0	0	0	0	0	30.27	2	*******	[hh:hk:kk]
219	100011263	<hkxhk>	0	0	0	0	0	0	0	0	10	11	9	0	0	0	0	0	0	0	2.2	2	-	[hh:hk:kk]
220	100011501	<hkxhk>	0	0	0	0	0	0	0	0	7	12	11	0	0	0	0	0	0	0	2.27	2	-	[hh:hk:kk]
221	100012099	<hkxhk>	0	0	0	0	0	0	0	0	10	9	11	0	0	0	0	0	0	0	4.87	2	*	[hh:hk:kk]

222	100012241	<hkxhk>	0	0	0	0	0	0	0	0	11	2	17	0	0	0	0	0	0	0	24.93	2	*******	[hh:hk:kk]
223	100013104	<hkxhk>	0	0	0	0	0	0	0	0	18	1	11	0	0	0	0	0	0	0	29.4	2	*******	[hh:hk:kk]
224	100013614	<hkxhk>	0	0	0	0	0	0	0	0	10	6	14	0	0	0	0	0	0	0	11.87	2	****	[hh:hk:kk]
225	100013722	<hkxhk>	0	0	0	0	0	0	0	0	10	11	9	0	0	0	0	0	0	0	2.2	2	-	[hh:hk:kk]
226	100014282	<hkxhk>	0	0	0	0	0	0	0	0	7	13	10	0	0	0	0	0	0	0	1.13	2	-	[hh:hk:kk]
227	100015588	<hkxhk>	0	0	0	0	0	0	0	0	11	9	10	0	0	0	0	0	0	0	4.87	2	*	[hh:hk:kk]
228	100015926	<hkxhk>	0	0	0	0	0	0	0	0	10	3	17	0	0	0	0	0	0	0	22.47	2	*******	[hh:hk:kk]
229	100017861	<hkxhk>	0	0	0	0	0	0	0	0	11	8	11	0	0	0	0	0	0	0	6.53	2	**	[hh:hk:kk]
230	100018457	<hkxhk>	0	0	0	0	0	0	0	0	16	1	13	0	0	0	0	0	0	0	26.73	2	*******	[hh:hk:kk]
231	100019339	<hkxhk>	0	0	0	0	0	0	0	0	9	13	8	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
232	100019573	<hkxhk>	0	0	0	0	0	0	0	0	11	7	12	0	0	0	0	0	0	0	8.6	2	**	[hh:hk:kk]
233	100021332	<hkxhk>	0	0	0	0	0	0	0	0	11	10	9	0	0	0	0	0	0	0	3.6	2	-	[hh:hk:kk]
234	100023083	<hkxhk>	0	0	0	0	0	0	0	0	10	3	17	0	0	0	0	0	0	0	22.47	2	*******	[hh:hk:kk]
235	100024467	<hkxhk>	0	0	0	0	0	0	0	0	20	0	10	0	0	0	0	0	0	0	36.67	2	*******	[hh:hk:kk]
236	100024679	<hkxhk>	0	0	0	0	0	0	0	0	9	11	10	0	0	0	0	0	0	0	2.2	2	-	[hh:hk:kk]
237	100025173	<hkxhk>	0	0	0	0	0	0	0	0	11	10	9	0	0	0	0	0	0	0	3.6	2	-	[hh:hk:kk]
238	100025260	<hkxhk>	0	0	0	0	0	0	0	0	8	7	15	0	0	0	0	0	0	0	11.8	2	****	[hh:hk:kk]
239	100026162	<hkxhk>	0	0	0	0	0	0	0	0	10	12	8	0	0	0	0	0	0	0	1.47	2	-	[hh:hk:kk]
240	100026335	<hkxhk>	0	0	0	0	0	0	0	0	9	10	11	0	0	0	0	0	0	0	3.6	2	-	[hh:hk:kk]
241	100026855	<hkxhk>	0	0	0	0	0	0	0	0	12	3	15	0	0	0	0	0	0	0	19.8	2	*******	[hh:hk:kk]
242	100027694	<hkxhk>	0	0	0	0	0	0	0	0	13	6	11	0	0	0	0	0	0	0	11.07	2	****	[hh:hk:kk]
243	100028344	<hkxhk>	0	0	0	0	0	0	0	0	8	9	13	0	0	0	0	0	0	0	6.47	2	**	[hh:hk:kk]
244	100030512	<hkxhk>	0	0	0	0	0	0	0	0	12	10	8	0	0	0	0	0	0	0	4.4	2	-	[hh:hk:kk]
245	100031871	<hkxhk>	0	0	0	0	0	0	0	0	11	8	11	0	0	0	0	0	0	0	6.53	2	**	[hh:hk:kk]
246	100032255	<hkxhk>	0	0	0	0	0	0	0	0	10	12	8	0	0	0	0	0	0	0	1.47	2	-	[hh:hk:kk]
247	100034000	<hkxhk>	0	0	0	0	0	0	0	0	12	9	9	0	0	0	0	0	0	0	5.4	2	*	[hh:hk:kk]
248	100034115	<hkxhk>	0	0	0	0	0	0	0	0	9	12	9	0	0	0	0	0	0	0	1.2	2	-	[hh:hk:kk]
249	100045296	<hkxhk>	0	0	0	0	0	0	0	0	10	8	12	0	0	0	0	0	0	0	6.8	2	**	[hh:hk:kk]

250	Locus015	<abxcd>	3	11	9	6	0	0	0	0	0	0	0	0	0	0	0	0	0	1	5.07	3	-	[ac:ad:bc:bd]
251	100005289	<hkxhk>	0	0	0	0	0	0	0	0	11	0	19	0	0	0	0	0	0	0	34.27	2	*******	[hh:hk:kk]
252	100008017	<hkxhk>	0	0	0	0	0	0	0	0	8	12	10	0	0	0	0	0	0	0	1.47	2	-	[hh:hk:kk]
253	100009499	<hkxhk>	0	0	0	0	0	0	0	0	7	15	8	0	0	0	0	0	0	0	0.07	2	-	[hh:hk:kk]
254	100010290	<hkxhk>	0	0	0	0	0	0	0	0	9	9	12	0	0	0	0	0	0	0	5.4	2	*	[hh:hk:kk]
255	100011533	<hkxhk>	0	0	0	0	0	0	0	0	7	5	18	0	0	0	0	0	0	0	21.4	2	*******	[hh:hk:kk]
256	100015788	<hkxhk>	0	0	0	0	0	0	0	0	19	1	10	0	0	0	0	0	0	0	31.53	2	*******	[hh:hk:kk]
257	100016828	<hkxhk>	0	0	0	0	0	0	0	0	8	5	17	0	0	0	0	0	0	0	18.73	2	*******	[hh:hk:kk]
258	100017275	<hkxhk>	0	0	0	0	0	0	0	0	8	11	11	0	0	0	0	0	0	0	2.73	2	-	[hh:hk:kk]
259	100017379	<hkxhk>	0	0	0	0	0	0	0	0	9	3	18	0	0	0	0	0	0	0	24.6	2	*******	[hh:hk:kk]
260	100018480	<hkxhk>	0	0	0	0	0	0	0	0	9	10	11	0	0	0	0	0	0	0	3.6	2	-	[hh:hk:kk]
261	100018738	<hkxhk>	0	0	0	0	0	0	0	0	13	6	11	0	0	0	0	0	0	0	11.07	2	****	[hh:hk:kk]
262	100019834	<hkxhk>	0	0	0	0	0	0	0	0	17	1	12	0	0	0	0	0	0	0	27.8	2	*******	[hh:hk:kk]
263	100021135	<hkxhk>	0	0	0	0	0	0	0	0	7	15	8	0	0	0	0	0	0	0	0.07	2	-	[hh:hk:kk]
264	100022073	<hkxhk>	0	0	0	0	0	0	0	0	12	7	11	0	0	0	0	0	0	0	8.6	2	**	[hh:hk:kk]
265	100022346	<hkxhk>	0	0	0	0	0	0	0	0	9	12	9	0	0	0	0	0	0	0	1.2	2	-	[hh:hk:kk]
266	100025359	<hkxhk>	0	0	0	0	0	0	0	0	7	6	17	0	0	0	0	0	0	0	17.47	2	******	[hh:hk:kk]
267	100025727	<hkxhk>	0	0	0	0	0	0	0	0	12	7	11	0	0	0	0	0	0	0	8.6	2	**	[hh:hk:kk]
268	100026536	<hkxhk>	0	0	0	0	0	0	0	0	9	15	6	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
269	100027020	<hkxhk>	0	0	0	0	0	0	0	0	18	0	12	0	0	0	0	0	0	0	32.4	2	*******	[hh:hk:kk]
270	100027130	<hkxhk>	0	0	0	0	0	0	0	0	12	8	10	0	0	0	0	0	0	0	6.8	2	**	[hh:hk:kk]
271	100027827	<hkxhk>	0	0	0	0	0	0	0	0	7	11	12	0	0	0	0	0	0	0	3.8	2	-	[hh:hk:kk]
272	100031448	<hkxhk>	0	0	0	0	0	0	0	0	7	13	10	0	0	0	0	0	0	0	1.13	2	-	[hh:hk:kk]
273	100032973	<hkxhk>	0	0	0	0	0	0	0	0	8	4	18	0	0	0	0	0	0	0	22.8	2	*******	[hh:hk:kk]
274	100036121	<hkxhk>	0	0	0	0	0	0	0	0	8	10	12	0	0	0	0	0	0	0	4.4	2	-	[hh:hk:kk]
275	100037727	<hkxhk>	0	0	0	0	0	0	0	0	8	16	6	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
276	100039438	<hkxhk>	0	0	0	0	0	0	0	0	7	13	10	0	0	0	0	0	0	0	1.13	2	-	[hh:hk:kk]
277	100040322	<hkxhk>	0	0	0	0	0	0	0	0	10	13	7	0	0	0	0	0	0	0	1.13	2	-	[hh:hk:kk]

278	100041763	<hkxhk>	0	0	0	0	0	0	0	0	7	11	12	0	0	0	0	0	0	0	3.8	2	-	[hh:hk:kk]
279	100066627	<hkxhk>	0	0	0	0	0	0	0	0	18	0	12	0	0	0	0	0	0	0	32.4	2	*******	[hh:hk:kk]
280	Locus101	<abxcd>	8	7	7	7	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0.1	3	-	[ac:ad:bc:bd]
281	100004968	<hkxhk>	0	0	0	0	0	0	0	0	13	9	8	0	0	0	0	0	0	0	6.47	2	**	[hh:hk:kk]
282	100005451	<hkxhk>	0	0	0	0	0	0	0	0	8	14	8	0	0	0	0	0	0	0	0.13	2	-	[hh:hk:kk]
283	100006380	<hkxhk>	0	0	0	0	0	0	0	0	8	13	9	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
284	100006658	<hkxhk>	0	0	0	0	0	0	0	0	8	4	18	0	0	0	0	0	0	0	22.8	2	*******	[hh:hk:kk]
285	100006995	<hkxhk>	0	0	0	0	0	0	0	0	12	9	9	0	0	0	0	0	0	0	5.4	2	*	[hh:hk:kk]
286	100013111	<hkxhk>	0	0	0	0	0	0	0	0	12	9	9	0	0	0	0	0	0	0	5.4	2	*	[hh:hk:kk]
287	100013194	<hkxhk>	0	0	0	0	0	0	0	0	8	3	19	0	0	0	0	0	0	0	27.27	2	*******	[hh:hk:kk]
288	100013215	<hkxhk>	0	0	0	0	0	0	0	0	8	16	6	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
289	100014903	<hkxhk>	0	0	0	0	0	0	0	0	7	10	13	0	0	0	0	0	0	0	5.73	2	*	[hh:hk:kk]
290	100015908	<hkxhk>	0	0	0	0	0	0	0	0	7	4	19	0	0	0	0	0	0	0	25.73	2	*******	[hh:hk:kk]
291	100016084	<hkxhk>	0	0	0	0	0	0	0	0	11	12	7	0	0	0	0	0	0	0	2.27	2	-	[hh:hk:kk]
292	100016123	<hkxhk>	0	0	0	0	0	0	0	0	10	10	10	0	0	0	0	0	0	0	3.33	2	-	[hh:hk:kk]
293	100016142	<hkxhk>	0	0	0	0	0	0	0	0	8	9	13	0	0	0	0	0	0	0	6.47	2	**	[hh:hk:kk]
294	100016873	<hkxhk>	0	0	0	0	0	0	0	0	6	16	8	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
295	100016879	<hkxhk>	0	0	0	0	0	0	0	0	18	6	6	0	0	0	0	0	0	0	20.4	2	*******	[hh:hk:kk]
296	100017033	<hkxhk>	0	0	0	0	0	0	0	0	6	16	8	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
297	100017663	<hkxhk>	0	0	0	0	0	0	0	0	18	1	11	0	0	0	0	0	0	0	29.4	2	*******	[hh:hk:kk]
298	100020308	<hkxhk>	0	0	0	0	0	0	0	0	6	8	16	0	0	0	0	0	0	0	13.2	2	****	[hh:hk:kk]
299	100020373	<hkxhk>	0	0	0	0	0	0	0	0	10	11	9	0	0	0	0	0	0	0	2.2	2	-	[hh:hk:kk]
300	100021016	<hkxhk>	0	0	0	0	0	0	0	0	10	7	13	0	0	0	0	0	0	0	9.13	2	**	[hh:hk:kk]
301	100023727	<hkxhk>	0	0	0	0	0	0	0	0	10	11	9	0	0	0	0	0	0	0	2.2	2	-	[hh:hk:kk]
302	100027016	<hkxhk>	0	0	0	0	0	0	0	0	11	13	6	0	0	0	0	0	0	0	2.2	2	-	[hh:hk:kk]
303	100028019	<hkxhk>	0	0	0	0	0	0	0	0	7	17	6	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
304	100029849	<hkxhk>	0	0	0	0	0	0	0	0	10	12	8	0	0	0	0	0	0	0	1.47	2	-	[hh:hk:kk]
305	100031764	<hkxhk>	0	0	0	0	0	0	0	0	12	4	14	0	0	0	0	0	0	0	16.4	2	******	[hh:hk:kk]

306	100034182	<hkxhk>	0	0	0	0	0	0	0	0	13	9	8	0	0	0	0	0	0	0	6.47	2	**	[hh:hk:kk]
307	100040593	<hkxhk>	0	0	0	0	0	0	0	0	6	14	10	0	0	0	0	0	0	0	1.2	2	-	[hh:hk:kk]
308	100043628	<hkxhk>	0	0	0	0	0	0	0	0	10	11	9	0	0	0	0	0	0	0	2.2	2	-	[hh:hk:kk]
309	100050450	<hkxhk>	0	0	0	0	0	0	0	0	11	10	9	0	0	0	0	0	0	0	3.6	2	-	[hh:hk:kk]
310	Locus010	<efxeg>	0	0	0	0	6	7	10	6	0	0	0	0	0	0	0	0	0	1	1.48	3	-	[ee:ef:eg:fg]
311	Locus035	<nnxnp>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12	18	0	1.2	1	-	[nn:np]
312	100003098	<hkxhk>	0	0	0	0	0	0	0	0	8	12	10	0	0	0	0	0	0	0	1.47	2	-	[hh:hk:kk]
313	100006599	<hkxhk>	0	0	0	0	0	0	0	0	7	17	6	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
314	100007118	<hkxhk>	0	0	0	0	0	0	0	0	10	10	10	0	0	0	0	0	0	0	3.33	2	-	[hh:hk:kk]
315	100007430	<hkxhk>	0	0	0	0	0	0	0	0	10	10	10	0	0	0	0	0	0	0	3.33	2	-	[hh:hk:kk]
316	100011759	<hkxhk>	0	0	0	0	0	0	0	0	9	14	7	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
317	100011766	<hkxhk>	0	0	0	0	0	0	0	0	10	11	9	0	0	0	0	0	0	0	2.2	2	-	[hh:hk:kk]
318	100012477	<hkxhk>	0	0	0	0	0	0	0	0	9	5	16	0	0	0	0	0	0	0	16.6	2	******	[hh:hk:kk]
319	100013986	<hkxhk>	0	0	0	0	0	0	0	0	7	6	17	0	0	0	0	0	0	0	17.47	2	******	[hh:hk:kk]
320	100015079	<hkxhk>	0	0	0	0	0	0	0	0	8	6	16	0	0	0	0	0	0	0	15.07	2	*****	[hh:hk:kk]
321	100016376	<hkxhk>	0	0	0	0	0	0	0	0	6	10	14	0	0	0	0	0	0	0	7.6	2	**	[hh:hk:kk]
322	100020345	<hkxhk>	0	0	0	0	0	0	0	0	9	7	14	0	0	0	0	0	0	0	10.2	2	***	[hh:hk:kk]
323	100022382	<hkxhk>	0	0	0	0	0	0	0	0	6	13	11	0	0	0	0	0	0	0	2.2	2	-	[hh:hk:kk]
324	100023274	<hkxhk>	0	0	0	0	0	0	0	0	9	6	15	0	0	0	0	0	0	0	13.2	2	****	[hh:hk:kk]
325	100026362	<hkxhk>	0	0	0	0	0	0	0	0	9	10	11	0	0	0	0	0	0	0	3.6	2	-	[hh:hk:kk]
326	100026411	<hkxhk>	0	0	0	0	0	0	0	0	8	13	9	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
327	100027536	<hkxhk>	0	0	0	0	0	0	0	0	9	14	7	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
328	100028285	<hkxhk>	0	0	0	0	0	0	0	0	10	11	9	0	0	0	0	0	0	0	2.2	2	-	[hh:hk:kk]
329	100031175	<hkxhk>	0	0	0	0	0	0	0	0	10	11	9	0	0	0	0	0	0	0	2.2	2	-	[hh:hk:kk]
330	100031678	<hkxhk>	0	0	0	0	0	0	0	0	8	13	9	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
331	100032113	<hkxhk>	0	0	0	0	0	0	0	0	8	13	9	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
332	100032239	<hkxhk>	0	0	0	0	0	0	0	0	10	11	9	0	0	0	0	0	0	0	2.2	2	-	[hh:hk:kk]
333	100033340	<hkxhk>	0	0	0	0	0	0	0	0	13	11	6	0	0	0	0	0	0	0	5.4	2	*	[hh:hk:kk]

334	100052570	<hkxhk>	0	0	0	0	0	0	0	0	11	0	19	0	0	0	0	0	0	0	34.27	2	*******	[hh:hk:kk]
335	100052860	<hkxhk>	0	0	0	0	0	0	0	0	9	12	9	0	0	0	0	0	0	0	1.2	2	-	[hh:hk:kk]
336	100065594	<hkxhk>	0	0	0	0	0	0	0	0	8	13	9	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
337	100072097	<hkxhk>	0	0	0	0	0	0	0	0	10	10	10	0	0	0	0	0	0	0	3.33	2	-	[hh:hk:kk]
338	100079612	<hkxhk>	0	0	0	0	0	0	0	0	9	13	8	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
339	Locus008	<efxeg>	0	0	0	0	7	5	7	11	0	0	0	0	0	0	0	0	0	0	2.53	3	-	[ee:ef:eg:fg]
340	Locus072	<hkxhk>	0	0	0	0	0	0	0	0	7	12	10	0	0	0	0	0	0	1	1.48	2	-	[hh:hk:kk]
341	Locus141	<nnxnp>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	19	1	2.79	1	*	[nn:np]
342	100006762	<hkxhk>	0	0	0	0	0	0	0	0	8	15	7	0	0	0	0	0	0	0	0.07	2	-	[hh:hk:kk]
343	100007884	<hkxhk>	0	0	0	0	0	0	0	0	14	10	6	0	0	0	0	0	0	0	7.6	2	**	[hh:hk:kk]
344	100011233	<hkxhk>	0	0	0	0	0	0	0	0	9	15	6	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
345	100011570	<hkxhk>	0	0	0	0	0	0	0	0	7	16	7	0	0	0	0	0	0	0	0.13	2	-	[hh:hk:kk]
346	100011669	<hkxhk>	0	0	0	0	0	0	0	0	10	14	6	0	0	0	0	0	0	0	1.2	2	-	[hh:hk:kk]
347	100011984	<hkxhk>	0	0	0	0	0	0	0	0	7	7	16	0	0	0	0	0	0	0	13.93	2	*****	[hh:hk:kk]
348	100012448	<hkxhk>	0	0	0	0	0	0	0	0	8	13	9	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
349	100014932	<hkxhk>	0	0	0	0	0	0	0	0	8	15	7	0	0	0	0	0	0	0	0.07	2	-	[hh:hk:kk]
350	100018407	<hkxhk>	0	0	0	0	0	0	0	0	6	14	10	0	0	0	0	0	0	0	1.2	2	-	[hh:hk:kk]
351	100018479	<hkxhk>	0	0	0	0	0	0	0	0	7	5	18	0	0	0	0	0	0	0	21.4	2	*******	[hh:hk:kk]
352	100019500	<hkxhk>	0	0	0	0	0	0	0	0	8	15	7	0	0	0	0	0	0	0	0.07	2	-	[hh:hk:kk]
353	100020443	<hkxhk>	0	0	0	0	0	0	0	0	8	15	7	0	0	0	0	0	0	0	0.07	2	-	[hh:hk:kk]
354	100021462	<hkxhk>	0	0	0	0	0	0	0	0	7	14	9	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
355	100022569	<hkxhk>	0	0	0	0	0	0	0	0	13	12	5	0	0	0	0	0	0	0	5.47	2	*	[hh:hk:kk]
356	100022622	<hkxhk>	0	0	0	0	0	0	0	0	8	6	16	0	0	0	0	0	0	0	15.07	2	*****	[hh:hk:kk]
357	100023498	<hkxhk>	0	0	0	0	0	0	0	0	15	9	6	0	0	0	0	0	0	0	10.2	2	***	[hh:hk:kk]
358	100026067	<hkxhk>	0	0	0	0	0	0	0	0	10	3	17	0	0	0	0	0	0	0	22.47	2	*******	[hh:hk:kk]
359	100028988	<hkxhk>	0	0	0	0	0	0	0	0	7	4	19	0	0	0	0	0	0	0	25.73	2	*******	[hh:hk:kk]
360	100030006	<hkxhk>	0	0	0	0	0	0	0	0	7	13	10	0	0	0	0	0	0	0	1.13	2	-	[hh:hk:kk]
361	100031281	<hkxhk>	0	0	0	0	0	0	0	0	7	14	9	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]

362	100032623	<hkxhk>	0	0	0	0	0	0	0	0	8	14	8	0	0	0	0	0	0	0	0.13	2	-	[hh:hk:kk]
363	100034107	<hkxhk>	0	0	0	0	0	0	0	0	8	13	9	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
364	100036775	<hkxhk>	0	0	0	0	0	0	0	0	12	11	7	0	0	0	0	0	0	0	3.8	2	-	[hh:hk:kk]
365	100040007	<hkxhk>	0	0	0	0	0	0	0	0	7	7	16	0	0	0	0	0	0	0	13.93	2	*****	[hh:hk:kk]
366	100042316	<hkxhk>	0	0	0	0	0	0	0	0	8	17	5	0	0	0	0	0	0	0	1.13	2	-	[hh:hk:kk]
367	100043908	<hkxhk>	0	0	0	0	0	0	0	0	6	15	9	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
368	100054433	<hkxhk>	0	0	0	0	0	0	0	0	7	14	9	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
369	100057592	<hkxhk>	0	0	0	0	0	0	0	0	6	15	9	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
370	100059912	<hkxhk>	0	0	0	0	0	0	0	0	8	13	9	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
371	Locus042	<hkxhk>	0	0	0	0	0	0	0	0	8	13	8	0	0	0	0	0	0	1	0.31	2	-	[hh:hk:kk]
372	100003676	<hkxhk>	0	0	0	0	0	0	0	0	6	9	15	0	0	0	0	0	0	0	10.2	2	***	[hh:hk:kk]
373	100004245	<hkxhk>	0	0	0	0	0	0	0	0	6	17	7	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
374	100004380	<hkxhk>	0	0	0	0	0	0	0	0	14	8	8	0	0	0	0	0	0	0	8.93	2	**	[hh:hk:kk]
375	100011599	<hkxhk>	0	0	0	0	0	0	0	0	7	15	8	0	0	0	0	0	0	0	0.07	2	-	[hh:hk:kk]
376	100012240	<hkxhk>	0	0	0	0	0	0	0	0	12	6	12	0	0	0	0	0	0	0	10.8	2	****	[hh:hk:kk]
377	100016552	<hkxhk>	0	0	0	0	0	0	0	0	17	2	11	0	0	0	0	0	0	0	24.93	2	*******	[hh:hk:kk]
378	100017408	<hkxhk>	0	0	0	0	0	0	0	0	8	9	13	0	0	0	0	0	0	0	6.47	2	**	[hh:hk:kk]
379	100017463	<hkxhk>	0	0	0	0	0	0	0	0	17	7	6	0	0	0	0	0	0	0	16.6	2	******	[hh:hk:kk]
380	100018039	<hkxhk>	0	0	0	0	0	0	0	0	6	16	8	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
381	100018582	<hkxhk>	0	0	0	0	0	0	0	0	8	9	13	0	0	0	0	0	0	0	6.47	2	**	[hh:hk:kk]
382	100023695	<hkxhk>	0	0	0	0	0	0	0	0	11	13	6	0	0	0	0	0	0	0	2.2	2	-	[hh:hk:kk]
383	100024054	<hkxhk>	0	0	0	0	0	0	0	0	8	11	11	0	0	0	0	0	0	0	2.73	2	-	[hh:hk:kk]
384	100024129	<hkxhk>	0	0	0	0	0	0	0	0	18	4	8	0	0	0	0	0	0	0	22.8	2	*******	[hh:hk:kk]
385	100024382	<hkxhk>	0	0	0	0	0	0	0	0	7	16	7	0	0	0	0	0	0	0	0.13	2	-	[hh:hk:kk]
386	100024642	<hkxhk>	0	0	0	0	0	0	0	0	10	14	6	0	0	0	0	0	0	0	1.2	2	-	[hh:hk:kk]
387	100024826	<hkxhk>	0	0	0	0	0	0	0	0	6	15	9	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
388	100026244	<hkxhk>	0	0	0	0	0	0	0	0	10	13	7	0	0	0	0	0	0	0	1.13	2	-	[hh:hk:kk]
389	100026911	<hkxhk>	0	0	0	0	0	0	0	0	8	15	7	0	0	0	0	0	0	0	0.07	2	-	[hh:hk:kk]

390	100026982	<hkxhk>	0	0	0	0	0	0	0	0	9	8	13	0	0	0	0	0	0	0	7.6	2	**	[hh:hk:kk]
391	100029325	<hkxhk>	0	0	0	0	0	0	0	0	6	11	13	0	0	0	0	0	0	0	5.4	2	*	[hh:hk:kk]
392	100030623	<hkxhk>	0	0	0	0	0	0	0	0	7	15	8	0	0	0	0	0	0	0	0.07	2	-	[hh:hk:kk]
393	100030711	<hkxhk>	0	0	0	0	0	0	0	0	8	16	6	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
394	100031333	<hkxhk>	0	0	0	0	0	0	0	0	8	14	8	0	0	0	0	0	0	0	0.13	2	-	[hh:hk:kk]
395	100033092	<hkxhk>	0	0	0	0	0	0	0	0	6	10	14	0	0	0	0	0	0	0	7.6	2	**	[hh:hk:kk]
396	100034632	<hkxhk>	0	0	0	0	0	0	0	0	17	4	9	0	0	0	0	0	0	0	20.4	2	*******	[hh:hk:kk]
397	100035096	<hkxhk>	0	0	0	0	0	0	0	0	11	13	6	0	0	0	0	0	0	0	2.2	2	-	[hh:hk:kk]
398	100038815	<hkxhk>	0	0	0	0	0	0	0	0	5	6	19	0	0	0	0	0	0	0	23.87	2	*******	[hh:hk:kk]
399	100040641	<hkxhk>	0	0	0	0	0	0	0	0	9	14	7	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
400	100040999	<hkxhk>	0	0	0	0	0	0	0	0	14	8	8	0	0	0	0	0	0	0	8.93	2	**	[hh:hk:kk]
401	100003569	<hkxhk>	0	0	0	0	0	0	0	0	10	6	14	0	0	0	0	0	0	0	11.87	2	****	[hh:hk:kk]
402	100007031	<hkxhk>	0	0	0	0	0	0	0	0	8	13	9	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
403	100007542	<hkxhk>	0	0	0	0	0	0	0	0	6	15	9	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
404	100008602	<hkxhk>	0	0	0	0	0	0	0	0	10	10	10	0	0	0	0	0	0	0	3.33	2	-	[hh:hk:kk]
405	100014134	<hkxhk>	0	0	0	0	0	0	0	0	9	12	9	0	0	0	0	0	0	0	1.2	2	-	[hh:hk:kk]
406	100016643	<hkxhk>	0	0	0	0	0	0	0	0	7	16	7	0	0	0	0	0	0	0	0.13	2	-	[hh:hk:kk]
407	100020706	<hkxhk>	0	0	0	0	0	0	0	0	8	10	12	0	0	0	0	0	0	0	4.4	2	-	[hh:hk:kk]
408	100022371	<hkxhk>	0	0	0	0	0	0	0	0	14	0	16	0	0	0	0	0	0	0	30.27	2	*******	[hh:hk:kk]
409	100022524	<hkxhk>	0	0	0	0	0	0	0	0	17	9	4	0	0	0	0	0	0	0	16.07	2	******	[hh:hk:kk]
410	100023652	<hkxhk>	0	0	0	0	0	0	0	0	9	15	6	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
411	100024498	<hkxhk>	0	0	0	0	0	0	0	0	9	13	8	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
412	100025754	<hkxhk>	0	0	0	0	0	0	0	0	6	16	8	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
413	100028213	<hkxhk>	0	0	0	0	0	0	0	0	10	15	5	0	0	0	0	0	0	0	1.67	2	-	[hh:hk:kk]
414	100029549	<hkxhk>	0	0	0	0	0	0	0	0	12	13	5	0	0	0	0	0	0	0	3.8	2	-	[hh:hk:kk]
415	100029581	<hkxhk>	0	0	0	0	0	0	0	0	9	15	6	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
416	100030474	<hkxhk>	0	0	0	0	0	0	0	0	10	15	5	0	0	0	0	0	0	0	1.67	2	-	[hh:hk:kk]
417	100032842	<hkxhk>	0	0	0	0	0	0	0	0	10	15	5	0	0	0	0	0	0	0	1.67	2	-	[hh:hk:kk]

418	100033418	<hkxhk>	0	0	0	0	0	0	0	0	9	12	9	0	0	0	0	0	0	0	1.2	2	-	[hh:hk:kk]
419	100036544	<hkxhk>	0	0	0	0	0	0	0	0	9	9	12	0	0	0	0	0	0	0	5.4	2	*	[hh:hk:kk]
420	100039783	<hkxhk>	0	0	0	0	0	0	0	0	8	13	9	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
421	100003852	<hkxhk>	0	0	0	0	0	0	0	0	7	8	15	0	0	0	0	0	0	0	10.8	2	****	[hh:hk:kk]
422	100003945	<hkxhk>	0	0	0	0	0	0	0	0	10	10	10	0	0	0	0	0	0	0	3.33	2	-	[hh:hk:kk]
423	100004033	<hkxhk>	0	0	0	0	0	0	0	0	10	13	7	0	0	0	0	0	0	0	1.13	2	-	[hh:hk:kk]
424	100011339	<hkxhk>	0	0	0	0	0	0	0	0	7	10	13	0	0	0	0	0	0	0	5.73	2	*	[hh:hk:kk]
425	100012300	<hkxhk>	0	0	0	0	0	0	0	0	8	14	8	0	0	0	0	0	0	0	0.13	2	-	[hh:hk:kk]
426	100013538	<hkxhk>	0	0	0	0	0	0	0	0	9	10	11	0	0	0	0	0	0	0	3.6	2	-	[hh:hk:kk]
427	100014191	<hkxhk>	0	0	0	0	0	0	0	0	8	14	8	0	0	0	0	0	0	0	0.13	2	-	[hh:hk:kk]
428	100015684	<hkxhk>	0	0	0	0	0	0	0	0	9	13	8	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
429	100018230	<hkxhk>	0	0	0	0	0	0	0	0	9	13	8	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
430	100019291	<hkxhk>	0	0	0	0	0	0	0	0	15	2	13	0	0	0	0	0	0	0	22.8	2	*******	[hh:hk:kk]
431	100020864	<hkxhk>	0	0	0	0	0	0	0	0	7	10	13	0	0	0	0	0	0	0	5.73	2	*	[hh:hk:kk]
432	100024229	<hkxhk>	0	0	0	0	0	0	0	0	9	14	7	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
433	100026307	<hkxhk>	0	0	0	0	0	0	0	0	8	7	15	0	0	0	0	0	0	0	11.8	2	****	[hh:hk:kk]
434	100030224	<hkxhk>	0	0	0	0	0	0	0	0	7	16	7	0	0	0	0	0	0	0	0.13	2	-	[hh:hk:kk]
435	100032264	<hkxhk>	0	0	0	0	0	0	0	0	7	17	6	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
436	100036165	<hkxhk>	0	0	0	0	0	0	0	0	9	12	9	0	0	0	0	0	0	0	1.2	2	-	[hh:hk:kk]
437	100038767	<hkxhk>	0	0	0	0	0	0	0	0	9	7	14	0	0	0	0	0	0	0	10.2	2	***	[hh:hk:kk]
438	100040112	<hkxhk>	0	0	0	0	0	0	0	0	11	10	9	0	0	0	0	0	0	0	3.6	2	-	[hh:hk:kk]
439	100040285	<hkxhk>	0	0	0	0	0	0	0	0	9	11	10	0	0	0	0	0	0	0	2.2	2	-	[hh:hk:kk]
440	100043095	<hkxhk>	0	0	0	0	0	0	0	0	7	8	15	0	0	0	0	0	0	0	10.8	2	****	[hh:hk:kk]
441	100003475	<hkxhk>	0	0	0	0	0	0	0	0	11	10	9	0	0	0	0	0	0	0	3.6	2	-	[hh:hk:kk]
442	100012497	<hkxhk>	0	0	0	0	0	0	0	0	7	14	9	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
443	100013446	<hkxhk>	0	0	0	0	0	0	0	0	8	12	10	0	0	0	0	0	0	0	1.47	2	-	[hh:hk:kk]
444	100013539	<hkxhk>	0	0	0	0	0	0	0	0	8	16	6	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
445	100014845	<hkxhk>	0	0	0	0	0	0	0	0	12	9	9	0	0	0	0	0	0	0	5.4	2	*	[hh:hk:kk]

446	100015153	<hkxhk>	0	0	0	0	0	0	0	0	10	13	7	0	0	0	0	0	0	0	1.13	2	-	[hh:hk:kk]
447	100019310	<hkxhk>	0	0	0	0	0	0	0	0	5	7	18	0	0	0	0	0	0	0	19.8	2	*******	[hh:hk:kk]
448	100020945	<hkxhk>	0	0	0	0	0	0	0	0	10	2	18	0	0	0	0	0	0	0	26.8	2	*******	[hh:hk:kk]
449	100024877	<hkxhk>	0	0	0	0	0	0	0	0	9	7	14	0	0	0	0	0	0	0	10.2	2	***	[hh:hk:kk]
450	100025098	<hkxhk>	0	0	0	0	0	0	0	0	16	3	11	0	0	0	0	0	0	0	20.87	2	*******	[hh:hk:kk]
451	100026803	<hkxhk>	0	0	0	0	0	0	0	0	10	14	6	0	0	0	0	0	0	0	1.2	2	-	[hh:hk:kk]
452	100026827	<hkxhk>	0	0	0	0	0	0	0	0	7	13	10	0	0	0	0	0	0	0	1.13	2	-	[hh:hk:kk]
453	100027107	<hkxhk>	0	0	0	0	0	0	0	0	9	16	5	0	0	0	0	0	0	0	1.2	2	-	[hh:hk:kk]
454	100034614	<hkxhk>	0	0	0	0	0	0	0	0	7	13	10	0	0	0	0	0	0	0	1.13	2	-	[hh:hk:kk]
455	100045489	<hkxhk>	0	0	0	0	0	0	0	0	10	7	13	0	0	0	0	0	0	0	9.13	2	**	[hh:hk:kk]
456	100048112	<hkxhk>	0	0	0	0	0	0	0	0	8	11	11	0	0	0	0	0	0	0	2.73	2	-	[hh:hk:kk]
457	100071689	<hkxhk>	0	0	0	0	0	0	0	0	8	16	6	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
458	100077687	<hkxhk>	0	0	0	0	0	0	0	0	8	16	6	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
459	Locus069	<hkxhk>	0	0	0	0	0	0	0	0	10	16	4	0	0	0	0	0	0	0	2.53	2	-	[hh:hk:kk]
460	100002178	<hkxhk>	0	0	0	0	0	0	0	0	7	7	16	0	0	0	0	0	0	0	13.93	2	*****	[hh:hk:kk]
461	100003936	<hkxhk>	0	0	0	0	0	0	0	0	9	13	8	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
462	100009838	<hkxhk>	0	0	0	0	0	0	0	0	9	14	7	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
463	100014683	<hkxhk>	0	0	0	0	0	0	0	0	7	15	8	0	0	0	0	0	0	0	0.07	2	-	[hh:hk:kk]
464	100017801	<hkxhk>	0	0	0	0	0	0	0	0	6	14	10	0	0	0	0	0	0	0	1.2	2	-	[hh:hk:kk]
465	100019936	<hkxhk>	0	0	0	0	0	0	0	0	8	15	7	0	0	0	0	0	0	0	0.07	2	-	[hh:hk:kk]
466	100020602	<hkxhk>	0	0	0	0	0	0	0	0	11	12	7	0	0	0	0	0	0	0	2.27	2	-	[hh:hk:kk]
467	100021246	<hkxhk>	0	0	0	0	0	0	0	0	6	12	12	0	0	0	0	0	0	0	3.6	2	-	[hh:hk:kk]
468	100024334	<hkxhk>	0	0	0	0	0	0	0	0	13	1	16	0	0	0	0	0	0	0	26.73	2	*******	[hh:hk:kk]
469	100027553	<hkxhk>	0	0	0	0	0	0	0	0	8	15	7	0	0	0	0	0	0	0	0.07	2	-	[hh:hk:kk]
470	100030760	<hkxhk>	0	0	0	0	0	0	0	0	8	11	11	0	0	0	0	0	0	0	2.73	2	-	[hh:hk:kk]
471	100031805	<hkxhk>	0	0	0	0	0	0	0	0	10	13	7	0	0	0	0	0	0	0	1.13	2	-	[hh:hk:kk]
472	100032688	<hkxhk>	0	0	0	0	0	0	0	0	4	12	14	0	0	0	0	0	0	0	7.87	2	**	[hh:hk:kk]
473	100052401	<hkxhk>	0	0	0	0	0	0	0	0	6	7	17	0	0	0	0	0	0	0	16.6	2	******	[hh:hk:kk]

474	Locus022	<efxeg>	0	0	0	0	10	6	6	7	0	0	0	0	0	0	0	0	0	1	1.48	3	-	[ee:ef:eg:fg]
475	100001903	<hkxhk>	0	0	0	0	0	0	0	0	10	13	7	0	0	0	0	0	0	0	1.13	2	-	[hh:hk:kk]
476	100011059	<hkxhk>	0	0	0	0	0	0	0	0	11	8	11	0	0	0	0	0	0	0	6.53	2	**	[hh:hk:kk]
477	100012595	<hkxhk>	0	0	0	0	0	0	0	0	18	7	5	0	0	0	0	0	0	0	19.8	2	*******	[hh:hk:kk]
478	100017125	<hkxhk>	0	0	0	0	0	0	0	0	7	12	11	0	0	0	0	0	0	0	2.27	2	-	[hh:hk:kk]
479	100018498	<hkxhk>	0	0	0	0	0	0	0	0	9	14	7	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
480	100020016	<hkxhk>	0	0	0	0	0	0	0	0	10	14	6	0	0	0	0	0	0	0	1.2	2	-	[hh:hk:kk]
481	100020790	<hkxhk>	0	0	0	0	0	0	0	0	6	15	9	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
482	100025077	<hkxhk>	0	0	0	0	0	0	0	0	9	15	6	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
483	100026245	<hkxhk>	0	0	0	0	0	0	0	0	11	13	6	0	0	0	0	0	0	0	2.2	2	-	[hh:hk:kk]
484	100026813	<hkxhk>	0	0	0	0	0	0	0	0	9	5	16	0	0	0	0	0	0	0	16.6	2	******	[hh:hk:kk]
485	100028783	<hkxhk>	0	0	0	0	0	0	0	0	10	14	6	0	0	0	0	0	0	0	1.2	2	-	[hh:hk:kk]
486	100031279	<hkxhk>	0	0	0	0	0	0	0	0	7	17	6	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
487	100032270	<hkxhk>	0	0	0	0	0	0	0	0	9	12	9	0	0	0	0	0	0	0	1.2	2	-	[hh:hk:kk]
488	100034039	<hkxhk>	0	0	0	0	0	0	0	0	17	7	6	0	0	0	0	0	0	0	16.6	2	******	[hh:hk:kk]
489	Locus108	<efxeg>	0	0	0	0	6	6	9	8	0	0	0	0	0	0	0	0	0	1	0.93	3	-	[ee:ef:eg:fg]
490	100010643	<hkxhk>	0	0	0	0	0	0	0	0	10	8	12	0	0	0	0	0	0	0	6.8	2	**	[hh:hk:kk]
491	100010854	<hkxhk>	0	0	0	0	0	0	0	0	12	11	7	0	0	0	0	0	0	0	3.8	2	-	[hh:hk:kk]
492	100012465	<hkxhk>	0	0	0	0	0	0	0	0	6	17	7	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
493	100016919	<hkxhk>	0	0	0	0	0	0	0	0	7	14	9	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
494	100017539	<hkxhk>	0	0	0	0	0	0	0	0	11	13	6	0	0	0	0	0	0	0	2.2	2	-	[hh:hk:kk]
495	100021716	<hkxhk>	0	0	0	0	0	0	0	0	9	8	13	0	0	0	0	0	0	0	7.6	2	**	[hh:hk:kk]
496	100022444	<hkxhk>	0	0	0	0	0	0	0	0	17	2	11	0	0	0	0	0	0	0	24.93	2	*******	[hh:hk:kk]
497	100025166	<hkxhk>	0	0	0	0	0	0	0	0	12	10	8	0	0	0	0	0	0	0	4.4	2	-	[hh:hk:kk]
498	100026613	<hkxhk>	0	0	0	0	0	0	0	0	17	2	11	0	0	0	0	0	0	0	24.93	2	*******	[hh:hk:kk]
499	100026725	<hkxhk>	0	0	0	0	0	0	0	0	8	11	11	0	0	0	0	0	0	0	2.73	2	-	[hh:hk:kk]
500	100027165	<hkxhk>	0	0	0	0	0	0	0	0	17	3	10	0	0	0	0	0	0	0	22.47	2	*******	[hh:hk:kk]
501	100027600	<hkxhk>	0	0	0	0	0	0	0	0	12	13	5	0	0	0	0	0	0	0	3.8	2	-	[hh:hk:kk]

502	100033516	<hkxhk>	0	0	0	0	0	0	0	0	12	12	6	0	0	0	0	0	0	0	3.6	2	-	[hh:hk:kk]
503	100009021	<hkxhk>	0	0	0	0	0	0	0	0	7	15	8	0	0	0	0	0	0	0	0.07	2	-	[hh:hk:kk]
504	100011387	<hkxhk>	0	0	0	0	0	0	0	0	9	14	7	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
505	100014531	<hkxhk>	0	0	0	0	0	0	0	0	9	10	11	0	0	0	0	0	0	0	3.6	2	-	[hh:hk:kk]
506	100019260	<hkxhk>	0	0	0	0	0	0	0	0	10	13	7	0	0	0	0	0	0	0	1.13	2	-	[hh:hk:kk]
507	100020136	<hkxhk>	0	0	0	0	0	0	0	0	8	14	8	0	0	0	0	0	0	0	0.13	2	-	[hh:hk:kk]
508	100020668	<hkxhk>	0	0	0	0	0	0	0	0	8	12	10	0	0	0	0	0	0	0	1.47	2	-	[hh:hk:kk]
509	100021484	<hkxhk>	0	0	0	0	0	0	0	0	10	12	8	0	0	0	0	0	0	0	1.47	2	-	[hh:hk:kk]
510	100025283	<hkxhk>	0	0	0	0	0	0	0	0	9	14	7	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
511	100026940	<hkxhk>	0	0	0	0	0	0	0	0	9	9	12	0	0	0	0	0	0	0	5.4	2	*	[hh:hk:kk]
512	100028717	<hkxhk>	0	0	0	0	0	0	0	0	9	14	7	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
513	100029737	<hkxhk>	0	0	0	0	0	0	0	0	10	13	7	0	0	0	0	0	0	0	1.13	2	-	[hh:hk:kk]
514	100033626	<hkxhk>	0	0	0	0	0	0	0	0	9	14	7	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
515	Locus001	<nnxnp>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	13	17	0	0.53	1	-	[nn:np]
516	100004932	<hkxhk>	0	0	0	0	0	0	0	0	7	16	7	0	0	0	0	0	0	0	0.13	2	-	[hh:hk:kk]
517	100010396	<hkxhk>	0	0	0	0	0	0	0	0	10	11	9	0	0	0	0	0	0	0	2.2	2	-	[hh:hk:kk]
518	100016715	<hkxhk>	0	0	0	0	0	0	0	0	6	10	14	0	0	0	0	0	0	0	7.6	2	**	[hh:hk:kk]
519	100017588	<hkxhk>	0	0	0	0	0	0	0	0	8	14	8	0	0	0	0	0	0	0	0.13	2	-	[hh:hk:kk]
520	100022046	<hkxhk>	0	0	0	0	0	0	0	0	11	1	18	0	0	0	0	0	0	0	29.4	2	*******	[hh:hk:kk]
521	100027043	<hkxhk>	0	0	0	0	0	0	0	0	14	7	9	0	0	0	0	0	0	0	10.2	2	***	[hh:hk:kk]
522	100034043	<hkxhk>	0	0	0	0	0	0	0	0	8	12	10	0	0	0	0	0	0	0	1.47	2	-	[hh:hk:kk]
523	100034443	<hkxhk>	0	0	0	0	0	0	0	0	10	2	18	0	0	0	0	0	0	0	26.8	2	*******	[hh:hk:kk]
524	100037623	<hkxhk>	0	0	0	0	0	0	0	0	10	13	7	0	0	0	0	0	0	0	1.13	2	-	[hh:hk:kk]
525	100055758	<hkxhk>	0	0	0	0	0	0	0	0	10	10	10	0	0	0	0	0	0	0	3.33	2	-	[hh:hk:kk]
526	100062795	<hkxhk>	0	0	0	0	0	0	0	0	9	13	8	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
527	Locus088	<hkxhk>	0	0	0	0	0	0	0	0	4	16	10	0	0	0	0	0	0	0	2.53	2	-	[hh:hk:kk]
528	100009403	<hkxhk>	0	0	0	0	0	0	0	0	12	2	16	0	0	0	0	0	0	0	23.6	2	*******	[hh:hk:kk]
529	100009495	<hkxhk>	0	0	0	0	0	0	0	0	17	2	11	0	0	0	0	0	0	0	24.93	2	*******	[hh:hk:kk]

530	100012752	<hkxhk>	0	0	0	0	0	0	0	0	8	17	5	0	0	0	0	0	0	0	1.13	2	-	[hh:hk:kk]
531	100015635	<hkxhk>	0	0	0	0	0	0	0	0	7	15	8	0	0	0	0	0	0	0	0.07	2	-	[hh:hk:kk]
532	100019228	<hkxhk>	0	0	0	0	0	0	0	0	9	8	13	0	0	0	0	0	0	0	7.6	2	**	[hh:hk:kk]
533	100020651	<hkxhk>	0	0	0	0	0	0	0	0	11	4	15	0	0	0	0	0	0	0	17.2	2	******	[hh:hk:kk]
534	100022240	<hkxhk>	0	0	0	0	0	0	0	0	17	4	9	0	0	0	0	0	0	0	20.4	2	*******	[hh:hk:kk]
535	100022719	<hkxhk>	0	0	0	0	0	0	0	0	7	17	6	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
536	100026564	<hkxhk>	0	0	0	0	0	0	0	0	10	13	7	0	0	0	0	0	0	0	1.13	2	-	[hh:hk:kk]
537	100033440	<hkxhk>	0	0	0	0	0	0	0	0	8	12	10	0	0	0	0	0	0	0	1.47	2	-	[hh:hk:kk]
538	100033733	<hkxhk>	0	0	0	0	0	0	0	0	8	14	8	0	0	0	0	0	0	0	0.13	2	-	[hh:hk:kk]
539	100002378	<hkxhk>	0	0	0	0	0	0	0	0	8	16	6	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
540	100009195	<hkxhk>	0	0	0	0	0	0	0	0	9	6	15	0	0	0	0	0	0	0	13.2	2	****	[hh:hk:kk]
541	100010326	<hkxhk>	0	0	0	0	0	0	0	0	9	14	7	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
542	100011298	<hkxhk>	0	0	0	0	0	0	0	0	11	10	9	0	0	0	0	0	0	0	3.6	2	-	[hh:hk:kk]
543	100014562	<hkxhk>	0	0	0	0	0	0	0	0	9	14	7	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
544	100018155	<hkxhk>	0	0	0	0	0	0	0	0	11	13	6	0	0	0	0	0	0	0	2.2	2	-	[hh:hk:kk]
545	100019070	<hkxhk>	0	0	0	0	0	0	0	0	7	9	14	0	0	0	0	0	0	0	8.07	2	**	[hh:hk:kk]
546	100027115	<hkxhk>	0	0	0	0	0	0	0	0	14	1	15	0	0	0	0	0	0	0	26.2	2	*******	[hh:hk:kk]
547	100027989	<hkxhk>	0	0	0	0	0	0	0	0	12	4	14	0	0	0	0	0	0	0	16.4	2	******	[hh:hk:kk]
548	100028672	<hkxhk>	0	0	0	0	0	0	0	0	7	18	5	0	0	0	0	0	0	0	1.47	2	-	[hh:hk:kk]
549	100071907	<hkxhk>	0	0	0	0	0	0	0	0	10	7	13	0	0	0	0	0	0	0	9.13	2	**	[hh:hk:kk]
550	100009440	<hkxhk>	0	0	0	0	0	0	0	0	7	11	12	0	0	0	0	0	0	0	3.8	2	-	[hh:hk:kk]
551	100014803	<hkxhk>	0	0	0	0	0	0	0	0	9	10	11	0	0	0	0	0	0	0	3.6	2	-	[hh:hk:kk]
552	100016000	<hkxhk>	0	0	0	0	0	0	0	0	8	4	18	0	0	0	0	0	0	0	22.8	2	*******	[hh:hk:kk]
553	100017218	<hkxhk>	0	0	0	0	0	0	0	0	8	11	11	0	0	0	0	0	0	0	2.73	2	-	[hh:hk:kk]
554	100017862	<hkxhk>	0	0	0	0	0	0	0	0	9	10	11	0	0	0	0	0	0	0	3.6	2	-	[hh:hk:kk]
555	100017985	<hkxhk>	0	0	0	0	0	0	0	0	7	13	10	0	0	0	0	0	0	0	1.13	2	-	[hh:hk:kk]
556	100023732	<hkxhk>	0	0	0	0	0	0	0	0	7	12	11	0	0	0	0	0	0	0	2.27	2	-	[hh:hk:kk]
557	100024315	<hkxhk>	0	0	0	0	0	0	0	0	8	14	8	0	0	0	0	0	0	0	0.13	2	-	[hh:hk:kk]

558	100028512	<hkxhk>	0	0	0	0	0	0	0	0	12	6	12	0	0	0	0	0	0	0	10.8	2	****	[hh:hk:kk]
559	100030282	<hkxhk>	0	0	0	0	0	0	0	0	11	12	7	0	0	0	0	0	0	0	2.27	2	-	[hh:hk:kk]
560	100005578	<hkxhk>	0	0	0	0	0	0	0	0	8	9	13	0	0	0	0	0	0	0	6.47	2	**	[hh:hk:kk]
561	100014998	<hkxhk>	0	0	0	0	0	0	0	0	12	6	12	0	0	0	0	0	0	0	10.8	2	****	[hh:hk:kk]
562	100016596	<hkxhk>	0	0	0	0	0	0	0	0	9	4	17	0	0	0	0	0	0	0	20.4	2	*******	[hh:hk:kk]
563	100019999	<hkxhk>	0	0	0	0	0	0	0	0	5	19	6	0	0	0	0	0	0	0	2.2	2	-	[hh:hk:kk]
564	100021561	<hkxhk>	0	0	0	0	0	0	0	0	9	7	14	0	0	0	0	0	0	0	10.2	2	***	[hh:hk:kk]
565	100023222	<hkxhk>	0	0	0	0	0	0	0	0	8	12	10	0	0	0	0	0	0	0	1.47	2	-	[hh:hk:kk]
566	100027636	<hkxhk>	0	0	0	0	0	0	0	0	10	6	14	0	0	0	0	0	0	0	11.87	2	****	[hh:hk:kk]
567	100036947	<hkxhk>	0	0	0	0	0	0	0	0	12	12	6	0	0	0	0	0	0	0	3.6	2	-	[hh:hk:kk]
568	100007258	<hkxhk>	0	0	0	0	0	0	0	0	7	14	9	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
569	100012971	<hkxhk>	0	0	0	0	0	0	0	0	7	14	9	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
570	100014217	<hkxhk>	0	0	0	0	0	0	0	0	11	10	9	0	0	0	0	0	0	0	3.6	2	-	[hh:hk:kk]
571	100022806	<hkxhk>	0	0	0	0	0	0	0	0	7	6	17	0	0	0	0	0	0	0	17.47	2	******	[hh:hk:kk]
572	100027270	<hkxhk>	0	0	0	0	0	0	0	0	7	14	9	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
573	100038458	<hkxhk>	0	0	0	0	0	0	0	0	5	16	9	0	0	0	0	0	0	0	1.2	2	-	[hh:hk:kk]
574	100050970	<hkxhk>	0	0	0	0	0	0	0	0	7	14	9	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
575	100009445	<hkxhk>	0	0	0	0	0	0	0	0	9	12	9	0	0	0	0	0	0	0	1.2	2	-	[hh:hk:kk]
576	100013515	<hkxhk>	0	0	0	0	0	0	0	0	6	16	8	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
577	100020498	<hkxhk>	0	0	0	0	0	0	0	0	7	16	7	0	0	0	0	0	0	0	0.13	2	-	[hh:hk:kk]
578	100023272	<hkxhk>	0	0	0	0	0	0	0	0	8	13	9	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
579	100026454	<hkxhk>	0	0	0	0	0	0	0	0	7	18	5	0	0	0	0	0	0	0	1.47	2	-	[hh:hk:kk]
580	100030763	<hkxhk>	0	0	0	0	0	0	0	0	9	11	10	0	0	0	0	0	0	0	2.2	2	-	[hh:hk:kk]
581	100054662	<hkxhk>	0	0	0	0	0	0	0	0	6	15	9	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
582	100010930	<hkxhk>	0	0	0	0	0	0	0	0	10	3	17	0	0	0	0	0	0	0	22.47	2	*******	[hh:hk:kk]
583	100017048	<hkxhk>	0	0	0	0	0	0	0	0	10	12	8	0	0	0	0	0	0	0	1.47	2	-	[hh:hk:kk]
584	100031005	<hkxhk>	0	0	0	0	0	0	0	0	10	6	14	0	0	0	0	0	0	0	11.87	2	****	[hh:hk:kk]
585	100032803	<hkxhk>	0	0	0	0	0	0	0	0	7	12	11	0	0	0	0	0	0	0	2.27	2	-	[hh:hk:kk]

586	100004168	<hkxhk>	0	0	0	0	0	0	0	0	7	12	11	0	0	0	0	0	0	0	2.27	2	-	[hh:hk:kk]
587	100005690	<hkxhk>	0	0	0	0	0	0	0	0	8	10	12	0	0	0	0	0	0	0	4.4	2	-	[hh:hk:kk]
588	100007522	<hkxhk>	0	0	0	0	0	0	0	0	8	9	13	0	0	0	0	0	0	0	6.47	2	**	[hh:hk:kk]
589	100007756	<hkxhk>	0	0	0	0	0	0	0	0	9	11	10	0	0	0	0	0	0	0	2.2	2	-	[hh:hk:kk]
590	100009221	<hkxhk>	0	0	0	0	0	0	0	0	6	9	15	0	0	0	0	0	0	0	10.2	2	***	[hh:hk:kk]
591	100009764	<hkxhk>	0	0	0	0	0	0	0	0	9	5	16	0	0	0	0	0	0	0	16.6	2	******	[hh:hk:kk]
592	100010581	<hkxhk>	0	0	0	0	0	0	0	0	8	9	13	0	0	0	0	0	0	0	6.47	2	**	[hh:hk:kk]
593	100011028	<hkxhk>	0	0	0	0	0	0	0	0	6	17	7	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
594	100011275	<hkxhk>	0	0	0	0	0	0	0	0	6	12	12	0	0	0	0	0	0	0	3.6	2	-	[hh:hk:kk]
595	100012029	<hkxhk>	0	0	0	0	0	0	0	0	13	1	16	0	0	0	0	0	0	0	26.73	2	*******	[hh:hk:kk]
596	100012096	<hkxhk>	0	0	0	0	0	0	0	0	9	7	14	0	0	0	0	0	0	0	10.2	2	***	[hh:hk:kk]
597	100013463	<hkxhk>	0	0	0	0	0	0	0	0	7	12	11	0	0	0	0	0	0	0	2.27	2	-	[hh:hk:kk]
598	100014226	<hkxhk>	0	0	0	0	0	0	0	0	8	12	10	0	0	0	0	0	0	0	1.47	2	-	[hh:hk:kk]
599	100014713	<hkxhk>	0	0	0	0	0	0	0	0	11	11	8	0	0	0	0	0	0	0	2.73	2	-	[hh:hk:kk]
600	100015298	<hkxhk>	0	0	0	0	0	0	0	0	8	6	16	0	0	0	0	0	0	0	15.07	2	*****	[hh:hk:kk]
601	100016156	<hkxhk>	0	0	0	0	0	0	0	0	9	5	16	0	0	0	0	0	0	0	16.6	2	******	[hh:hk:kk]
602	100016912	<hkxhk>	0	0	0	0	0	0	0	0	7	6	17	0	0	0	0	0	0	0	17.47	2	******	[hh:hk:kk]
603	100017066	<hkxhk>	0	0	0	0	0	0	0	0	8	15	7	0	0	0	0	0	0	0	0.07	2	-	[hh:hk:kk]
604	100017148	<hkxhk>	0	0	0	0	0	0	0	0	13	3	14	0	0	0	0	0	0	0	19.27	2	*******	[hh:hk:kk]
605	100017321	<hkxhk>	0	0	0	0	0	0	0	0	10	8	12	0	0	0	0	0	0	0	6.8	2	**	[hh:hk:kk]
606	100018055	<hkxhk>	0	0	0	0	0	0	0	0	9	9	12	0	0	0	0	0	0	0	5.4	2	*	[hh:hk:kk]
607	100019463	<hkxhk>	0	0	0	0	0	0	0	0	7	18	5	0	0	0	0	0	0	0	1.47	2	-	[hh:hk:kk]
608	100020294	<hkxhk>	0	0	0	0	0	0	0	0	7	7	16	0	0	0	0	0	0	0	13.93	2	*****	[hh:hk:kk]
609	100022789	<hkxhk>	0	0	0	0	0	0	0	0	7	18	5	0	0	0	0	0	0	0	1.47	2	-	[hh:hk:kk]
610	100022885	<hkxhk>	0	0	0	0	0	0	0	0	6	16	8	0	0	0	0	0	0	0	0.4	2	-	[hh:hk:kk]
611	100023134	<hkxhk>	0	0	0	0	0	0	0	0	8	10	12	0	0	0	0	0	0	0	4.4	2	-	[hh:hk:kk]
612	100023897	<hkxhk>	0	0	0	0	0	0	0	0	7	7	16	0	0	0	0	0	0	0	13.93	2	*****	[hh:hk:kk]
613	100025241	<hkxhk>	0	0	0	0	0	0	0	0	12	8	10	0	0	0	0	0	0	0	6.8	2	**	[hh:hk:kk]

614	100025807	<hkxhk>	0	0	0	0	0	0	0	0	8	12	10	0	0	0	0	0	0	0	1.47	2	-	[hh:hk:kk]
615	100026352	<hkxhk>	0	0	0	0	0	0	0	0	7	7	16	0	0	0	0	0	0	0	13.93	2	*****	[hh:hk:kk]
616	100027794	<hkxhk>	0	0	0	0	0	0	0	0	8	10	12	0	0	0	0	0	0	0	4.4	2	-	[hh:hk:kk]
617	100028149	<hkxhk>	0	0	0	0	0	0	0	0	8	8	14	0	0	0	0	0	0	0	8.93	2	**	[hh:hk:kk]
618	100028785	<hkxhk>	0	0	0	0	0	0	0	0	8	12	10	0	0	0	0	0	0	0	1.47	2	-	[hh:hk:kk]
619	100028859	<hkxhk>	0	0	0	0	0	0	0	0	14	1	15	0	0	0	0	0	0	0	26.2	2	*******	[hh:hk:kk]
620	100029145	<hkxhk>	0	0	0	0	0	0	0	0	14	11	5	0	0	0	0	0	0	0	7.53	2	**	[hh:hk:kk]
621	100029491	<hkxhk>	0	0	0	0	0	0	0	0	13	8	9	0	0	0	0	0	0	0	7.6	2	**	[hh:hk:kk]
622	100029895	<hkxhk>	0	0	0	0	0	0	0	0	10	7	13	0	0	0	0	0	0	0	9.13	2	**	[hh:hk:kk]
623	100030002	<hkxhk>	0	0	0	0	0	0	0	0	8	17	5	0	0	0	0	0	0	0	1.13	2	-	[hh:hk:kk]
624	100031543	<hkxhk>	0	0	0	0	0	0	0	0	8	11	11	0	0	0	0	0	0	0	2.73	2	-	[hh:hk:kk]
625	100033364	<hkxhk>	0	0	0	0	0	0	0	0	6	11	13	0	0	0	0	0	0	0	5.4	2	*	[hh:hk:kk]
626	100033504	<hkxhk>	0	0	0	0	0	0	0	0	8	9	13	0	0	0	0	0	0	0	6.47	2	**	[hh:hk:kk]
627	100035308	<hkxhk>	0	0	0	0	0	0	0	0	11	9	10	0	0	0	0	0	0	0	4.87	2	*	[hh:hk:kk]
628	100037955	<hkxhk>	0	0	0	0	0	0	0	0	14	10	6	0	0	0	0	0	0	0	7.6	2	**	[hh:hk:kk]
629	100043754	<hkxhk>	0	0	0	0	0	0	0	0	12	5	13	0	0	0	0	0	0	0	13.4	2	****	[hh:hk:kk]
630	100045448	<hkxhk>	0	0	0	0	0	0	0	0	12	10	8	0	0	0	0	0	0	0	4.4	2	-	[hh:hk:kk]
631	100047150	<hkxhk>	0	0	0	0	0	0	0	0	14	3	13	0	0	0	0	0	0	0	19.27	2	*******	[hh:hk:kk]
632	100048168	<hkxhk>	0	0	0	0	0	0	0	0	8	13	9	0	0	0	0	0	0	0	0.6	2	-	[hh:hk:kk]
633	100065700	<hkxhk>	0	0	0	0	0	0	0	0	8	6	16	0	0	0	0	0	0	0	15.07	2	*****	[hh:hk:kk]
634	Locus025	<abxcd>	4	13	9	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7.6	3	*	[ac:ad:bc:bd]
635	100004221	<hkxhk>	0	0	0	0	0	0	0	0	13	12	5	0	0	0	0	0	0	0	5.47	2	*	[hh:hk:kk]
636	100004835	<hkxhk>	0	0	0	0	0	0	0	0	16	10	4	0	0	0	0	0	0	0	12.93	2	****	[hh:hk:kk]
637	100028289	<hkxhk>	0	0	0	0	0	0	0	0	15	1	14	0	0	0	0	0	0	0	26.2	2	*******	[hh:hk:kk]
638	100037906	<hkxhk>	0	0	0	0	0	0	0	0	7	12	11	0	0	0	0	0	0	0	2.27	2	-	[hh:hk:kk]
639	100045831	<hkxhk>	0	0	0	0	0	0	0	0	6	10	14	0	0	0	0	0	0	0	7.6	2	**	[hh:hk:kk]
640	100046903	<hkxhk>	0	0	0	0	0	0	0	0	3	16	11	0	0	0	0	0	0	0	4.4	2	-	[hh:hk:kk]
641	Locus145	<nnxnp>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	17	13	0	0.53	1	-	[nn:np]

Appendix 7: List of 290 date palm accessions used in this study with laboratory code, gender, country of origin and observed allelic size range.

No.	Lab Code	Accession Name	Gender	Country of Origin	Observed Size (bp)	No.	Lab Code	Accession Name	Gender	Country of Origin	Observed Size (bp)
1	QB	Qash Bunaring*a*	Female	Oman/GB	339/339	27	NBdh	Nashu Ba'oodh	Female	Oman/GB	339/339
2	MA	Miznag Ahmar	Female	Oman/GB	339/339	28	Mwaz	Mawaz	Female	Oman/GB	339/339
3	QT	Qash Tabaq	Female	Oman/GB	339/339	29	NGn	Nashu Ghoson	Female	Oman/GB	339/339
4	Frd	Fard	Female	Oman/GB	339/339	30	QSuhi	Qash Suwaih	Female	Oman/GB	339/339
5	Ksb	Khsab	Female	Oman/GB	339/339	31	QQat	Qash Qataari	Female	Oman/GB	339/339
6	QH	Qash Hagr	Female	Oman/GB	339/339	32	QAAB	Qash Ain Al Bakar	Female	Oman/GB	339/339
7	QBedsit	Qash Beladsait	Female	Oman/GB	339/339	33	QATb	Qash Al Teeb	Female	Oman/GB	339/339
8	KH	Khinaizi Halaw	Female	Oman/GB	339/339	**34**	**QHreer**	**Qash Hareer**	**Female**	**Oman/GB**	**339/346**
9	HH	Hilali Hassa	Female	Oman/GB	339/339	35	HilA	Hilali Ahmer	Female	Oman/GB	339/339
10	Zbd	Zabad	Female	Oman/GB	339/339	36	QZmel	Qash Zamel	Female	Oman/GB	339/339
11	QN	Qash Na'im	Female	Oman/GB	339/339	37	Ghrbo	Ghrabo	Female	Oman/GB	339/339
12	QHw	Qash Halaw	Female	Oman/GB	339/339	38	Shari	Shaeeri	Female	Oman/GB	339/339
13	QBuR	Qash Bu Rashid	Female	Oman/GB	339/339	39	QAKbal	Qash Abu Keebal	Female	Oman/GB	339/339
14	Bentm	Bentaami	Female	Oman/GB	339/339	40	Bydh	Bayadh	Female	Oman/GB	339/339
15	Frid	Farid	Female	Oman/GB	339/339	41	QMshrb	Qash Mushrab	Female	Oman/GB	339/339
16	Batsh	Battash	Female	Oman/GB	339/339	42	QGhsn	Ghssan Qash	Female	Oman/GB	339/339
17	KA	Khinaizi Arabi	Female	Oman/GB	339/339	43	QHmd	Hmdan Qash	Female	Oman/GB	339/339
18	QQrot	Qash Qaroot	Female	Oman/GB	339/339	44	QHbeb	Habeeb Qash	Female	Oman/GB	339/339
19	Klbi	Kalbi	Female	Oman/GB	339/339	45	QASf	Qash Abu Saif	Female	Oman/GB	339/339
20	UmAs	Umm Alssila	Female	Oman/GB	339/339	46	QHbsh	Habisha Qash	Female	Oman/GB	339/339
21	Mzrm	Mazrm	Female	Oman/GB	339/339	**47**	**Hsas**	**Hessas**	**Female**	**Oman/GB**	**339/346**
22	Minz	Minaz	Female	Oman/GB	339/339	48	QMnzf	Manzef Qash	Female	Oman/GB	339/339
23	Mhlb	Mahlabbi	Female	Oman/GB	339/339	49	Zad	Zaad	Female	Oman/GB	339/339
24	Khmr	Khamri	Female	Oman/GB	339/339	50	Nghl	Naghal	Female	Oman/GB	339/339
25	Shbrot	Shabroot	Female	Oman/GB	339/339	51	QQntra	Qantara Qash	Female	Oman/GB	339/339
26	Mek	Mekhildi	Female	Oman/GB	339/339	52	Man	Ma'an	Female	Oman/GB	339/339
53	Bunrng	Bunaringa	Female	Oman/GB	339/339	84	QSba	Qash Sba	Female	Oman/GB	339/339
54	Jbri	Jebri	Female	Oman/GB	339/339	85	QHshm	Qash Hiyshmi	Female	Oman/GB	339/339
55	HilMkrn	Hilali Makran	Female	Oman/GB	339/339	86	QATbi	Qash Al Teebi	Female	Oman/GB	339/339
56	QHbGb	Qash Humaid G.	Female	Oman/GB	339/339	87	QRsh	Qash Rsheed	Female	Oman/GB	339/339
57	Menh	Menhi	Female	Oman/GB	339/339	88	NAkzm	Nashu Al khzma	Female	Oman/GB	339/339
58	Mbsl	Mebseli	Female	Oman/GB	339/339	89	Berzb	Berzeban	Female	Oman/GB	339/339
59	Brshi	Bershi	Female	Oman/GB	339/339	90	NAWk	Nashu Al Wakhrh	Female	Oman/GB	339/339
60	Res	Rees	Female	Oman/GB	339/339	91	NghtKhls	Naghlt Khalas	Female	Oman/GB	339/339

61	QAHrem	Qash Al Hareem	Female	Oman/GB	339/339	92	KhlsOm	Khalas Oman	Female	Oman/GB	339/339
62	NEn	Nashu Ewan	Female	Oman/GB	339/339	93	QGhrof	Qash Gha'roof	Female	Oman/GB	339/339
63	MDn	Malkt Deeni	Female	Oman/GB	339/339	94	QGhfan	Qash Ghafan	Female	Oman/GB	339/339
64	Mysi	Mayasi	Female	Oman/GB	339/339	95	QNsrh	Qash Nas'rah	Female	Oman/GB	339/339
65	Hdgi	Hadaqi	Female	Oman/GB	339/339	96	QGheny	Qash Gheniyah	Female	Oman/GB	339/339
66	AQren	Abu Qareen	Female	Oman/GB	339/339	97	QNwhi	Qash Nwaihi	Female	Oman/GB	339/339
67	Selh	Selahni	Female	Oman/GB	339/339	98	QAMsbt	Qash Al Masbt	Female	Oman/GB	339/339
68	QSrh	Qash Sahrh	Female	Oman/GB	339/339	99	QtNghl	Qasht Naghal	Female	Oman/GB	339/339
69	NbotS	Naboot Saif	Female	Oman/GB	339/339	100	QAli	Qash Ali	Female	Oman/GB	339/339
70	QAY	Qash Al Yamam	Female	Oman/GB	339/339	101	QHreb	Qash Hareb	Female	Oman/GB	339/339
71	QHmd	Qash Humaid	Female	Oman/GB	339/339	102	QNasir	Qash Nasir	Female	Oman/GB	339/339
72	**NLul**	**Naghl Lulu**	**Female**	**Oman/GB**	**339/346**	103	QSfiyh	Qash Safiyh	Female	Oman/GB	339/339
73	QARb	Qash Al Rabeea	Female	Oman/GB	339/339	104	QFkrh	Qash Fakhrh	Female	Oman/GB	339/339
74	Mzn	Mazni	Female	Oman/GB	339/339	105	QSwid	Qash Suwaid	Female	Oman/GB	339/339
75	QAwb	Qash Al wa'b	Female	Oman/GB	339/339	106	QBaOr	Qash Ba'Omar	Female	Oman/GB	339/339
76	QASgh	Qash Al Saghiay	Female	Oman/GB	339/339	107	NShms	Nashu Shamiss	Female	Oman/GB	339/339
77	AAAoq	Abu Al Audooq	Female	Oman/GB	339/339	108	Shahl	Shahl	Female	Oman/GB	339/339
78	QShf	Qash Shafer	Female	Oman/GB	339/339	109	Rmli	Ramli	Female	Oman/GB	339/339
79	Alak	Alak	Female	Oman/GB	339/339	110	Shihm	Shiham	Female	Oman/GB	339/339
80	Snah	Snah	Female	Oman/GB	339/339	111	Seedi	Seedi	Female	Oman/GB	339/339
81	QHatm	Qash Hatami	Female	Oman/GB	339/339	112	Khshkr	Khashkar	Female	Oman/GB	339/339
82	QKtrh	Qash Katerh	Female	Oman/GB	339/339	113	NSleh	Nashu Saleh	Female	Oman/GB	339/339
83	QARbab	Qash Al Rabab	Female	Oman/GB	339/339	114	NMneh	Nashu Maneh	Female	Oman/GB	339/339
115	Lulu	Lulu	Female	Oman/GB	339/339	146	HiliOm	Hilali Omani	Female	Oman/GB	339/339
116	Rabai	Rabai	Female	Oman/GB	339/339	147	QGhnuw	Qash Ghinuwi	Female	Oman/GB	339/339
117	QSwlm	Suwailim Qash	Female	Oman/GB	339/339	148	QASAArz	Qash A'Saba Al Ar	Female	Oman/GB	339/339
118	Bata	Bata	Female	Oman/GB	339/339	149	Serna	Serna	Female	Oman/GB	339/339
119	Brny	Barny	Female	Oman/GB	339/339	150	KhlsAZrh	Khalas Al Zahra	Female	Oman/GB	339/339
120	NFhod	Nashu Fahood	Female	Oman/GB	339/339	151	Haithm	Haithami	Female	Oman/GB	339/339
121	Mutrh	Muttrahi	Female	Oman/GB	339/339	152	Kh1-GB	Khori 1	Male	Oman/GB	339/346
122	Bidaa	Bidaa	Female	Oman/GB	339/339	153	Kh2-GB	Khori 2	Male	Oman/GB	339/346
123	Medrk	Medairki	Female	Oman/GB	339/339	154	Kh3-GB	Khori 3	Male	Oman/GB	339/346
124	Kibkb	Kibkab	Female	Oman/GB	339/339	155	Kh4-GB	Khori 4	Male	Oman/GB	339/346
125	**Huzfh**	**Huzaifah**	**Female**	**Oman/GB**	**339/346**	156	Ngh1-GB	Naghayli 1	Male	Oman/GB	339/346
126	NAKshy	Nashu Al Khash	Female	Oman/GB	339/339	157	Ngh2-GB	Naghayli 2	Male	Oman/GB	339/346
127	Hawam	Hawam	Female	Oman/GB	339/339	158	Ngh3-GB	Naghayli 3	Male	Oman/GB	339/346
128	Qdmi	Qadmi	Female	Oman/GB	339/339	159	Md-GB	Medgahdel	Male	Oman/GB	339/346
129	QGmh	Qash Gammah	Female	Oman/GB	339/339	160	Bh1-GB	Bahlani 1	Male	Oman/GB	339/346
130	QSima	Qash Saima	Female	Oman/GB	339/339	161	Bh2-GB	Bahlani 2	Male	Oman/GB	339/346

131	Mdlok	Medlooki	Female	Oman/GB	339/339	162	Bh3-GB	Bahlani 3	Male	Oman/GB	339/346
132	Damos	Damoos	Female	Oman/GB	339/339	163	Bh4-GB	Bahlani 4	Male	Oman/GB	339/346
133	QHrez	Qash Hareez	Female	Oman/GB	339/339	164	Ghr1-GB	Ghareef 1	Male	Oman/GB	339/346
134	QALoz	Qash Al Looz	Female	Oman/GB	339/339	165	Ghr2-GB	Ghareef 2	Male	Oman/GB	339/346
135	QASmnh	Qash Al Semnah	Female	Oman/GB	339/339	166	Ghr3-GB	Ghareef 3	Male	Oman/GB	339/346
136	QHmryh	Qash Hamreiyah	Female	Oman/GB	339/339	167	Ghr4-GB	Ghareef 4	Male	Oman/GB	339/346
137	**QBalob**	**Qash Baloobiya**	**Female**	**Oman/GB**	**339/346**	168	BM1-GB	Al Fahel Al dhakm 1	Male	Oman/GB	339/346
138	QAASohn	Qash Abu Al Soh.	Female	Oman/GB	339/339	169	BM2-GB	Al Fahel Al dhakm 2	Male	Oman/GB	339/346
139	QMish	Qash Mishah	Female	Oman/GB	339/339	170	Unkn1-GB	Unknown Male 1	Male	Oman/GB	339/346
140	QADhiyh	Qash Al Dahiyah	Female	Oman/GB	339/339	171	Unkn2-GB	Unknown Male 2	Male	Oman/GB	339/346
141	QARmly	Qash Al Ramliyah	Female	Oman/GB	339/339	172	Unkn3-GB	Unknown Male 3	Male	Oman/GB	339/346
142	QAWli	Qash Al Wali	Female	Oman/GB	339/339	173	Bo's1-GB	Bu'Sab'ah 1	Male	Oman/GB	339/346
143	**Belq**	**Bel'aq**	**Female**	**Oman/GB**	**339/346**	174	Bo's2-GB	Bu'Sab'ah 2	Male	Oman/GB	339/346
144	Jebren	Jebreen	Female	Oman/GB	339/339	175	Bo's3-GB	Bu'Sab'ah 3	Male	Oman/GB	339/346
145	QBusmn	Qash Bussemen	Female	Oman/GB	339/339	176	R1-GB	Rghad 1	Male	Oman/GB	339/346
177	R2-GB	Rghad 2	Male	Oman/GB	339/346	208	5B	BC1 population	Female	Oman/WQ	340/340
178	R3-GB	Rghad 3	Male	Oman/GB	339/346	209	6B	BC1 population	Male	Oman/WQ	340/347
179	A'r1-GB	A'reesh 1	Male	Oman/GB	339/346	210	7B	BC1 population	Male	Oman/WQ	340/347
180	A'r2-GB	A'reesh 2	Male	Oman/GB	339/346	211	8B	BC1 population	Male	Oman/WQ	340/347
181	A'n1-GB	An'bati 1	Male	Oman/GB	339/346	212	10B	BC1 population	Male	Oman/WQ	340/347
182	A'n2-GB	An'bati 2	Male	Oman/GB	339/346	213	11B	BC1 population	Male	Oman/WQ	340/347
183	A'n3-GB	An'bati 3	Male	Oman/GB	339/346	214	12B	BC1 population	Male	Oman/WQ	340/347
184	Almq1-GB	Al Maquidha 1	Male	Oman/GB	339/346	215	13B	BC1 population	Male	Oman/WQ	340/347
185	Almq2-GB	Al Maquidha 2	Male	Oman/GB	339/346	216	14B	BC1 population	Female	Oman/WQ	339/339
186	Almq3-GB	Al Maquidha 3	Male	Oman/GB	339/346	217	15B	BC1 population	Male	Oman/WQ	340/347
187	Soq1-GB	Soo'qum 1	Male	Oman/GB	339/346	218	16B	BC1 population	Female	Oman/WQ	339/339
188	Soq2-GB	Soo'qum 2	Male	Oman/GB	339/346	219	17B	BC1 population	Female	Oman/WQ	339/339
189	Khzi1-GB	Khzini 1	Male	Oman/GB	339/346	220	18B	BC1 population	Male	Oman/WQ	340/347
190	Khzi2-GB	Khzini 2	Male	Oman/GB	339/346	221	19B	BC1 population	Female	Oman/WQ	339/339
191	Khzi3-GB	Khzini 3	Male	Oman/GB	339/346	222	20B	BC1 population	Male	Oman/WQ	340/347
192	Do1-GB	Do'wairah 1	Male	Oman/GB	339/346	223	23B	BC1 population	Female	Oman/WQ	339/339
193	Do2-GB	Do'wairah 2	Male	Oman/GB	339/346	224	24B	BC1 population	Female	Oman/WQ	339/339
194	Alls1-GB	Al Lasah 1	Male	Oman/GB	339/346	225	25B	BC1 population	Male	Oman/WQ	340/347
195	Alls2-GB	Al Lasah 2	Male	Oman/GB	339/346	226	26B	BC1 population	Female	Oman/WQ	339/339
196	Alls3-GB	Al Lasah 3	Male	Oman/GB	339/346	227	27B	BC1 population	Female	Oman/WQ	340/340
197	620	Parent1 BC	Female	Oman	340/340	228	28B	BC1 population	Male	Oman/WQ	340/347
198	81	Parent1 BC_R	Female	Oman	340/340	229	29B	BC1 population	Female	Oman/WQ	339/339
199	622	Parent4 BC	Female	Oman	340/340	230	30B	BC1 population	Male	Oman/WQ	340/347
200	85	Parent2	Male	Oman	340/347	231	32B	BC1 population	Male	Oman/WQ	340/347

201	Male13	Parent2_R	Male	Oman	340/347	232	33B	BC1 population	Female	Oman/WQ	339/339
202	Um	Accession	Female	Oman	340/340	233	34B	BC1 population	Male	Oman/WQ	340/347
203	Um.SQU	Parent3 F	Female	Oman	340/340	234	35B	BC1 population	Female	Oman/WQ	339/339
204	1B	BC1 population	Male	Oman/WQ	340/347	235	36B	BC1 population	Male	Oman/WQ	340/347
205	2B	BC1 population	Female	Oman/WQ	339/339	236	37B	BC1 population	Female	Oman/WQ	339/339
206	3B	BC1 population	Male	Oman/WQ	340/347	237	38B	BC1 population	Female	Oman/WQ	339/339
207	4B	BC1 population	Male	Oman/WQ	340/347	238	39B	BC1 population	Male	Oman/WQ	340/347
239	40B	BC1 population	Female	Oman/WQ	340/340	270	17F	F1 population	Female	Oman/B	340/340
240	42B	BC1 population	Female	Oman/WQ	339/339	271	18F	F1 population	Female	Oman/B	339/339
241	43B	BC1 population	Unknown	Oman/WQ	339/339	272	19F	F1 population	Female	Oman/B	339/339
242	44B	BC1 population	Female	Oman/WQ	339/339	273	20F	F1 population	Female	Oman/B	339/339
243	45B	BC1 population	Female	Oman/WQ	339/339	274	21F	F1 population	Female	Oman/B	339/339
244	46B	BC1 population	Female	Oman/WQ	339/339	275	22F	F1 population	Female	Oman/B	339/339
245	47B	BC1 population	Male	Oman/WQ	340/346	276	23F	F1 population	Male	Oman/B	3339/347
246	48B	BC1 population	Male	Oman/WQ	340/347	277	24F	F1 population	Male	Oman/B	340/347
247	49B	BC1 population	Male	Oman/WQ	340/347	278	25F	F1 population	Male	Oman/B	340/347
248	51B	BC1 population	Female	Oman/WQ	339/339	279	26F	F1 population	Male	Oman/B	340/347
249	52B	BC1 population	Female	Oman/WQ	339/339	280	27F	F1 population	Female	Oman/B	339/339
250	53B	BC1 population	Female	Oman/WQ	339/339	281	28F	F1 population	Female	Oman/B	339/339
251	54B	BC1 population	Male	Oman/WQ	340/347	282	29F	F1 population	Female	Oman/B	339/339
252	55B	BC1 population	Male	Oman/WQ	340/347	283	30F	F1 population	Male	Oman/B	340/347
253	57B	BC1 population	Male	Oman/WQ	340/347	284	31F	F1 population	Male	Oman/B	340/347
254	58B	BC1 population	Male	Oman/WQ	340/347	285	34F	F1 population	Male	Oman/B	340/347
255	59B	BC1 population	Female	Oman/WQ	339/339	286	35F	F1 population	Female	Oman/B	339/339
256	60B	BC1 population	Female	Oman/WQ	339/339	287	37F	BC1 population	Male	Oman/B	340/347
257	59B	BC1 population	Female	Oman/WQ	339/339	288	38F	F1 population	Female	Oman/B	339/339
258	60B	BC1 population	Female	Oman/WQ	339/339	289	41F	F1 population	Female	Oman/B	339/339
259	1F	F1 population	Male	Oman/B	340/347	290	42F	F1 population	Female	Oman/B	339/339
260	2F	F1 population	Male	Oman/B	340/347						
261	6F	F1 population	Female	Oman/B	339/339						
262	7F	F1 population	Male	Oman/B	340/347						
263	8F	F1 population	Male	Oman/B	340/347						
264	9F	F1 population	Male	Oman/B	340/347						
265	10F	F1 population	Female	Oman/B	339/339						
266	13F	F1 population	Male	Oman/B	340/347						
267	14F	F1 population	Female	Oman/B	339/339						
268	15F	F1 population	Female	Oman/B	339/339						
269	16F	F1 population	Female	Oman/B	339/339						

Appendix 8: Lists of 96 accessions from Sanremo, Bordighera, USDA-ARS, France and other origins including Iraq, Libya, Sudan and Iran used in this study. Laboratory code, gender, country of origin and observed allelic size range are mentioned.

Sample No.	Lab Code	Accession Name	Gender	Country of Origin	Observed Size (bp)	Sample No.	Lab Code	Accession Name	Gender	Country of Origin	Observed Size (bp)
1	406	-	Female	Sanremo	340/340	25	447	-	Female	Sanremo	339/339
2	407	-	Female	Sanremo	340/340	26	449	-	Female	Sanremo	339/339
3	408	-	Female	Sanremo	340/340	27	450	-	Female	Sanremo	339/339
4	409	-	Female	Sanremo	339/339	**28**	**462**	-	**Female**	**Bordighera**	**340/347**
5	412	-	Female	Sanremo	339/339	29	474	-	Male	Bordighera	340/347
6	414	-	Female	Sanremo	339/339	**30**	**478**	-	**Male**	**Sanremo**	**340/340**
7	**415**	-	**Male**	**Sanremo**	**340/340**	31	485	-	Male	Bordighera	339/346
8	**419**	-	**Male**	**Sanremo**	**340/340**	32	494	-	Female	Bordighera	340/340
9	420	-	Female	Sanremo	339/339	33	495	-	Male	Bordighera	340/347
10	421	-	Male	Sanremo	340/347	34	500	-	Female	Bordighera	340/340
11	422	-	Unkown	Sanremo	339/339	35	501	-	Female	Bordighera	340/340
12	423	-	Male	Sanremo	340/347	36	502	-	Female	Bordighera	340/340
13	426	-	Male	Sanremo	340/347	**37**	**503**	-	**Female**	**Bordighera**	**340/346**
14	428	-	Male	Sanremo	340/347	38	514	-	Female	Bordighera	340/340
15	429	-	Female	Sanremo	340/340	**39**	**517**	-	**Female**	**Bordighera**	**340/347**
16	431	-	Female	Sanremo	340/340	40	521	-	Female	Bordighera	340/340
17	433	-	Female	Sanremo	340/340	41	522	-	Female	Bordighera	339/339
18	434	-	Female	Sanremo	340/340	42	523	-	Female	Bordighera	339/339
19	439	-	Male	Sanremo	339/346	43	527	-	Female	Bordighera	339/339
20	441	-	Female	Sanremo	339/339	**44**	**529**	-	**Female**	**Bordighera**	**340/347**
21	442	-	Male	Sanremo	340/347	45	541	-	Male	Bordighera	340/347
22	443	-	Male	Sanremo	340/347	46	564	-	Unkown	Bordighera	340/346
23	444	-	Male	Sanremo	340/347	47	574	-	Unkown	Bordighera	340/347
24	446	-	Female	Sanremo	340/340	**48**	**93001**	-	**Male**	**Sanremo**	**340/340**
49	Gondaila'	Gondaila'	Female	Sudan	340/340	80	Sa-Ly	Saidi	Female	Libya	338/338
50	Barakawi'	Barakawi'	Female	Sudan	339/339	81	Aq-Ly	Aqudool	Female	Libya	339/339
51	Khalass	Khalass	Female	USDA-ARS	339/339	82	Med-Sdn	Medina	Female	Sudan	339/339
52	Fardh #4	Fardh #4	Male	USDA-ARS	340/347	83	Gnd-Sdn	Gondaila	Female	Sudan	339/339
53	Thory	Thory	Female	USDA-ARS	339/339	84	Bar-Sdn	Barakawi	Female	Sudan	339/339
54	Hilali	Hilali	Male	USDA-ARS	340/340	85	Bit-Sdn	Bitamoda	Female	Sudan	339/339
55	Barhee	Barhee	Female	USDA-ARS	340/340	86	Do-Sdn	Dogna	Female	Sudan	339/339
56	Medjool	Medjool	Female	USDA-ARS	339/339	87	Iran1	Barhi	Female	Iran	339/339

57	Fran1	-	Female	France	339/339	**88**	**Iran3**	**Bentossbae**	**Female**	**Iran**	**339/346**
58	**Fran2**	-	**Female**	**France**	**339/346**	89	Iran4	Ddagal-e Zard	Female	Iran	339/339
59	**Fran3**	-	**Female**	**France**	**339/346**	90	Iran6	Shekar	Female	Iran	339/339
60	Fran4	-	Female	France	339/339	91	Iran9	Gentaar	Female	Iran	339/339
61	Fran5	-	Female	France	339/339	92	Iran13	Zahedi	Female	Iran	339/339
62	Fran6	-	Female	France	339/339	93	Iran22	Soweidance	Female	Iran	339/339
63	Fran7	-	Female	France	339/339	94	Iran33	Nashenaas	Female	Iran	339/339
64	DA-Iq	Daml Asfer	Female	Iraq	339/339	95	Iran34	Ghanaami	Female	Iran	339/339
65	B-Iq	Badmi	Female	Iraq	339/339	96	Iran40	Halilehei	Female	Iran	339/339
66	Sar-Iq	Sarmadti	Female	Iraq	339/339						
67	Khd-Iq	Khadrawy	Female	Iraq	339/339						
68	Mkm-Iq	Maktoom	Female	Iraq	339/339						
69	Bdm-Iq	Bdmalki	Female	Iraq	339/339						
70	Ben-Iq	Benosh	Female	Iraq	339/339						
71	Ash-Iq	Ashrasi	Female	Iraq	339/339						
72	Khs-Iq	Khastawi	Female	Iraq	339/339						
73	Say-Iq	Saylani	Female	Iraq	339/339						
74	Bhm-Iq	Bahram	Female	Iraq	339/339						
75	Aw-Ly	Awreeq	Female	Libya	339/339						
76	Kh-Ly	Khmag	Female	Libya	339/339						
77	Ta-Ly	Taghiyat	Female	Libya	339/339						
78	Am-Ly	Amreer	Female	Libya	339/339						
79	Tal-Ly	Talees	Female	Libya	338/338						