डी एन ए चिह्नकों का उपयोग कर भारतीय गेहूँ (*ट्रिटीकम* प्रजातियों) भूप्रजातियों में आबादी संरचना एवं जातिवृत्त का अध्ययन

Study of population structure and phylogeny in Indian wheat (*Triticum* spp.) landraces using DNA markers

KUMARASWAMY, H. H.

DIVISION OF MOLECULAR BIOLOGY AND BIOTECHNOLOGY
NATIONAL RESEARCH CENTRE ON PLANT BIOTECHNOLOGY
INDIAN AGRICULTURAL RESEARCH INSTITUTE
NEW DELHI – 110 012

2019

Study of population structure and phylogeny in Indian wheat (*Triticum* spp.) landraces using DNA markers

By

KUMARASWAMY, H. H.

A thesis
Submitted to the Post Graduate School,
ICAR-Indian Agricultural Research Institute, New Delhi,
In partial fulfillment of requirements
for the award of the degree of

DOCTOR OF PHILOSOPHY

In

MOLECULAR BIOLOGY AND BIOTECHNOLOGY

2019

Approved by Advisory committee:

Chairperson: **Dr. N. K. Singh**

Co-Chairperson: **Dr. Kishor Gaikwad**

Members: **Dr. S. Gopala Krishnan**

Dr. C. Viswanathan

ICAR-NATIONAL RESEARCH CENTRE ON PLANT BIOTECHNOLOGY
DIVISION OF MOLECULAR BIOLOGY AND BIOTECHNOLOGY
ICAR-INDIAN AGRICULTURAL RESEARCH INSTITUTE,
NEW DELHI-110 012

Prof. (Dr.) N. K. Singh
Principal Scientist &
National Professor – Dr. B. P. Pal Chair

CERTIFICATE

This is to certify that the thesis entitled "**Study of population structure and phylogeny in Indian wheat (*Triticum* spp.) landraces using DNA markers**" submitted to the Faculty of the Post-Graduate School, ICAR-Indian Agricultural Research Institute, New Delhi, in partial fulfillment of the requirements for the award of degree of **DOCTOR OF PHILOSOPHY in MOLECULAR BIOLOGY AND BIOTECHNOLOGY**, embodies the results of *bona fide* research work carried out by **Mr. Kumaraswamy, H. H.,** under my guidance and supervision, and that no part of this thesis has been submitted for award of any other degree or diploma.

It is further certified that any assistance and help availed during the course of investigation as well as sources of information, have been duly acknowledged by him.

Date: 28/1/19

Place: New Delhi-12

(N. K. Singh)
Chairperson, Advisory Committee

DEDICATED TO

LORD BALAJI

AND

MY PARENTS & TEACHERS

ACKNOWLEDGEMENTS

Any research for that matter is not a one man's show, it is a concerted performance and consortium of supports. Hence, it is my sincere duty to acknowledge all the persons who directly or indirectly encouraged, supported or helped while doing the research and/or preparation of the manuscript. Without my supporters, this thesis would not have taken its present shape. At the outset, I thank one and all.

I consider myself privileged and fortunate for having worked under the able guidance and super vision of ***Prof* (*Dr.*) *N.K. Singh****, Principal Scientist (Biotechnology) and National Professor-Dr. B.P. Pal Chair, ICAR-National Research Centre on Plant Biotechnology* (*ICAR-NRCPB*), *New Delhi*, *and the Chairperson of my advisory committee, Division of Molecular Biology and Biotechnology, Post Graduate School, ICAR-Indian Agricultural Research Institute (ICAR-IARI), PUSA Campus, New Delhi. I am sincerely obliged and indebted to him for his invaluable guidance, constant encouragement, scholarly suggestions, untiring enthusiasm, affectionate behavior and he had been a source of constant inspiration and support.*

I wish to place on record of my sincere and heartfelt gratitude to ***Dr. Kishor Gaikwad****, Principal Scientist (Biotechnology), ICAR-National Research Centre on Plant Biotechnology (ICAR-NRCPB), New Delhi, and the Co-chairperson of my advisory committee, for his constant encouragement, scientific suggestions, valuable guidance and all the help throughout the programme of this work.*

It is a great privilege for me to express my esteemed and profound gratitude to ***Dr. Viswanathan Chinnusamy****, Head & Principal Scientist, Division of Physiology, ICAR- IARI, New Delhi, and Member of my advisory committee, for his scholarly guidance, experienced words, constant encouragement, and help throughout this study.*

With endless pleasure, it is a great occasion for me to express my deep sense of gratitude to ***Dr. S. Gopala Krishnan****, Principal scientist, Division of Genetics, ICAR-IARI, New Delhi, and Member of my advisory committee, for his cooperative attitude, encouragement and help.*

It gives me a tremendous pleasure to express my gratitude and reverence to ***Dr. T. Mohapatra****, Secretary, DARE, Ministry of Agriculture and Farmers' Welfare, Government of India, and Director General, ICAR, New elhi; and,* ***Dr. T.R. Sharma****,*

Director, NABI, Chandigarh; *who were, in the beginning, Co-chairpersons of my advisory committee, and earstwhile Principal Scientists, ICAR-NRCPB*; *for their critical comments, personal interest, encouraging words and moral support throughout the study*.

It gives me a great pleasure to express my heartfelt thanks to ***Dr. G.P. Singh****, Director, ICAR-IIWBR, Karnal, and former Member of my advisory committee, and then Principal Scientist* (*Genetics*), *Division of Genetis, Post Graduate School, ICAR-Indian Agricultural Research Institute (ICAR-IARI), PUSA Campus, New Delhi*; *for his generous help while planning and performing the field experiments, encouraging words and moral support throughout the study*.

I acknowledge, with thanks, the facility and financial support from ***ICAR-NPTC project***, ***Post Graduate School***, ***ICAR-Indian Agricultural Research Institute***, *New Delhi, and* ***ICAR-National Research Centre on Plant Biotechnology***, *New Delhi*; *the kind supply of plant materials used in this research by late* ***Dr. S.K. Mishra***, *ICAR-National Bureau of Plant Genetic Resources;* ***land facility*** *for field experiments provided by* ***Division of Genetics***, *ICAR-Indian Agricultural Research Institute, New Delhi; and,* ***study leave*** *granted by* ***Director***, ***ICAR-IIOR***, *Hyderabad, and* ***Indian Council of Agricultural Research***, *New Delhi*.

I thank ***Dr. Amita Charu***, *Scientist (Biotechnology), ICAR-NRCPB, PUSA Campus, New Delhi, for facilitating my capillary gel electrophoresis requirements during this study.*

I take this opportunity to thank all my lab members ***Dr. Vikram Singh***, ***Mr. Vijay Singh***, ***Dr. Balwant Singh***, ***Dr. Sonia Goel***, ***Dr. Renu Singh***, ***Ms. Neera Yadav***, ***Dr. Kavita Singh***, ***Dr. Shefali Mishra***, ***Dr. Vagesh Mishra***, and **all labmates**, *for their encouragements and help in carrying out experiments. Also, I place on record the help of* ***Mr. Munesh***, ***Mr. Abhimanyu***, ***Mr. Omkar (Dablu)***, ***Mr. Amit*** *and* ***Mr. Rohit*** *in carrying out lab and field experiments.*

My acknowledgements are due to ***Mr. Devoji Srikanth***, ***Mr. Sravan***, ***Mr. Kiran***, *and* ***Mrs. Raji*** *of ICAR-IIOR, Hyderabad, for their help during the preparation of manuscript of this thesis.*

I sincerely thank ***Dr. Jyoti Kumari****, Senior Scientist* (*Plant Breeding*) *and* ***Dr. Sunil Archak****, National Fellow, ICAR-NBPGR, New Delhi, for their kind help in getting information about wheat accessions used in the present study.*

It is my heart-felt offer of sincere thanks all ***my teachers*** *at P.G. School, ICAR-IARI, New Delhi; and the* ***entire consortium of teachers****, in my academic, personal, spiritual, and professional life, for initiation and expansion of my overall knowledge, without which, at the outset, this study would not have been possible.*

I place on record the due acknowledgements that ***Dr. Dhandapani****, Principal Scientist* (*Ag. Statistics*)*, ICAR-National Academy of Agricultural Research Management, Hyderabad;* ***Dr. C. Sarada****, Principal Scientist* (*Ag. Statistics*) *and* ***Dr. S. Senthil****, Principal Scientist* (*Plant breeding*)*, ICAR-Indian Institute of Oilseeds Research, Hyderabad;* ***Mr. Neeraj Bundhalkoti****, Scientist* (*Bioinformatics*)*, ICAR-Indian Agricultural Statistical Research Institute, New Delhi; and,* ***Dr. Nagaraja Reddy****, Scientist* (*Biotechnology*)*, ICAR-Directorate of Medicinal & Aromatic Plants Research, Anand, Gujarat; kindly helped in the processing of data using various statistical and bioinformatic tools****.***

I thank all ***my classmates*** *and* ***hostel-mates*** *for their great company, and making my days of Ph.D. study at IARI memorable.*

I thank my beloved wife ***Mrs. M.S. Dhanalakshmi****, daughter* ***Ms. H.K. Punyakoti****, and son* ***Mr. H.K. Vishnu Bhagawan****, brother* ***Mr. H.H. Umesha****, nephews* ***Mr. Panduranga*** *and* ***Mr. Madhu****, niece* ***Mrs. Pavithra Yoganand****, for their abundant love, inspiration, selfless sacrifice and moral support.*

I am overwhelmed with feelings of indebtedness to my beloved father ***Late Shri. Hanumegowda*** *and mother* ***Late Smt. Jayalakshmamma****, for connecting my soul to this beautiful world without, at the outset, which this study would not have been possible.*

Devotionally profused thanks to ***Lord Tirupati Balaji****, for everything.*

Dated: 28-01-2019

Place: New Delhi-110012

(*KUMARASWAMY, H. H.*)

TABLE OF CONTENTS

Contd…

TABLE OF CONTENTS (Contd…)

Contd…

TABLE OF CONTENTS (Contd…)

Contd…

4. RESULTS 87-109

Contd…

TABLE OF CONTENTS (Contd…)

Contd…

TABLE OF CONTENTS (Contd…)

LIST OF TABLES

Contd…

LIST OF TABLES (contd…)

Contd…

LIST OF TABLES (contd…)

Contd…

LIST OF TABLES (contd…)

LIST OF FIGURES

Contd…

LIST OF FIGURES (contd…)

ABBREVIATIONS AND SYMBOLS

AMOVA	:	Analysis of molecular variance
bp	:	base pair(s)
CTAB	:	Cetyl Trimethyl Ammonium Bromide
Conc	:	concentration
DNA	:	Deoxy ribose nucleic acid
EDTA	:	Ethylene diamine tetra-acetic acid
h	:	hour
Kb	:	kilo base pair
µM	:	micro molar
µL	:	micro liter
min	:	minutes
mg	:	milligram
mM	:	milli molar
M	:	Molar
OD	:	Optical density
PCR	:	Polymerase Chain Reaction
RT	:	Room Temperature
rpm	:	revolutions per minute
NGS	:	Next generation sequencing
TBE	:	Tris borate EDTA
W/v	:	weight/volume
V	:	volt
V/v	:	volume/volume
TE	:	Tris EDTA buffer
Tris	:	Tris (hydroxy methyl) amino methane

Note: The abbreviations for chemicals and symbols follow the tentative rules of IUPAC-IUB Commission on Biochemical Nomenclature.

CHAPTER 1

INTRODUCTION

1. INTRODUCTION

Bread wheat (*Triticum aestivum* L.) is a self-pollinating annual plant, belonging to the sub-tribe Triticineae of the Tribe Triticeae in the the family Poeceae (grasses). It is a segmental allo-hexaploid having three closely related genomes in two copies each (AABBDD) with chromosome number of 42 (2n=6x=42) and a genome size of 16 GB. The progenitor of subgenome A of hexaploid wheat is confirmed to be *T. urartu* (2n = 2x = 14, genome AA) and that of B is *Aegilops speltoides* (2n = 2x = 14, genome SS). Regarding the progenitor of constituent genome B of hexaploid wheat, it is still not confirmed as the donor might have lost its identity in the course of evolution. However, *Aegilops tauschii* (2n = 2x = 14, genome DD) is speculated (Gupta *et al.*, 2005) to have been the donor of subgenome D.

Wheat is the oldest and one of the most important cereal crops. Wheat cultivation has been practiced since ancient times, 10000-8000 BC (Heun *et al*., 1997; Nesbitt and Samuel, 1998; Luo *et al.*, 2007). It is the staple food for majority of the world's population and, therefore, it is the most widely grown crop. Since several centuries, wheat has been playing an important role in human nutrition, as it is a major source of dietary fiber, protein, and energy. Wheat alone provides nearly 55% percent carbohydrates and 20 percent of the food calories, and proteins in human nutrition, consumed globally, and staple food of about two billion people (*i.e*. approximately 40% of the world population) (Breiman and Graur, 1995). Thus, wheat, being the first cultivated cereal in the world (>200 mha annually), plays a crucial role in sustaining food security (Shiferaw *et al.*, 2013). Currently, while nearly 95% of the globally cultivated wheat is hexaploid (bread) wheat, durum wheat constitutes the remaining 5% (Shewry, 2009). In addition to guaranteeing access to food, wheat is a key commodity to eradicate global hunger by ensuring sufficient production to feed the ever-growing world population that is projected to reach 9 billion by 2050 (FAO, 2010).

Wheat cultivation in India has been practiced in the last 5000 years (Feldman, 2001). Wheat is used in various forms such as baking, making bread, pastas and biscuits. Wheat grain is also a source of alcoholic beverages and wheat starch is used for making pastes. Wheat bran from flour milling and vegetative plant parts are used as livestock feed.

Wheat has the potential to be used as non-food raw material in many industries such as shiffening/ surface coating agent in the manufacturing of paper and board; as source of adhesive in the manufacturing of corrugated material; as fermentation substrate in the production of nutraceutical supplements such as vitamins, antibodies, etc (Jones *et al.*, 1987; Sharma, 2016). Apart from its usage as fuel, traditional construction material such as thatches and panels, used for building houses and in making various handicraft articles in rural areas, are made from wheat straw. Globally wheat trade is a major and impacting issue in political and economic relationships among nations (Sharma, 2016). Wheat is consumed as a staple food in different forms in India: mainly as chapati, bread, noodles, macaroni, sphaghetti, cakes, pizzas, doughnuts, etc. It is also consumed as semolina, locally known as rava or sooji, used for the preparation of kheer, upma, sooji, halwa, etc. Vermicelli made from durum wheat is used for making kheer.

Three types of wheat are grown in India: 1) *Triticum aestivum*, popularly called bread wheat, 2) *Triticum durum*, known as durum wheat, and 3) *Triticum dicoccum*, also called dicoccum wheat). Being a *Rabi* crop in India, wheat is planted during winter as the crop needs cool winter and hot summer.

The global production of wheat has been recorded 729 million tonnes harvested from 220 million hectares with the average productivity of 3.3 t/ ha (FAOSTAT, 2014). India accounts for 13% of the total global wheat production (FAOSTAT, 2014). In the year 2014, India was the world second-largest producer, next to China, with the record production of 95.58 million tonnes harvested from 30.47 million hectares with an average productivity of 3.15 t/ha (FAOSTAT, 2014). However, the wheat average productivity in India is lower than the world average. Enhancing the production in the face of changing climate *inter alia* requires protection against biotic and abiotic stresses (Singh *et al.*, 2008; Singh *et al.*, 2011). Germplasm resources are gold mines of genes for resistance/ tolerance to biotic and abiotic stresses.

The diverse genetic resources with wider adaptation and broad utility have rendered cultivated wheat (*Triticum aestivum* L.) a status of a worldwide food crop. Domestication and modern plant breeding have presumably narrowed the genetic base of bread wheat (*Triticum aestivum*), which could jeopardize future crop improvement (Reif *et al.*, 2005).

Further, it has been argued that modern plant breeding has led to genetic homogeneity and possible vulnerability to biotic and abiotic stresses (Vellve, 1993); Russel, 2000; Roussel *et al*., 2004; Fu, 2005). While there is a multitude of high yielding modern wheat varieties in use, a variety that is successful today can overnight be rendered ineffective by changes in the agro-ecosystem and/or outbreaks of new races of pests and diseases. Landrace cultivars (LC) undoubtedly represent an important source of genetic variation in wheat. One of the prime examples is the use of *Rht* dwarfing genes that became available through the Japanese wheat 'Norin10', derived from the LC Shiro Daruma (Kihara, 1983). Approximately, 266589 hexaploid wheat collections are available worldwide of which 24% constitute landraces (FAO, 2002).

Further, due to an increasing pressure on arable land caused by increasing paces of human population and urbanization; and, since there are environmental bottle-necks due to challenges of climate change; crop production needs to be increased to secure future food supplies, while minimizing its impact on ecosystems. Therefore, in the face of these challenges, genomes and genetic diversity need to be understood in great details (Bevan *et al.*, 2017). Wheat production must, therefore, adapt to more variable environments and the substantial impact it has on the environment needs to be reduced. Productivity must also increase at a much greater rate than in the past to meet the needs of Earth's growing population (Godfray *et al.*, 2010). While there are increasing number of biotic- and abiotic stresses coupled with climate changes, accelerated by global warming, there will be a need to increase cereal production by 1.7-fold to meet future challenges of food security (Mba *et al*., 2012). Wheat crop has relative advantage due to its easy adaptability to variable climate and low-cost production. The International Food Policy Research Institute (IFRI) has taken up wheat improvement programme as more effective measure to meet the food security of the world (Andersen and Cohen, 2001). The National Commission on Agriculture also has estimated that India needs approximately 110 million tons of wheat by 2020 (Joshi, 2007). All these scenarios have necessitated the search for new genes from germplasm accessions which were adapted to local condition for centuries.

Detailed information on intra- and inter-germplasm genetic variability is crucial as they offer practical opportunities not only for future collection trips but also for the

conservation and utilization of existing genetic resources (Karp *et al*., 1997; Warburton and Hoisington, 2001; Matus and Hayes, 2002).

Genetic diversity analysis is an initial, foremost, and important step for exploitation of any genetic resources. The systematic evaluation of the molecular diversity encompassed in wheat genetic resources is a prerequisite for its efficient utilization in breeding as well as improvement programmes. Some applications of diversity analysis using molecular marker tools includes: identification of genetic diversity (Hamrick and Godt, 1990), determining collection priorities and sampling strategies (Schoen and Brown, 1991), guiding the designation of *in situ* or on-farm conservation strategies (Bonierbale *et al.,* 1997), monitoring genetic erosion (Robert *et a1.*,1991) or vulnerability (Adams, 1977), guiding the management of *ex situ* collections (Kresovich *et al.,* 1997), maximizing the genetic diversity in core collections (Gepts, 1995), comparing agronomically useful regions of genomes of different crops (Paterson *et al.,* 1995), defining the identity of improved varieties or other plant genetic resources (Lee *et al.,* 1995), monitoring the movement of plant genetic resources (Hardon *et al.,* 1994) and assisting in taxonomic evolution and enhancing the understanding of relationships between crop gene pools (Gepts, 1995), achieving precise, unambiguous and accurate identification of germplasm at species/ subspecies levels (Wang and Tanksley, 1989; Virk *et al.,* 1995; Zhu *et al.,* 1998), and identifying duplicates within the collections particularly in gene banks (Virk *et al.,* 1995).

The application of different techniques to genetic diversity analysis have been well reviewed by many authors (Malyshev and Kartel, 1997; Newbury and Ford-Lloyd, 1997; Westman and Kresovich, 1997; Karp *et al.,* 1997; Idrees and Irshad, 2014). Assessment of level and distribution of polymorphism (usually conceptualized as ' allelic richness' and ' allelic evenness') in a crop permit the sampling and utilization of genetic resources in a more systematic and efficient manner; and, allow an enhanced understanding of evolutionary relationships useful for both breeding and conservation efforts (Aruna, 2003).

Genetic diversity can be estimated based on various criteria: pedigree analysis (Barrett *et al.,* 1998), morphological traits (Schut and Stam, 1997; Maric *et al.,* 1998; Casadesus *et al.,* 2007; Zarkti *et al.,* 2012; Malik *et al.,* 2014) as well as biochemical

markers (Cox *et al.,* 1985; Metakovsky and Branlard, 1998) or molecular markers (Rao and Riley, 1994; Karp *et al.,* 1996; Gupta *et al.,* 1999; Manifesto *et al.,* 2001; Pagnotta *et al.,* 2005). As morphological traits are the best alternative to their molecular counterparts for assessing genetic diversity (Hailu *et al.,* 2010; Mengistu *et al.,* 2015), agromorphological characterization is primary step for conservation and utilization of plant genetic resources (Hafida *et al.,* 2012). Thus, diversity can be assessed using both morphological traits (genetic diversity) and molecular/DNA markers (genomic diversity).

Data on agro-morphological and phenological variability in crop species, including their wild and weedy relatives, are useful for basic studies on crop evolution (Jain and Singh, 1972; Harlan, 1981; Simmonds, 1976; Jain, 1977), planning efficient germplasm collection expeditions (Moseman and Craddock, 1976; Cristopher *et al.,* 1984), and selecting parents to use them in plant breeding programmes (Ward, 1962; Munck *et al.,* 1970; Bartual *et al.,* 1985; Dale *et al.,* 1985). For several crop species, evaluation of entries in germplasm collections has provided a large data on the plant traits that can be utilized to study patterns of genetic diversity (Harlan, 1981; Frankel and Hawkes, 1975; Polignano and Spagnoletti Zeuli, 1985).

Thus, classical methods of estimating genetic diversity and / or relatedness among groups of plants relied upon phenotypic traits. However, these have two disadvantages: firstly, the traits are subject to environmental influences, and, secondly, the levels of polymorphism (allelic variation) are limited. These limitations were significantly overcome by development of environment- neutral biochemical makers (Isozymes) and protein electrophoresis (Hunter and Markert, 1957), and molecular markers that focus directly on variations controlled by genes. The higher resolution of molecular markers makes them a valuable tool for a variety of purposes, *viz*., fingerprinting; appropriate choice of parents for breeding programs; analysis of quantitative traits; location and detection of quantitative trait loci (QTLs); gene mapping, marker assisted selection and gene transfer; understanding evolutionary pathways; and, the assessment of genetic diversity of plant germplasm. Hillis (1987) recommended that morphological work on large samples combined with molecular analysis on smaller samples maximize both information and usefulness. Kresovich and his colleagues (1997) believed that molecular markers could resolve biological, operational and logistical questions dealing with four

broad areas of germplasm characterization: the determination of correct identity of an individual to ascertain, for onstance, whether it is true-to-type, or is there any duplication, etc.; the estimation of degree of similarity among individuals; understanding of hierarchical structure and partitioning of variations among individuals, accessions, populations, and species; and, identification and detection of the presence of particular alleles in individuals, accessions, populations, chromosomes or cloned DNA segments. The range of molecular markers that can be, relatively, easily used in most plant germplasm is quite extensive (Gupta and Varshney, 2000). Techniques for identifying polymorphism in actual DNA sequence use DNA-hybridization methods such as RFLP (Restriction Fragments Length Polymorphism); or PCR-based masrker systems such as RAPD (Random Amplified Polymorphic DNA), SSR (Simple Sequence Repeat); or the method based on both hybridization and PCR such as AFLP (Amplified Fragment Length Polymorphism). There are various methods that differ with respect to their cost, ease of application, type of data generated, dominant or co-dominant markers, the degree of polymorphism that they are capable of revealing, the way they resolve genetic differences, and the taxonomic levels at which they can be most appropriately used (Karp and Edwards, 1997). DNA markers can be used as powerful tools to study phylogenetic architecture of the plant populations (Zhang *et al*., 2006). In the past decade, genetic diversity study of wheat has been benefited from the application of SSR markers. Dreisigacker *et al*. (2005) has analysed genetic variation in CIMMYT wheat landraces where they studied not only the diversity among landraces but also within each of them. Roussel *et al*. (2004) reported on chronological trends and crop improvement programmes analyzed with the help of genetic variations assessed among French wheat accessions. Zhang *et al.* (2006) utilized SSR markers to assess genetic and phylogenetic relatedness among Oman landraces. Population structure and phylogeny of Indian landraces of wheat have not been studied, whereas such studies are carried out in landraces of other countries. Zhang *et al.* (2010) assessed Genetic diversity and population structure in U.S. elite winter wheat.

A large number of methodologies exist for the assessment of genetic diversity in plant species. A combination of morphological traits and protein profiling methods; *viz*., isozymes (Nevo *et al.*, 1986), allozymes (Hamrick and Godt, 1990), and seed storage proteins (Doll and Brown, 1979); have conventionally been applied. However, such traits

are influenced by environmental factors, and, therefore, the results elucidated based on such studies do not provide a true measure of genetic diversity. The advent of environmentally neutral molecular markers will allow better quantification of genetic diversity (Clegg, 1984; Gepts, 1995; Aruna, 2003). Recent technological advancements have expanded the range of DNA polymorphism assays for fingerprinting, and investigating genetic diversity and genetic relatedness. Assessment of levels and distribution of polymorphism, usually conceptualized as 'allelic richness' and 'allelic evenness', respectively, permits the sampling and utilization of genetic resources in a more systematic, effective and efficient manner, and also allow an enhanced understanding of evolutionary relationships that are essential for both breeding and conservation efforts (Aruna, 2003). These technologies include restriction fragment length polymorphism (RFLP) (Botstein and White, 1980; Rafalski and Vogel, 1996), random amplified polymorphic DNA markers (RAPD) (Bowcock, 1994), amplified fragment length polymorphism (AFLP) (Zabeau and Vos, 1993), and simple sequence repeats or micro-satellites (SSR) (Tautz, 1989; Weber and May, 1989). These methods detect polymorphism by assaying subsets of the total amount of DNA sequence variation in a genome (Aruna, 2003). Polymorphisms detected by using AFLP and RFLP marker systems is based on variations in restriction fragment sizes. RFLPs have been used to characterize the genetic diversity in cultivated pigeonpea and its wild relatives (Beckmann and Soller, 1983; Miller and Tanskley, 1990; Wang *et al.*, 1992; Sivaramakrishnan *et al.*, 2001). There are many advantages of RFLPs in the estimation of genetic relationships (Melchinger *et al.*, 1991; Smith *et al.*, 1991; Stuber, 1992). AFLP methodology has been used to assess genetic diversity in rice (Zhu *et al.*, 1998), lactuca (Hill *et al.*, 1996), Soybean (Sharma *et al*, 1996; Maughan *et al.*, 1997), and sunflower (Hongtrakul, 1997). The genetic variability of 38 grape fruits (*Citrus paradbr*) and three pummelos (*C. maxima*) accessions was evaluated using RAPD and SSR maker analyses; and it was found that approximately 49% of the 198 RAPD were polymorphic, and 4.6 alleles per SSR loci were identified (Corazza-Nunes, 2002). The AFLP system (Vos *et al.*, 1995) has been shown to be effective and reproducible in analysis of genetic linkage and gene mapping (MacKill *et al.*, 1996; Voorrips *et al.*, 1997), map-based cloning (Cnops *et al.*, 1996), and plant evolution (Huem *et al.*, 1997). Biodiversity technology has been applied to wheat in localized situations (Parker *et al.*,

1998; Barrett *et al*., 1998; Law *et al.*, 1998; Ma and Lapitan, 1998; Hartl *et al.*, 1999; Bai *et al.*, 1999; Bohn *et al.*, 1999; Singh *et al.*, 1999). RFLP analysis, using genomic single copy probes, has been generally used to characterize the variation among wild and cultivated species (Miller and Tanskley, 1990; Jena and Kochert, 1991; Gawel *et al.*, 1992; Janet *et al.*, 1992). Variations in chloroplast DNA (Close *et al.*, 1989) and mitochondrial DNA (Deu *et al.*, 1995; Moeykens *et al.*, 1995; Tozuka *et al.*, 1998) have been used to study the diversity of cytoplasms in soyabean and pigeonpea. Compared to chloroplast genome, the mitochondrial genome has many variations within, and/or between, closely related species (Close *et al.*, 1989; Grabau *et al.*, 1992; Deu *et al.*, 1995.).

The ways by which genetic diversity in populations is estimated and represented are important. The effectiveness of this strategy depends on the suitability of the marker system for analysing the diversity of a particular sample collection under investigation. Different marker methods can give different views of diversity, depending upon the evolutionary parameters of the underlying DNA sequence variation. For instance, simple sequence repeats (SSRs), give high resolution views on relatedness; and, single nucleotide polymorphism (SNP)-based variation is more suited to reveal deeper relationships, reflecting the slow mutation rate for this type of sequence variation (Jing *et al.*, 2007).

SSRs are highly polymorphic and are becoming the markers of choice in both animal and plant species (Condit and Hubell, 1991; Akkaya *et al.*, 1992; Morgante and Oliveri, 1993; Wang *et al.*, 1994). Song *et al.* (2005) and Torada *et al.* (2006) have developed SSR-based linkage maps in bread wheat and many other researchers have carried out genomic analysis and marker development for specific loci using bin-mapped ESTs and flanking bacterial artificial chromosome (BAC) contigs (Lu *et al.*, 2006). Currently, bin-mapped ESTs and gene distribution covering all the seven chromosomes of each of the three genomes of bread wheat is available (Qi *et al.*, 2003; Qi *et al.*, 2004;). Further, Singh *et al.* (2007) have reported the location of 213 single copy (SC) rice gene homologs of wheat in the telomeric bins of 21 wheat chromosomes, based on the conserved collinearity.

Simple sequence repeat (SSR) markers remain one of the best marker systems for wheat research because of their chromosome specificity, high polymorphism (Plaschke *et*

al., 1995; Huang *et al*., 2002), high reproducibility, codominant inheritance patterns (Roder *et al*., 1998), and their capability to generate large datasets in a short period of time and, thereby, facilitating the evaluation of large number of germplasm accessions in seed banks (Tang and Knapp, 2003).

While choosing genome wide markers, the following findings have to be kept in mind. Among the features of the wheat genome found in previous mapping studies are the suppression of genetic recombination near the centromeres, and the apparently higher density of genes toward the distal regions of the chromosomes relative to the centromeric regions (Gill *et al.*, 1996b), consistent with findings in other cereals such as rice (Wu *et al.*, 2002). Due to presence of three homoeologous genomes, gene duplication is higher than expected, with many cases of more than three copies (paralogs) per gene (Anderson *et al.*, 1992; Dubcovsky *et al.*, 1996). Gene loss and other types of rearrangements have also been documented where one or two of the three orthologous copies of a gene are missing or have moved out of the expected homoeologous regions in one of the three genomes (Li and Gill, 2002). Structural changes in some of the chromosomes relative to their homoeologs have been described and the physical mapping of wheat ESTs has supported earlier reports. Among them are complex inversions at chromosome 4A (relative to 4B and 4D), a pericentric inversion in 4B, and two reciprocal translocations of portions of 4AL with 5AL and then with 7BS (Naranjo *et al.*, 1987; Chen and Gustafson, 1997).

In the recent past, genetic diversity among wheat cultivars, including Indian bread wheat cultivars, have been studied by various researchers. But such studies on Indian farmers' varieties/land races are lacking. One source of genes that can be utilized to meet these constantly emerging challenges to food production is the ancient farmer varieties (landraces).

No previous work on the study of diversity and population structure is done in large number of Indian wheat accessions. However, studies were carried out in few landraces from Oman (Zhang *et al.*, 2006) and US elite winter wheat (Zhang *et al.*, 2006). CYMMIT landraces are analysed for genetic diversity (Dreisigacker *et al*., 2005). So far as Indian wheats are concerned, Routray *et al.* (2007) has reported genetic diversity of 27 Indian landraces of bread wheat from hilly areas of Uttaranchal, and Mir *et al.* (2011) studied

genetic diversity among 263 Indian bread wheat cultivars, released during last 100 years, using simple sequence repeats (SSRs) but included mostly modern Indian wheat cultivars. In India, National Bureau of Plant Genetic Resources under Indian Council of Agricultural Research (ICAR-NBPGR) has a large collection, comprising wild and weedy relatives, of wheat germplasms including about 30000 accessions collected from India (indigenous collection; IC) and from other countries (exotic collections; EC) (Kumar *et al*., 2016b). However, so far, no systematic study has been carried out on genomic diversity and population structure involving large number of accessions that helps to understand the pattern and the spectrum of diversity, both at molecular and agro-morphological level, and supports the decision-making on exploration, preservation, conservation, and utilization of the germplasm accessions in wheat improvement programme.

Thus, with an aim of understanding the extent and pattern of diversity in a panel of six hundred and eighteen Indian wheat landrace accessions at genetic (using morphological traits) and genomic (using DNA markers) levels, the current study entitled "**Study of population structure and phylogeny in Indian wheat (*Triticum* spp.) landraces using DNA markers"** was undertaken with the following objectives:

1. Analysis of genetic diversity of Indian wheat landraces using DNA markers.
2. Study of population structure and phylogeny among the Indian wheat landraces.

CHAPTER 2

REVIEW OF LITERATURE

2. REVIEW OF LITERATURE

2.1 Wheat as a crop plant

2.1.1 History, origin, evolution and domestication

Wheat (*Triticum* spp. L.) is a self-pollinating annual plant, belonging to the sub-tribe Triticineae of the Tribe Triticeae in the family Poaceae (grasses). Further, being the oldest and the most widely grown crop since 10000-8000 BC (Heun *et al.*, 1997; Luo *et al.,* 2007; Nesbitt and Samuel, 1998) in the world, it is grown in India since 5000 years (Feldman, 2001). Wheat is believed to be originated from fertile half crescent located in Western Asia (Belderok *et al.,* 2000). Some of the earliest remains of the crop have been found in Syria, Jordan, and Turkey and date back almost 9,000 years. The evolution and domestication events of bread wheat is graphically illustrated (Figure 2.1) by Bevan *et al.* (2017). About 7 Million years (Myr) ago, an ancestral Triticeae gave rise to the AA and BB diploid precursor genomes of wheat. These formed the diploid precursor DD genome of goat-grass 5 Myr ago. Around 800 thousand years (kyr) ago, the tetraploid AABB genome of wild emmer wheat formed, which was domesticated in the region that is now south-west Turkey. About 8 kyr ago, a random hexaploidization event occurred between domesticated emmer wheat and a wild goat-grass, which contributed its DD genome, that formed the contemporary bread wheat genome (Wicker *et al.,* 2009; Bevan *et al.,* 2017).

Common wheat originated due to crossing between two independent polyploid wheats. The first event occurred during the geographical expansion of agriculture during which an ancestor called domesticated einkorn (*T. urartu*, 2n = 2x = 14, genome AA) was spread across Asia, Europe and Africa. In the course of time, it crossed with an unknown species presumably related to *Aegilops speltoides* (2n = 2x = 14, genome SS). This event might have given rise to cultivated species emmer wheat (*T. turgidum* ssp. dicoccum, 2n = 4x = 28, genomes AABB) (Yong *et al.,* 2006). Later the domesticated emmer grown in northeast Turkey was crossed with another grass species known as *Aegilops tauschii* (2n = 2x = 14, genome-DD) giving way to emergence of the hexaploid common wheat (T. *aestivum*, 2n = 6x = 42, genomes-AABBDD) with three similar genomes (Dubcovsky and

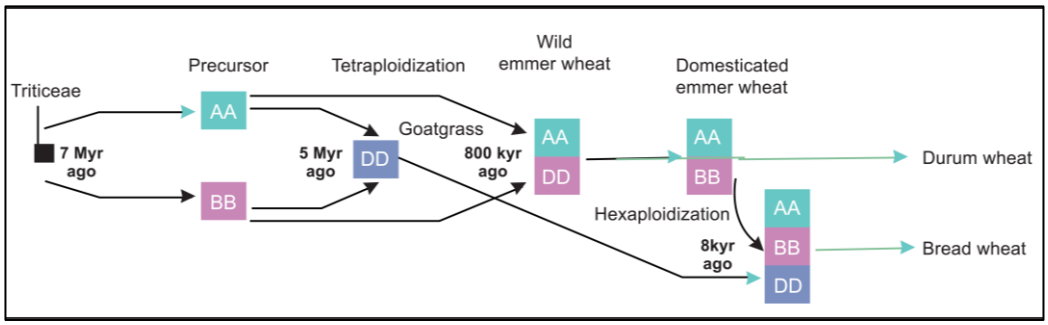


Figure 2.1 Evolution and domestication of common wheat (Adopted from Bevan *et al*., 2017).

Dvorak, 2007). However, later-studies involved DNA sequence data for analysis of these subgenomes (Yen *et al.,* 1996; Levy and Fieldman, 2002). On the basis of DNA sequence of two genes, *viz.,* Plastid acetyl-CoA carboxylase (*Acc1*) and Plastid 3-phosphoglycerate kinase (*Pgk1*) (Huang *et al.,* 2002), it is known with some degree of certainty that *T. uratu* (2n=14) is the donor of subgenome 'A', while the donor of subgenome 'D' is *Ae. tauschii* (Syn. *Ae. squarrosa*). Regarding subgenome 'B', though *Ae. speltoides* was once considered a probable donor (Yen *et al.,* 1996; Maestra and Naranjo, 1998), subsequent studies showed that *Ae. speltoides* resembles the *T. timopheevii* subgenome 'G' more closely than subgenome 'B' of bread- and durum- wheats. Therefore, Analysis of DNA sequence of genes was also not helpful in identifying the progenitor of subgenome 'B'. Thus, still there is no unanimity regarding the progenitor of subgenome 'B' of bread wheat (Nevo *et al.,* 2002) and there are speculations that donor of the B-subgenome might have lost its identity in the course of evolution and may never be discovered (Gupta *et al.,* 2005).

2.1.2 Taxonomy

In grass family *Poaceae*, about 35 genera including *Aegilops, Dasypyrum, Lophopyrum, Secale, Thinopyrum* and *Triticum* belong to recently originated sub-tribe Tritineae of Triticineae. Direct exchange of genetic material or polyploidization is easy in these genera as their species are easily hybridizable (Suvarna, 2001). According to Linnaeus (1953), seven cultivated species are included within the genus *Triticum*, while the genus *Aegilops* has wild relatives. This dichotomy was accepted among Taxonomists for over 200 years until two genera were merged into a single genus *Triticum* which was started by Stebbins (1956) and subsequently supported by Bowden (1959), and Kimber and Feldman (1987). The nomenclature suggested by Miller (1987), along with ploidy-levels and genomic constitutions of wild and cultivated wheat species, is illustrated in Table 2.1.

2.1.3 Cytogenetics

Cytogenetic studies by Sakamura (1918) and Kihara (1924) led to the understanding that various species of genus *Triticum* are polyploid series originated from a basic haploid set of seven chromosomes. The polyploid wheats (tetraploid and hexaploid) are amphiploids, composed of two or more different genomes, each derived from a distinct ancestor and behaving like a diploid. In hexaploid wheat, the seven corresponding chromosomes of the

Table 2.1 Ploidy-levels and component genomes of wild and cultivated wheat species (Adopted from Miller, 1987).

Species	Common Name	Genomic Constitution
Diploid (2n=14)		
Triticum bioeticum	Wild eikorn	AA
Triticum uratu	Two grained eikorn	AA
Triticum monococcum	Culltivated eikorn	AA
Triticum speltoides	Wild grass	SS
Triticum tauschii	Wild grass	DD
Tetraploid (2n=28)		
Triticum dicoccoides	Wild emmer	AABB
Triticum dicoccum	Cultivated emmer	AABB
Triticum durum	Macaroni/ durum wheat	AABB
Triticum turanium	Kamul	AABB
Triticum polonicum	Polish wheat	AABB
Hexaploid (2n=42)		
Triticum spelta	Spelt wheat (cultivated)	AABBDD
Triticum aestivum	Bread/common wheat	AABBDD

different genomes (A, B and D) are called homoeologous and twenty-one pairs of chromosomes are called homologous.

Polyploids behave like diploids due to presence of a gene '*Ph*' carried by chromosome 5B. This gene prevents the homoeologous chromosomes from pairing among themselves and, therefore, ite is called pairing inhibitor gene or diploidizing gene (Riley and Chapman, 1958). By ensuring regular segregation of chromosomes, diploidization not only prevent structural changes in the wheat chromosomes but also provides for high fertility due to presence of genetic stability and prevalence of heterosis. Genomic formulae of *Triticum Urartu* and *Aegilops speltoids*, for instance, was designated as AA and SS, respectively, based on the reports of Kihara (1924) as well as Kihara (1954) that there was no regular pairing of chromosomes of different diploid species in interspecific hybrids. All the three species, *Triticum durum* (Duram wheat), *Triticum dicoccum* (Cultivated emmer wheat) and *Triticum dicoccoides* (Wild wheat), have the same genomic formula AABB.

Though early cytogeneticists believed that A-genome was contributed by the *Triticum monococcum* (Einkorn wheat) as the A-genome donor (Kihara, 1954), subsequent molecular studies proved that *Triticum Urartu*, a biological species of eikorn wheat, is the donor of A-genome found in emmer and durum wheat (Dvorak *et al.*, 1993; Jiang and Gill, 1994). Other studies suggest a possible polyphyletic origin of the B genome to be from hybridization between two distinct allotetraploids with a common A genome (AAXX and AAYY) followed by rearrangement of the X and Y genomes within their descents (Zhang, 2006). Despite the availability of geographical, cytological and morphological evidences, the B-genome donor identity is still elusive in that *Triticum speltoides* (SS) or its closely related species was presumed to be the B-genome ancestor (Riley *et al.*, 1958) until it was confirmed through molecular evidences that B-genome is contributed by a closest extant relative of *Triticum speltoides* (Dvorak and Appels, 1982a; Dvorak and Zhang, 1982b; Waines and Barnhart, 1992; Dvorak *et* al., 1993).

Genome 'D' found in bread wheat is mostly contributed by a progenitor related to *Aegilops squarrosa* (*L.*) which resulted in tetraploid genome 'AABB' gaving rise to hexaploid genome 'AABBDD'. Genome D has majorly impacted the world-wide distribution of wheat as it harbors genes that impart the bread making characteristics and

dough rheology of wheat flour. Hence, the most important wheat is the hexaploid bread wheat and, therefore, it is also called common wheat (*Triticum aestivum L*. Syn. *T. vulgare* Host).

2.1.4 Economic importance

Wheat (*Triticum spp*. L.) is an edible cereal and is one of the world's most important food crops. Being a staple in dozens of country's diets, especially in temperate zones, it is consumed throughout the world. Wheat is often cultivated in wind-swept environments with low rainfall and cold temperatures that tropical crops like rice and corn cannot withstand (Gibson and Benson, 2002).

2.1.4.1 Wheat economy

The harvest grain has been consumed by man for centuries and is also used as animal feed. Wheat plays an important role in human nutrition since centuries, as it is a major source of energy, protein and dietary fiber. Wheat alone provides nearly 55% percent carbohydrates and 20 percent of the food calories (Wiese, 1987) and proteins in human nutrition, consumed globally, and staple food of about two billion people (i.e. approximately 40% of the world population) (Breiman and Graur, 1995). Thus, wheat is not only the top most cultivated cereal crop in the world (>200 mha annually), it also helps to sustaining food security (Shiferaw *et al.,* 2013). While approximately 95% is hexaploid bread wheat, duram wheat shares the remaining 5% of the global cultivated wheats.

Due to its importance and increasing demand, it is a key commodity to eradicate global hunger not only by ensuring sufficient production to feed a world population that will grow by 50% and reach 9 billion by 2050, but also by guaranteeing access to food (FAO 2010). Furher, wheat trade plays a key role in influencing international relationship among global economies in addition to impacting political and economic relationships among nations (Sharma, 2016).

2.1.4.2 Wheat products

Wheat is used in various forms such as baking, making bread, pastas, and biscuits. Wheat grain is also a source of alcoholic beverages and wheat starch is used for making pastes. Wheat bran from flour milling and vegetative plant parts are used as livestock feed.

Also. wheat can be potentially used as non-food raw material in manufacturing industries including paper, corrugated boxes and boards, shiffening/ surface coating agents, adhesive, as fermentation substrate in the production of pharmaceutical products and beverages, etc., (Jones *et al.,* 1987; Sharma, 2016). Wheat straw serves as building material as it is used in traditional way of making thatches, panels and in crafts apart from being used as fuel for warming during winter and cooking. In India, wheat forms the major staple food consumed in different type of recipes including chapati, sphaghetti, noodles, bread, doughnuts, pizzas, etc. Kheer is made from vermicelli of durum wheat while semolina is made from what is called in vernacular as rava or sooji.

2.1.4.3 Wheat production

While global production of wheat has been recorded 729 million tonnes harvested from 220 million hectares with the average productivity of 3.3 t/ ha, India accounts for 13% of the total global wheat production (FAOSTAT, 2014). In the year 2014, India was the world second-largest producer, next to China, of the record production of 95.58 million tonnes harvested from 30.47 million hectares with an average productivity of 3.15 t/ha (FAOSTAT, 2014). However, the wheat average productivity in India is lower than the world average.

2.1.5 Wheat cultivation in India

Wheat cultivation in India started 5000 years ago (Feldman, 2001). Based on requirement of vernalization; a period of low temperatures necessary to induce flowering in crops, for growth habit and development; bread wheat is divided into two ecotypes: spring and winter wheats. Winter wheat requires vernalization by exposure to a temperature in the range of 3-8° C for several weeks, although the exact temperature and duration varies with different genotypes. Wheat has genetic variability for genes that control flowering time, which imparts adaptability to a wide range of environments (McMaster *et al.,* 2008). Vernalization is an adaptation to ensure that seed is produced after the harsh winter-climate. Spring wheat has little or no vernalization requirements; so, it can be cultivated during spring as well as summer. In India, spring wheat is cultivated in northern states during winter season where the winter is too cold and severe for winter wheat, and in southern states of India where temperatures are very mild. Winter wheat is planted in the

fall and requires exposure to certain periods of cold, depending on the variety, to initiate flowering (Malla *et al.*, 2011). It is grown in areas where the summer is too hot and/or dry for wheat to grow but the winter is suitable.

Bread wheat (*Triticum aestivum*), durum/macaroni wheat (*Triticum durum*) (durum wheat), and dococcum wheat (*Triticum dicoccum*) are the three types of wheat cultivated in India. Wheat is grown in India as, as *Rabi* crop and the sowing is taken-up during winter, as cool winter at sowing and hot summer at harvesting is favorable. Indian wheat belts fall under the following five agro-climatic zones.

1. *The North-Western Plains Zone*: This zone covers region of Jammu, political states of Punjab, Rajasthan, Haryana, and western part of Uttar Pradesh. The terrain is plain and wheat is sown in late October - November and the harvesting begins in the middle of April.
2. *The North Eastern Plains Zone*: This zone comprises eastern part of Uttar Pradesh, states of Bihar, West Bengal, Assam, Odisha, Manipur, Tripura, Meghalaya, Nagaland, Mizoram, Arunachal Pradesh and Sikkim. Since rice is harvested later in this area, wheat can be sown only in late November or early December, while Harvesting is done during March - April.
3. *The Central Zone*: This zone covers Bundelkhand area of Uttar Pradesh, south-eastern part of Rajasthan, states of Madhya Pradesh and Gujarat. Almost 75 per cent of the wheat cultivated in this zone depends on rainfal and the best quality durum wheat is produced in this zone.
4. *The Peninsular Zone*: This zone covers peninsular States of Maharashtra, Karnataka, Tamil Nadu, Andhra Pradesh, and Telangana. Sowing is usually completed by early November, and harvesting begins in the second half of February. Wheat is produced the earliest in this zone.
5. *The Northern Hill Zone*: This zone comprises the hilly states of Assam, Sikkim, Kashmir, Himachal Pradesh, Uttar Pradesh and West Bengal. Wheat is sown in October and harvested in May-June. The crop remains dormant in the cold months of November to March and starts growing as the temperature rises in April.

2.2 Genetic diversity

The term "diversity" refers to the range of variation, variety or differences found among a set of parameters (morphological or molecular) of different individuals and/or populations. Such variations contribute to biodiversity, which specifically refers to the variety and variations within assemblages including the genetic differences among them, whether naturally occurring or induced artificially. Biodiversity has also been defined as the variety of life and its processes that encompasses various living organisms, communities, and ecosystem in which they occur (Keystone-Center, 1991).

The basic building blocks are the genes contained in plants and animals which, by their diversity, can enable the whole organisms adapting to the changing environment. Plant genetic diversity, a useful parameter that can be transmitted genetically from parents to offspring, is the source of tremendous variations in plants that support all other forms of life on earth, covering a wide range at the evolutionary as well as ecological levels. Food and human needs for millennia have been heavily dependant on diversity in plants; and, it continues to be so as the diversity is the basis for development of plant traits of economic importance (Tanto and Demissie, 1996). It is important for broadening the genetic base and can be exploited *via* heterosis (Melchinger, 1999).

2.2.1 Molecular basis of wheat genome diversity

Insertion of enormous amount of retrotransposable element lead to the complexity of the wheat genome. Micro (sub-megabase) level collinearity among the cereals is disrupted due to the fast replacement of repetitive elements and gene deletions. In view of genomic changes that are possible within a given time span, conservation of genes between species tends to imply an important functional, or regional, constraints that do not permit changes in genomic structure (Breen *et al.,* 2010). Reduced height (or dwarfness) in wheat is imparted by two very important genes, *Rht1* and *Rht2,* found in a Japanese cultivar “Norin 10”, which inherited these two genes from Japanese landrace “Shiro Daruma” (Kihara, 1983). When Norman Borlaug of the joint Rockefeller-Mexico project observed extremely low wheat productivity in Mexico as the tall varieties were prone to lodging and non-repsonsive to added fertilizer inputs, he speculated that by reducing the height of the wheat plant, it can be made to suffer less lodging even under the higher input levels. By

introgressing *Rht1* and *Rht2* genes into the new varieties, Borlaug, ultimately, was able to prove his speculation, and eventually resulted in what has now been termed the "Green Revolution" (Krull and Borlaug, 1970; Borlaug, 1988). Though it was originally believed that these genes contributed to higher production simply through reduced height that reduced lodging, it is now understood that they impart better nutrient uptake and tillering capacity and directly contribute to enhanced yield (Krull and Borlaug, 1970; Gale *and* Youssefian, *1985*). Thus, depending on their geographical regions, landraces have specific genetic background that can be used in genetic research program (Allard, 1996). In addition, landraces are important genetic resources that improve gene pools of modern cultivars by introducing new alleles (Feldman and Sears, 1981; Nevo and Payne, 1987). Therefore, for effective utilization of germplasm in crop improvement programmes, substantial knowledge of full genetic potential of the germplasm, in terms of genetic and genomic diversity, is essentially needed (Maity and Das, 2015).

2.2.2 Gene distribution in wheat genome

Gene distribution study in wheat by various research groups (Gill *et al.,* 1996a, b; Sandhu *et al.,* 2001) suggests that genes are not distributed randomly in wheat genome: there are gene-rich regions (GRRs) and gene-empty regions (GERs) (Gill, 2004). GRRs, each with less than 50 kb size in contrast to 300 kb in rice (Gill, 2004), are interspersed within gene-poor regions. The genome contains 48 GRRs with an average of about seven such GRRs per homeologous group. GRRs vary with respect to gene size and gene density with a general trend of increased gene density towards a distal part of chromosome arms. An average recombination rate of 0.003 cM/kb, ranging from 0.0 to 0.007 cM/kb is reported in wheat (Sandhu, 2004; Gupta *et al.,* 2005).

2.3 Need for molecular and morphological diversity studies in wheat

Genetic diversity in populations of wild precursors of wheat has been eroded by domestication, due to which a limited range of diversity is present in landraces that were initially selected and adopted for cultivation (Figure 2.2). These landraces, initially selected and adopted for cultivation, have a limited range of diversity compared to populations of their wild precursors. Subsequent crop improvement efforts of modern agriculture

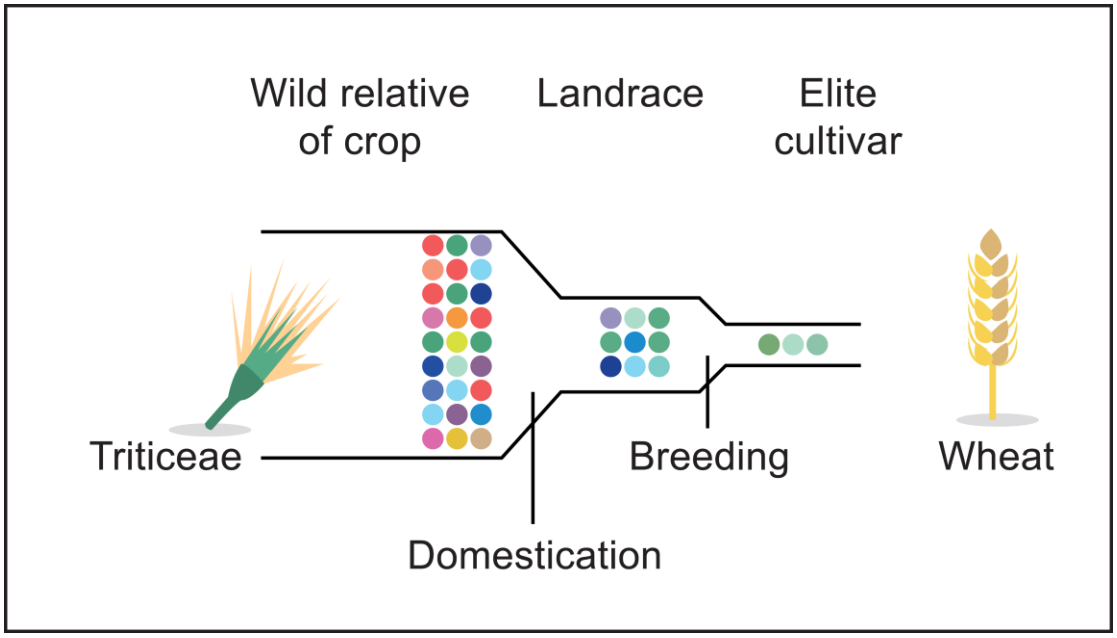


Figure 2.2 Genetic erosion during domestication of common wheat (Adopted from Bevan *et al*., 2017).

produced elite cultivars that have further narrowed-down the already limited range of genetic diversity that was present in their precursor landraces (Bevan *et al.,* 2017).

Since genetic diversity in the collections is critical to the world's fight against hunger, the global effort to assemble, document, and utilize these resources is enormous in making a range of genetic diversity accessible to breeding programs. ''Green Revolution'', realized through reduced plant height and decreased disease, demonstrated the tremendous impact that genetic resources can have on production (Hoisington *et al.,* 1999).

The International Maize and Wheat Improvement Center (CIMMYT)'s wheat breeding program has been increasingly incorporating wild wheat germplasm and its allelic diversity into their breeding programs in the form of synthetic hexaploids (Trethowan *et al.,* 2008; Mujeeb-Kazi, 2008). CIMMYT evaluates non -adapted genetic resources, such as landraces or wild species, to search for stress-adaptive traits. CIMMYT also uses high-throughput screening tools such as infrared-thermometry to discover novel genes, characterize parents for use in crosses, and to select in early generations.

2.4 Methods of genetic and genomic diversity studies

Genetic diversity can be estimated based on various criteria: pedigree analysis (Barrett *et al.,* 1998), morphological traits (Schut and Stam, 1997; Maric *et al.,* 1998; Casadesus *et al.,* 2007; Zarkti *et al.,* 2012; Malik *et al.,* 2014), biochemical markers (Cox *et al.,* 1985; Metakovsky and Branlard, 1998), and/or molecular markers (Rao and Riley, 1994; Karp *et al.,* 1996; Gupta *et al.,* 1999; Manifesto *et al.,* 2001; Pagnotta *et al.,* 2005). As morphological traits are the best alternative to their molecular counterparts for assessing genetic diversity (Hailu *et al.,* 2010; Mengistu *et al.,* 2015); agromorphological characterization is the primary step for conservation and utilization of plant genetic resources (Hafida *et al.,* 2012). Thus, diversity can be assessed using both morphological traits (genetic diversity) and molecular/DNA markers (genomic diversity). Wheat has gained agronomically and nutritionally important status among other cereal crops because of its large genome size 16 x 199 bp (Bennett and Smith, 1976) and a wide range of uses. This has been reflected in the development of many cultivars by the wheat breeders. Narrow genetic background has restricted improved varieties of main crop tolerance to biotic and abiotic stresses. Early farmers used to grow genetic blended cultivars (landraces)

on a very large scale and thereby extended genotypes with effective genetic make-up against various plant diseases. Today it is realized that the use of genetically diverse varieties is an effective strategy to minimize genetic vulnerability (Smale *et al.*, 2002) to biotic and abiotic stresses.

2.4.1 Different types of marker systems useful for studying genomic diversity

The normal growth and development of the plant requires genome, a fundamental and physiological unit in terms of a set of chromosomes (Kihara, 1930), and the term 'genome' was coined by Winkler (1920). Subsequently, the field of `genome analysis' rapidly advanced and, recently, constitutive genomes of organisms and their related genera and species are being analysed with the help of various tools; such as biochemical markers including isozymes (proteins) that are detected through electrophoresis or immunochemical reactions; cytological markers including karyotype analysis, DNA-DNA hybridization, chromosome staining, chromosome-pairing, etc.; and, DNA markers including AFLP, RAPD, SSRs, SNP, etc.

2.4.1.1 Morphological markers

Morphological markers are visually characterized as phenotypic traits such as flower color, seed shape, growth habits or pigmentation, etc. In conventional plant breeding programs, breeders generally select desired plant types based on direct morphological data as well as morphological markers tightly linked to a trait of interest. However, the morphological markers are not widely used due to their limited availability, low polymorphism and influence of environmental conditions, pleiotropism, expressivity, etc.

2.4.1.2 Biochemical markers

Allozyme was the first true molecular marker established to distinguish protein variants in enzymes (Schlotterer, 2004) and was first developed and widely used in population genetics (Hamrick and Godt, 1990). Isozymes are structurally different molecular forms of an enzyme with the same catalytic function. Mutation leads to altered amino acid that brings about alterations in net charge or the spatial structure (conformation) of the enzyme molecules which causes differential electrophoretic mobility of altered protein compared to the original protein. Since differential electrophoretic mobility of

proteins can be easily detected, isozymes are used as markers in the past several decades for various analyses/purposes in life sciences, *viz*., genetic variability and taxonomy, selections in plant breeding, phylogenetic relationships, population genetics, developmental biology and plant genetic diversity (Bretting and Widrlechner, 1995).

Two main seed-storage proteins in wheat, glutenins and gliadins, which display allelic variation at various loci and, therefore, partly responsible for the differences observed in the cooking properties of wheat, also served as valuable tools to assess genetic diversity. The high molecular weight (HMW) glutenin; in durum wheat from Israel (Branlard *et al*., 1989) and in *T. dicoccoides* form Jordan (Nevo and Payne, 1987; Levy and Feldman, 1988; Levy *et al*., 1998) and Turkey (Ciaffi *et al*., 1993); have been thoroughly analyzed for the variations in their subunits. Though these studies demonstrated the presence of wide spectrum of polymorphism at *Glu*A1 and *Glu*B1 loci, glutenin allelic variatins in case of *T. dicoccoides* is reported to be partly due to eco-geographical variations. However, protein-based markers suffer disadvantages such as environmental influence, limited in number and low allelic variations (i.e., low level of polymorphism).

Like morphological markers, the biochemical markers are also limited in number and are influenced by environmental factors or the developmental stage of the plant (Winter and Kahl, 1995).

2.4.1.3 Cytological markers

Relative chromosome lengths (Schlegal and Mettin, 1978) and arm-ratios (Kimber, 1971), were analyzed in wheat and established its phylogenetic relationships with other related species. Better understanding of the content and distribution of repeated DNA in wheat was started with the development and application of differential staining (chromosome banding) and DNA hybridization methods (Flavell and Smith, 1976; Ranjekar *et al.*, 1978). The later techniques such as genomic *in situ* hybridization (GISH) and fluorescent in-situ hybridization (FISH) revolutionized plant cytogenetics apart from finding their application in genome mapping project as well as genome sequencing projects, particularly in complex genomes like that of bread wheat (Wemer *et al.*,1992; Jiang and Gill, 1994). Further, GISH and FISH were extensively deployed for detection of alien chromosome segments (Mukai and Gill, 1991; Schwarzacher *et al.*, 1992;

Schwarzacher, 1996), analyzing DNA sequence organization, species relationships and evolution, and nuclear architecture (Heslop-Harrison *et al.*, 1990). However, presently available integrated wheat genetic map is based on several variants of DNA marker systems such as RFLP, RAPD, microsatellites, AFLPs (Anderson *et al.*,1992; Xie *et al.*,1993; Hohmann *et al.*, 1994; Roder *et al.*, 1995; Nelson *et al.*,1995a; Van-Deynze *et al.*, 1995), SNPs and DArT markers.

2.4.1.4 DNA/Molecular markers

The assessment of genetic diversity using molecular markers assumes a greater significance as it is a prerequisite for any sound breeding program (Surwenshi *et al.*, 2011). Since DNA marker systems, applied in genetic analysis for the first time in the 1980s, offer several advantages, over other alternative systems of genetic markers, such as unlimited number, not affected by the environment, and tissue or developmental stage independendence, etc., (Park *et al.*, 2009).

Plant DNA-level variation forms the basis of variations in its morphological traits. Therefore, it is important to analyse and understand DNA-level variations to decide upon the conservation and utilization of a given genetic resource. Plant DNA-level variations, also termed as molecular variations/ DNA polymorphism, can be analysed using several types of DNA markers. Since molecular markers depend on the naturally occurring DNA polymorphism, their presence can be readily detected, and inheritance can be easily monitored. In addition, unlike protein markers and morphological markers, DNA markers, which are independent of the environmental influence, follow Mendelian inheritance as single gene. Therefore, DNA markers are the best candidates for accurate detection, selection, evaluation and utilization of genetic diversity in plant breeding. Genetic polymorphism, defined as the simultaneous occurrence of two or more forms of a trait among the individuals within in the same population (Maurya, 2014), is a prerequisite for validation and application of molecular markers (DNA markers).

Since DNA extraction is cost- and labour-efficient; and, as it can be effectively obtained from any plant material at any developmental stage; its analysis has become practically possible and economically feasible (Kumar *et al.*, 2009). The detection and exploitation of DNA sequence variability using molecular markers mark the most

significant developments in the field of molecular genetics (Semagn *et al.,* 2006). It has also been proved to be helpful in understanding the genetic relationships, evolutionary trends and finger-printing of varieties.

A molecular marker, a DNA segment that represents the differences at the genomic level, may or may not be associated with phenotypic expression of a trait. Molecular markers offer numerous advantages, over their morphological counterparts, due to their stability and detectability regardless of tissues type, development, growth stage, differentiation, or defense status of the cell, in addition to non-confounding effect by the environmental and ecological variations.

In general, the term 'marker' is used to refer 'locus marker' (Stansfield, 1986). Each gene has a particular place called 'locus' along the genome. Upon mutation, genes become modified to one or more alternate forms called 'alleles' (or allelic forms). All the allelic forms of a given gene are called alleles of that gene if, and only if, they are occurring at the same locus on homologous chromosomes. When allelic forms of a locus (i.e., among homologous chromosomes) are identical, the genotype of an individual is called 'homozygote' (for this locus), whereas non-identical allelic forms make it 'heterozygote'.

The first application of DNA based information in wheat genomic diversity analysis can be attributed to the works of Kihara (1930) where he developed methods for genome analysis which consisted of crossing diploid species with each other and with polyploids to understand species diversity and relatedness based on amount of chromosome-pairing and fertility in hybrids. Later-studies, on characterization of germplasm diversity in wheat, were based on morphological characters such as variation in heading time, earliness, winter habit, and coleoptile length (Scarascia-Mugnozza and Porceddu, 1974; Qualset and Puri, 1974a, 1974b; Jain *et al.,* 1975; Porceddu, 1976).

In subsequent years, the molecular markers, especially DNA-based markers, have been extensively used in many areas of research such as gene mapping and tagging (Kliebenstein *et al.,* 2001; Karp and Edwards, 1997), characterization of sex (Flachowsky *et al.,* 2001; Martinez *et al.,* 1999), analysis of genetic diversity (Erschadi *et al.,* 2000; Palacios *et al.,* 1999; Lerceteau and Szmidt, 1999; Godt and Hamrick, 1999), genetic

relatedness (Rao *et al.*, 1997; Brookfield, 1992), linkage map construction, and marker assisted breeding (Kalia *et al.*, 2011).

2.5. Classification of DNA markers

DNA markers can be broadly classified into first generation markers that are based on hybridization/enzyme-mediated analysis, second generation markers that are based on polymerase chain reaction (PCR)-mediated analysis, and third/next generation markers that are based on nucleotide sequence analysis.

In hybridization-based markers, the DNA profiles are visualized by southern-blotting of a labeled probe, obtained from a DNA fragment of known sequence, with restriction enzyme-digested test-DNA sample. In case of PCR based markers, the primers of known sequence and length are used to prime the amplification of a test-DNA template followed by visualization of amplicon profile upon gel electrophoresis and staining with DNA inter-calating dye such as ethidium bromide. The invention of highly versatile polymerase chain reaction (PCR) (Saiki *et al.*, 1985), and the subsequent exploitation of the thermo-stable DNA polymerases for this extremely thermo-sensitive PCR-reaction (Saiki *et al.*, 1988) introduced a multitude of new possibilities in molecular marker research changing the total scenario of molecular biology.

Each marker has its own advantages and disadvantages (Table 2.2), suggesting that the choice of the marker is dictated by the purpose of the study, prior sequence-information and the laboratory facilities available. Some of the hybridization as well as PCR-based marker systems are dealt in the subsequent sections.

2.5.1. First generation (hybridization/enzyme-based markers)

The first-generation DNA marker system employed southern blot-based markers such as RFLPs (Restriction Fragment Length Polymorphism). RFLP analysis involves restriction endonuclease-mediated digestion of test DNA samples, gelelectrophoretion to resolve digested DNA according to different fragment lengths, and southern hybridization with radio-active labelled probe-sequence (Southern, 1975) to visualize banding patterns.

The first application of RFLPs for analyzing DNA sequence polymorphism was by Grodzicker *et al*. (1975) for mapping temperature- sensitive mutation in adeno-virus

Table 2.2 Advantages and disadvantages of different types of DNA markers (Modified from Idrees and Irshad, 2014).

S.No.	Advantages	Disadvantages
1.	*Random Amplified Polymorphic DNA* (RAPD) (Williams *et al*., 1990)	
	Inexpensive and easy to perform, requires less amount of DNA, no sequence information, moderate equipment required, and of anonymous origin.	Dominant, fail to distinguish homo and heterozygotes, no ooportunity for candidate gene mapping, and low to medium reproducibility.
2.	*Restriction Fragment Length Polymorphism* (RFLP) (Botstein *et al*., 1980)	
	Codominant, offers opportunity for genome, QTL and comparative mapping, requires moderate equipment, anonymous and genetic origin, andvery high reproducibility.	Difficult, limited by the restriction site, requires very high amount of DNA, and offers limited candidate gene mapping potential.
3.	*Amplified Fragment Length Polymorphism* (AFLP) (Vos *et al*., 1995)	
	Moderate ease of development, very good genome and QTL mapping potential, moderate to expensive equipment required, anonymous origin, and medium to high reproducibility.	Moderate to difficult, dominant, limited by the restriction site, medium to High DNA concentration required, very limited comparative mapping potential, and useless for candidate gene mapping potential.
4.	*Arbitrary Primed PCR* (AP-PCR) (Welsh and MacClelland, 1990)	
	Similar to RAPD except the size of the primer variation (5-8 oligonucleotides), easy, requires less amount of DNA, no sequence information, inexpensive, moderate equipment required and of anonymous origin.	Dominant, do not distinguish homo and heterozygotes, no potential for candidate gene mapping, and low to medium reproducibility.

Contd…

Table 2.2 Contd…

S.No.	Advantages	Disadvantages
5.	*Simple Sequence Repeats* (SSRs) (Hearne *et al.*, 1992)	
	Codominant, easy to moderate, genome and QTL mapping potential is good, easy, medium to high reproducibility, and of anonymous origin.	Limited to the size of the genome and number of simple sequence repeats, medium to high amount of DNA required, limited comparative mapping potential, useless for candidate gene mapping, moderate to expensive equipment required, and development is difficult.
6.	*Sequence Characterised Amplified Regions* (SCARs) (Paran and Michelmore, 1993)	
	Specific locus representing a single RAPD fragment, codominant, and highly reproducible.	Prior sequence information is required, and specific primers are required.
7.	*Single Primer Amplification Reaction* (SPARs) (Gupta *et al*., 1994)	
	Core motifs of microsatellite DNA more or less similar to those of RAPDs	Similar to disadvantages of RAPDs.
8.	*Single Stranded Conformational Polymorphism*(SSCP) (Orita *et al*., 1989)	
	Similar sized DNA fragments can be distinguished, based on the mobility of the single stranded DNA, and cost effective.	Polymorphism levels will be restricted.
9.	*Sequence Tagged Microsatellites* (STMS) (Beckmann and Soller, 1990)	
	Single locus, multiallelic, co-dominant, highly reproducible, and flanking regions of the micro satellites are used for designing primers.	Prior sequence information is required.

Contd…

Table 2.2 Contd…

S.No.	Advantages	Disadvantages
10.	*Sequence Tagged Site/Expression Sequence Tagged Site* (STS/EST) (Paran and Michelmore, 1993)	
	Codominant, high reproducibility, moderate ease of development, good genome and QTL mapping potential, and excellent candidate gene mapping potential.	Limited by the number of enzyme genes and histochemical enzyme assays, less DNA required, easy, development is expensive, and moderate to expensive equipment is required.
11.	*Thermal Gel Gradient Electrophoresis* (TGGE) (Rosenbaum and Riesner, 1987).	
	Separtion of fragments according to their mobilities, and better resolution can be achieved.	Denaturation conditions are high which may generate a lot of heat.
12.	*Variable Nucleotide Tandem Repeats* (VNTR) (Nakamura *et al*., 1987)	
	Variable number of tandem repeats are generated mostly on microsatellites.	Samples of identical or closely related species only can be analyzed.
13.	*Cleaved Amplified Polymorphic Sequences* (CAPS) (Lyamichev *et al*., 1993)	
	Specific for a particular locus and codominant.	Limited by the restriction site, and prior sequence information is required.
14.	*Sequence Amplified Polymorphic Loci* (SAMPL) (Morgante and Vogel, 1994)	
	Large scale analysis of genome, yields the best fingerprints when prior information on SSRs is not available, hypervariable loci are targeted, and more discriminative.	Variation in between very closely related genotypes can be determined.
15.	*Microsatellite Primed Polymerase Chain Reaction* (MP-PCR) (Meyer *et al.*, 1993)	
	Primer length is varied (10-20 oligonucleotides) similar to SSRs.	Limitations similar to those of SSR markers.
16.	*Arbitrary Primed Polymerase Chain Reaction* (AMP-PCR) (Meyer *et al*., 1993)	
	Specificity of the primer annealing is enhanced.	Large number of microsatellite primers must be screened, and optimal ones must be selected.

Contd…

Table 2.2 Contd…

S.No.	Advantages	Disadvantages
17.	*Inter Simple Sequence Repeats* (ISSR) (Hearne *et al*., 1992)	
	MP-PCR with 3' anchored primers are referred as ISSR	Polymorphism levels may be low due to absence of hypervariable regions.
18.	*Amplicon Length Polymorphism* (ALP) (Ghareyazie *et al*., 1995)	
	More amount of variation can be detected.	No specificity is observed.
19.	*Allele Specific PCR* (AS-PCR) (Sarkar *et al*., 1990)	
	Specificity is more and is applicable in applied aspects.	Wide variations cannot be detected, and it is very specific.
20.	*DNA Amplified Fingerprints* (DAF) (Caetano-Anolles *et al*., 1991)	
	Short primers (5 bp), and a large number of variations can be detected.	No specificity is observed, and primer binds to the DNA at random.
21.	*Randomly Amplified Microsatellites* (RAMS) (Ender *et al*., 1996)	
	Amplification is at random, and large variations which may be useful are detected.	No specificity observed, and primer binds at random throughout the entire genome.
22.	*Retrotransposan Microsatellite Amplified Polymorphism* (REMAP) (Kalender *et al*., 1999)	
	Useful in detecting variations in the transposable elements and it can be further used in expression studies.	The variation in the transposon sometimes may not be useful which may be futile.
23.	*Specific-Amplicon Polymorphism* (SAP) (Williams *et al*., 1991)	
	Very specific and helpful in detecting the variations at specific level.	Variations at random cannot be detected.

Contd…

Table 2.2 Contd…

S.No.	Advantages	Disadvantages
24.	*Single Nucleotide Polymorphism* (SNP) (Nikiforov *et al*., 1994)	
	Very specific and can detect polymorphism at the nucleotide level which may be helpful in appied aspects.	Not applicable for diversity analysis.
25.	*Microsatellite Simple Sequence Length Polymorphism* (SSLP) (Rongwen *et al*., 1995)	
	Detects variations at random with short sequences of primers.	No specificity is observed, and variations detected are at random.
26.	*Minisatellite Simple Sequence Length Polymorphism* (SSLP) (Jarwan and Wells, 1989)	
	There is less specificity as the primer length is comparatively more, and variations can be useful at applied level.	No specificity is observed, the variations are detected at random.
27.	*Selective amplification of start codon polymorphic loci* (SASPL) (Alsammanl *et al*., 2017)	
	High genomic coverage, amenable for *in silico* identification of potential genes near primers	Costly, laborious and shares all the disadvantages of other markers.

serotypes. Subsequently, after Botstein *et al.* (1980) deployed RFLPs for the first time to construct a genetic linkage map in human, and adopted for plant genomes (Helentjaris *et al.,* 1986; Weber and Helentjaris, 1989), they were most widely used as they follow simple Mendelian inheritance, in addition to being co-dominant in nature, and are very reliable markers in linkage analysis and breeding.

However, there are many disadvantages inherent to the very technique of RFLP markers, *viz*., involvement of radioactive isotopes making the analysis relatively hazardous, time-consuming, expensive and laborious; and its limited informativeness making it necessary to try many markers, one at a time, to get sufficient level of polymorphism. Because of these drawbacks, RFLP markers did not become popular. However, RFLP marker system, with further modifications, can be converted to PCR-based marker systems as described under second generation markers.

2.5.2 Second generation (PCR- based markers)

The second-generation DNA-based molecular markers were driven by the invention of polymerase chain reaction (PCR) (Mullis *et al.,* 1986). PCR is the most accurate *in vitro* technique that can be utilized to enzymatically replicate (amplify) small quantities of DNA without going for cloning, and without requiring a large amount of ultra-pure DNA. It is used to amplify a short (usually up to 10 kb), well-defined part of a DNA strand from single gene or just a part of a gene.

Since the following two major disadvantages of southern blot-based markers were overcome by PCR-technology, genetic and ecological analysis saw a revolution. First, it requires only small amount of DNA making it possible to carry out analysis at very early stages. Second, it enabled large scale experiments to be conducted in less time through multiplexing and high-throughput analysis; making the research simple and inexpensive. The various type of PCR-based markers such as RAPD, AFLP, ISSR, IRAPs and SSR, etc., relies on the use of PCR primers that bind to multiple, or specific, sites on template DNA. This can be achieved by using either short PCR primers (Randomly Amplified Polymorphism, RAPD) (Williams *et al.,* 1990), PCR primers that are complementary to repetitive elements such as microsatellites or simple sequence repeats (SSRs), Inter-Simple-Sequence-Repeats (ISSRs) (Zietkiewicz *et al.,* 1994), or retrotransposans (Inter-

Retrotransposon Amplified Polymorphism, IRAPs) (Kalendar *et al.*, 1999). Alternatively, restriction fragments could be amplified by adding linkers and subsequent selective amplification as in case of AFLP (Zabeau and Vos, 1993; Vos *et al.*, 1995).

2.5.2.1 First generation-derived PCR-based markers

With modification of RFLP marker system, the following PCR-based marker systems can be obtained.

*Sequence tagged sites (STS***)**

A sequence-tagged site (or STS) is a short (200 to 500 base pair) DNA sequence that is not repeated in the genome and whose location and base sequence are known. Nucleotide sequence information of RFLP probes specifically linked to a desired trait/ locus can be utilized to design PCR primer sequences giving a polymorphic banding pattern specific to linked trait/locus. This type of PCR-based markers is called sequence tagged sites (STS), and is extremely useful for studying the relationship among several species at a specific locus (Bustos *et al.*, 1999) and can be used in the construction of physical map.

Allele specific associated primers (ASAP)

In this type of markers, primers are designed specific to an allele whose sequence information is known and PCR-amplification is carried out using stringent annealing temperature. Allele specific associated primers (ASAP) are useful to tag genotype and the gene of interest (Gu *et al.*, 1995).

Expressed sequence tag markers (EST)

Expressed Sequence Tag Markers (EST) are obtained by partial sequencing of random cDNA clones. Based on the sequence information, EST-specific primers are designed and used to amplify genomic DNA to detect EST location in the genome using genomic library.

Subsequent to their introduction by Adams *et al.*, (1991), EST-tagged markers were extensively used in genome sequencing and mapping programs as they help to map human genes onto chromosomes after PCR amplification.

Single strand conformation polymorphism (SSCP)

This is a powerful technique that can identify heterozygosity among DNA fragments of identical molecular weight, and, therefore, popularly used for detecting point mutations (Orita *et al.*, 1989).

Random amplified polymorphic DNA markers (RAPDs) and RAPD-derived markers

A single arbitrary primer anneals to complimentary strands of the template DNA at two different sites in opposite direction. Introduced by Welsh and McClelland (1990), this method employs a single marker that act as both forward and reverse primer sothat a discrete PCR-amplicon product is obtained if and only if two priming sites are within the amplification range PCR chemistry. Therefore, RAPD markers produce multi-locus amplification of several discrete loci.

By introducing slight variations in the stringency of the PCR technique, primer sequence, or reaction assay; one can make the RAPD-derived marker systems, *viz.*, DAF, SCAR, CAPS, AFLP, AP- PCR, and RAMPO that are briefly discussed in the following paragraphs.

DNA amplification fingerprinting (DAF)

This technique, introduced by Caetano-Anolles *et al.* (1991), gives simple banding pattern, when very optimized reaction conditions drive the assay reaction involving a single arbitrary primer of only 5 bases to PCR-amplify test DNA samples. Such banding patterns are analyzed by polyacrylamide gel electrophoresis. It has potentiality to be useful for DNA fingerprinting.

Arbitrary primed polymerase chain reaction (AP-PCR)

Welsh *et al.* (1991) developed this technique in which a single, 10 to 50 base-long primer is used to produce DNA amplification patterns, under non-stringent conditions of annealing.

Sequence characterized amplified regions (SCAR)

SCAR markers are developed after sequencing the polymorphic band and re-designing the primer making it sinlge-locus specific. Therefore, like in case of STS

markers, SCAR markers are more reproducible as compared to RAPD. These markers, introduced by Michelmore *et al*. (1991) and Martin *et al.* (1991), are mostly dominant markers but can also act as co-dominant markers when amlicons are digested with tetra-cutting restriction enzymes, as done in case of sex identification in papaya using SCAR marker (Parasnis *et al.*, 2000).

Cleaved amplification polymorphic sequence (CAPS)

Using the prior knowledge of sequence information available, for instance, in sequence data-bank, primers are designed, and the resulting PCR products are restriction digested to obtain electrophoretic patterns (Koniecyzn and Ausubel, 1993; Jarvis *et al.,* 1994).

Randomly amplified microsatellite polymorphism (RAMPO)

In this marker technique, first RAPD primers are used to amplify genomic DNA followed by electrophoretic separataion and driying of gel for hybridizing with microsatellite oligonucleotide probes. Therefore, RAMPO markers offer advantages of several markers, *viz*., RAPD (Williams *et al.,* 1990), microsatellite-primed PCR (Weising *et al.*, 1995; Gupta *et al.,* 1994), and oligonucleotide fingerprinting (Epplen, 1992); regarding sensitivity, speed, level of variability detected, and non-requirement of prior sequence information (Richardson *et al.,* 1995).

Amplified fragment length polymorphism (AFLP)

The AFLP technique, developed by Zabeau and Vos (1993), is a combination of PCR and RFLP, where PCR amplification products are restriction digested followed by gel electrophoretion to obtain fingerprinting patterns that are extremely useful in detection of polymorphism between even closely related genotypes (Saiki *et al.*, 1988; Ehrlich *et al.*, 1991).

However, AFLP procedure is tedious and complex as it involves selective amplification of restriction-digested DNA using specific primers to create multiple bands. Double stranded adapters are ligated to DNA digested with one or two restriction enzymes such that the adjacent restriction site and sequence of the adapters serve as primer-binding sites for subsequent PCR amplification resulting in 50 to 100 bands per individual sample

upon analysis through denaturing polyacrylamide gel electrophoresis. Thus, polymorphism detected in AFLP results not only from diffrences in primer-binding sites but also from altered restriction sites. This offers, a great opportunity for capturing polymorphism arising from mutations such as insertions or deletions between two restriction sites, or inversions changing the relative position of one of the two restriction sites, ultimately leading to either abolition or creation of a restriction site

Like RAPD analysis, AFLP assay also does not require prior-knowledge of DNA sequence information, but helps detecting more number of loci upto 10-fold greater than that of RAPD analysis. Having the greater capacity to rapidly screen thousands of independent genetic loci, AFLP markers typically follow Mendelian inheritance and, therefore, has a high potential for typing, identification, and mapping of genetic loci; all making the AFLP markers more popular in developing saturated genetic linkage maps in various crop plants.

2.5.2.2 Repetitive sequence-derived PCR-based marker systems

It is known that, on an average, about 30-90% of the genome of virtually all eukaryotic species has one or the other kind of repetitive DNA. Repetitive DNA is highly polymorphic in nature offering an opportunity to exploit it as gold-mine for devising several genetic marker systems. Microsatellite, the term coined by Litt and Lutty (1989), and minisatellite, the term coined by Jeffrey (1985b), are the two most common and important forms of repetitive DNA. The main difference between the microsatellite and ministellite is in terms of length, in base pairs, of tandem repeats that constitute these satellites. While simple sequence repeat units of 1-6 bp length is found in microsatellite, monomer repeat length of about 11-60 bp characterizes the repetitive DNA as minisatellite.

Micro- and mini satellites form ideal marker systems because of advantages of locus-specific amplification, co-dominant nature, abundant availability, easy-to-use PCR-based marker system, and obedience to Mendelian inheritance.

The following are the various types of marker systems developed based on microsatellite and minisatellite sequences discovered in the genome.

Sequence tagged microsatellites sites (STMS)

Since DNA polymorphism is detected using specific primers designed from sequence data of a specific locus, primers complementary to the regions flanking the simple sequence repeat (SSR) locus yield high level of polymorphism. Though di-, tri- and tetra-nucleotide repeat units found in microsatellites are more popular for STMS analysis, as they give a clear banding pattern (Hearne *et al.,* 1992), di-nucleotide SSRs are generally abundant in the genome are useful for diversity analysis (Rafalski and Tingey, 1993).

Directed amplification of minisatellite region DNA (DAMD)

Minisatellites are used as primer-binding sites for DNA amplification. Introduced by Heath *et al*. (1993), DAMD has potential to be useful for cultivar identification and species differentiation (Somers *et al.,* 1996).

Inter simple sequence repeat markers (ISSR)

This technique, developed by Zietkiewicz *et al.* (1994), involves microsatellites anchored at the 3' end of primers used in PCR-amplification of genomic DNA. Being mostly a dominant marker, ISSR with various combinations of di-, tri-, tetra- and penta-nucleotides provides an opportunity to design many primers. many primers can be synthesized with few anchors for various combinations of di-, tri-, tetra- and penta-nucleotides. For instance, 64 combinations for tri-nucleotide repeats (because, $4^3 = 64$) and 256 for tetra-nucleotide repeats (4^4=256).

2.5.3 Third generation DNA markers (nucleotide sequence-based)

In recent years, there has been an emphasis on the development of newer and more efficient high throughput molecular marker systems involving inexpensive ans non-gel-based assays with high throughput detection systems, *viz*., SNPs (Single Nucleotide Polymorphism) (Gupta *et al.,* 2001) and Microarrays (Linman *et al.,* 2009).

Invasive cleavage, primer extension, oligonucleotide ligation assays (OLA), and allele-specific oligonucleotides (ASO), etc., can be deployed for high-throughput analysis of polymorphism due to single base differences (Maurya, 2014) called single nucleotide polymorphism (SNP). The main advantage of SNPs lies in their amenability for an automated highthroughput analysis at moderate cost (Chen and Sullivan, 2003).

In the year 2012, Hoshino *et al.* (2012) conduced a study to compare trend in the application of different DNA markers in terms of number of research articles published betwen 2007 and 2012. They used 'Web of Science' facility (Thomson Reuters, http://webofknowledge.com) to collect the information, summarized as follows:

Area of Marker Application	**Type of Molecular Marker**				
	Microsatellite/ SSR	**SNP**	**RAPD**	**AFLP**	**RFLP**
Biochemistry	1178	185	83	96	38
Molecular Biology	989	33	18	78	11
Evolutionary Biology	989	23	17	49	16
Ecology	1134	493	124	131	32
Genetics/Heredity/Bio-diversity/Conservation	405	4	3	8	3
Total	**4690**	**1.269**	**925**	**668**	**531**

It was found that there was an increasing trend that the microsatellite (SSR) had been the most used molecular marker to assess genetic diversity.

2.6 Simple sequence repeat (SSR) or microsatellite markers

Simple sequence repeats (SSRs), also known as microsatellites, are molecular markers that have become available to researchers for routine use (Quarrie *et al.*, 2005). Condit and Hubbel (1991) first reported SSRs in plants, suggesting their abundance in plant systems. The term microsatellite (simple sequence repeat), coined by Litt and Lutty (1989), refers to tandemly repeated DNA sequences of short motifs (e.g. poly CA, poly CT, poly AT and other repeated sequences of 1-6 bp). The first widespread markers to avail full advantage of PCR technology was microsatellites (Litt and Luty, 1989; Tautz, 1989; Weber and May, 1989). Initially Akkaya *et al.* (1992) reported length polymorphisms of SSRs in soybean, which opened a new a source of PCR-based molecular markers for other plant genomes also. Later, only with the advent of microsatellite or simple sequence repeat (SSR) markers, the use of molecular markers gained momentum during the last 20 years and thus SSRs are recognized as highly informative and locus-specific markers (Kumar, 2012).

The genomes of higher organisms contain three types of multiple copies of simple repetitive DNA sequences, satellite DNAs, *viz*., minisatellites and microsatellites, arranged in arrays of vastly differing sizes (Armour *et al.*, 1999; Hancock, 1999). Microsatellites; variously known as short tandem repeats (STR), simple sequence repeats (SSRs), or simple sequence length polymorphism (SSLPs); are the tandem repeats occuring in the form of

iterations of repeat units ranging from a single base pair to thousands of base pairs (Litt and Luty, 1989). Some researchers (Armour *et al.*, 1999) define microsatellites as 2-8 bp repeats, others (Goldstein and Pollock, 1997) as 1-6 or even 1-5 bp repeats (Schlotterer, 1998). Mono-, di-, tri- and tetra-nucleotide repeats are the main types of microsatellite but repeats of five (penta-) or six (hexa-) nucleotides are usually classified as microsatellites as well (Maurya, 2014). In 1960s, several researchers observed a distinct buoyant density during density gradient centrifugation which appeared as a 'satellite peak' and it was attribute to a fraction of sheared DNA, subsequently identified as large centromeric tandem repeats which was later termed as satellite DNA. Microsatellites were first identified in humans in 1981 by sequence analysis of alleles at the β globin locus (Miesfeld *et al.*, 1981; Spritz, 1981) and subsequently found to be naturally occurring and ubiquitous in prokaryotic and eukaryotic genomes (Tautz and Renz, 1984; Jeffreys *et al.*, 1985a; Tautz, 1989; Thoren *et al.*, 1995; Toth *et al.*, 2000). The repeats of longer units form minisatellites or, in the extreme case, satellite DNA (Ellegren and Hans, 2004). Though a typical satellite DNA is a centromeric sequence with a 100bp repeat, isolation and sequencing showed that repeat motifs have variable length ranging from just a single base to as many as thousands of bases (Pardue and Gall, 1970). Subsequently, satellite isolated in mammals was called minisatellite as its repeat motifs was 10-30 bp long (Jeffreys *et al.*, 1985a). Further, satellites with varying repeat motif length were observed in various sources by various resesarchers. Finally, satellites called microsatellites with even shorter repeat motifs were isolated after the discovery of di-nucleotide repeats of poly-(CA) and poly-(GT) in genomes of diverse eukaryotes by Hamada and his colleagues in 1982. Later, Weber and May (1989) demonstrated that using two PCR- primers designed to bind at genomic regions flanking repetitive regions, SSR polymorphisms (SSRPs) could be easily detected through PCR-amplification; and, this possibility has led to the development of SSRs and their assignment to specific chromosomes in various mammalian species.

In plants, hybridization of oligonucleotide probes of poly-(GT) and poly-(AG) with genomic libraries (using phage) first time revealed the presence of SSRs (Maurya, 2014). Analysis of published DNA sequences in diverse plant genomes showed the abundance of SSRs (Morgante and Olivieri, 1993). It is evident that non-coding regions are richer in SSRs compared to exonic region (Hancock, 1995) and different type of SSRs are preferred

by different taxa (Beckmann and Weber, 1992; Lagercrantz *et al.,* 1993; Tautz and Schlotterer, 1994).

Microsatellites are highly polymorphic, abundant and fairly evenly distributed throughout the euchromatic part of the genomes. These properties have made microsatellites one of the most popular genetic markers for mapping, paternity testing and population genetics (Goldstein and Schlotterer, 1999). SSRs are non-randomly distributed within expressed sequence tags (ESTs), UTR regions, introns and coding regions. SSR variations within these regions can cause frame-shift, alteration in gene expression, inactivation of gene, change of function and eventually resulting in phenotypic changes that can cause deleterious effects in organisms (Li *et al.,* 2004). The SSRs also play a significant functional role essential for gene transcription, translation, chromatin organization, recombination, DNA replication, cell cycle, etc., as depicted in Figure 2.3.

2.6.1 Evolutionary mechanism of origin/genesis of microsatellites

The genesis/origin of SSRs is an evolutionarily dynamic process and has proven to be exceedingly complex phenomenon (Ellegren and Hans, 2004; Pearson *et al.,* 2005). Most of the researchers believed that the following were the pre-requisites for genesis of microsatellites: (i) proto-microsatellite (ii) retro-transposon and (iii) insertion/deletion of 2-4 nucleotides. These mechanisms are described below.

Proto-microsatellite

During the process of DNA replication, whenever DNA polymerase encounters repeat sequence on the template DNA it experiences replication-slippage leading to extension or constriction of the repeat sequence by adding or deleting few repeat-units. However, for the occurrence of replication-slippage, there must exist a minimum number of repeats called proto-microsatellite (Rose and Falush, 1998). In mammals, for instance, the study has shown that $(ATGT)_4$ in African monkeys and $(ATGT)_5$ in humans were evolved from a pro-microsatellite $(ATGT)_2$ which in turn was created due to GA mutations at the η-globin locus because of which sequence “ATGTGTGT” was changed to “ATGTATGT”. Proto-microsatellite $(ATGT)_2$ was evolved in primates which is the common ancestor for other species such as gorillas, chimpanzees and humans (Messier *et al.,* 1996).

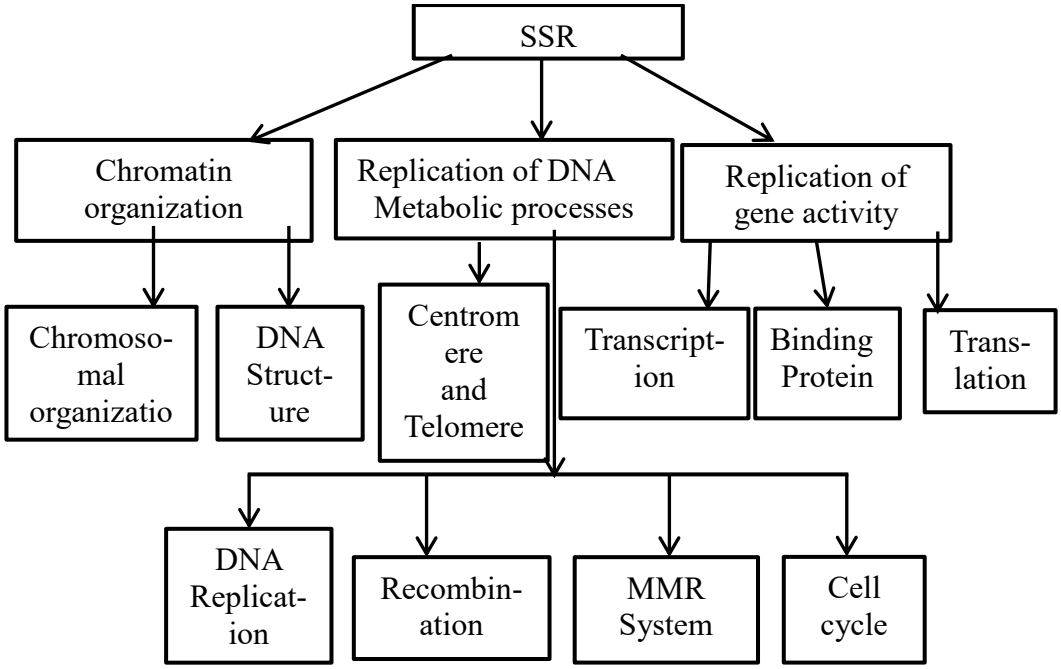


Figure 2.3 Putative functions/effects of SSR (Reproduced from Li *et al*., 2002).

Retro-transposon

Whenever any RNA molecule with repetitive DNA element is reverse-transcribed and are inserted back into chromosomes, they become retro-transposons. SSR generation was found to be accompanied by retro-transposition events by analysis of a portion-sequenced of human and rice genomes (Nadir *et al.,* 1996; Temnykh *et al.,* 2001). Primate genome has largely dispersed *Alu* SINES (Short Interspersed Nuclear Elements) family that has A-rich regions at the 3' terminal and within the sequence, which probably contributed to the genesis of microsatellites (Arcot *et al.*, 1995).

Insertion/deletion of 2-4 nucleotides

Zhu *et al.* (2000), based on their study on mutated genes in human genome, observed that most of the 2 to 5 nucleotide new repeats are due to random sequences giving rise to new microsatellites, either by way of insertion or deletion, and are not extensions of pre-existing repeats.

2.6.2 Evolution of SSRs in wheat

Tang *et al*. (2009) studied F_1 hybrids and newly synthesized allopolyploids derived from *Triticum aestivum* Chinese Spring × *Secale cereale* Jinzhou-heimai. They deployed 163 wheat simple sequence repeat (SSR) markers to investigate the variation of wheat microsatellites after allopolyploidization and variation of the PCR products of 29 SSR markers was observed. The lengths of the products amplified from amphiploids were different those of parents and F_1 hybrids. Sequencing indicated that the length variation of the 14 microsatellites stemmed mainly from variation in the number of repeat units. The alteration in repeat lengths occurred in both perfect and compound repeats. In some compound SSR loci, one motif was observed to expand whereas another to contract. Almost all the microsatellite evolution observed in this study could be explained by the slipped-strand mis-pairing model.

Li *et al.* (2002) illustrated the origin of microsatellite from random sequences as given in the following figure.

The results of this study seem to indicate that stress caused by allopolyploidization might be one of the factors that induce microsatellite evolution. Normally, microsatellites rapidly evolve in diploids. This provides direct experimental evidence on how microsatellites evolve following allopolyploidization.

2.6.3 Molecular mechanism of SSR length variation

The most accepted mechanism of microsatellite length variation is mutational mechanism including errors during recombination (Levinson and Gutman, 1987), unequal crossing-over (Harding *et al.*, 1992), and polymerase slippage (Wolff *et al.*, 1991; Stephan and Kim, 1998) during DNA replication or repair. In the event of unequal crossing-over, there takes place loss of many repeats on one non-sister chromatid while gain of many repeats occurs on the other non-sister chromatid.

Replication slippage

Length changes in microsatellite DNA generally occur due to replication slippage (Figure 2.4) during DNA replication and/or repair processes.

The DNA polymerase experiences replication-slippage when one DNA strand temporarily dissociates from the other and rapidly rebinds to it at a different position. As a result of this, base-pairing errors occur either on parent strand or new strand. If the base-pairing errors occur on new strand, there will be an increase (i.e., additions) in the number of repeats, whereas if it occurs on parent strand, there will be decreased (i.e. deletions) number of repeats (Goldstein and Schlottrer, 1999; Ellegren and Hans, 2004). Inverse relationship between rate of slippage and amount of changes in the number of repeats has been demonstrated by various studies. The main reason for this is that alteration in DNA polymerase, or its cofactors, can result in increased slippage rates or there is no effective repair system for DNA loops, ultimately resulting in destabilized microsatellites.

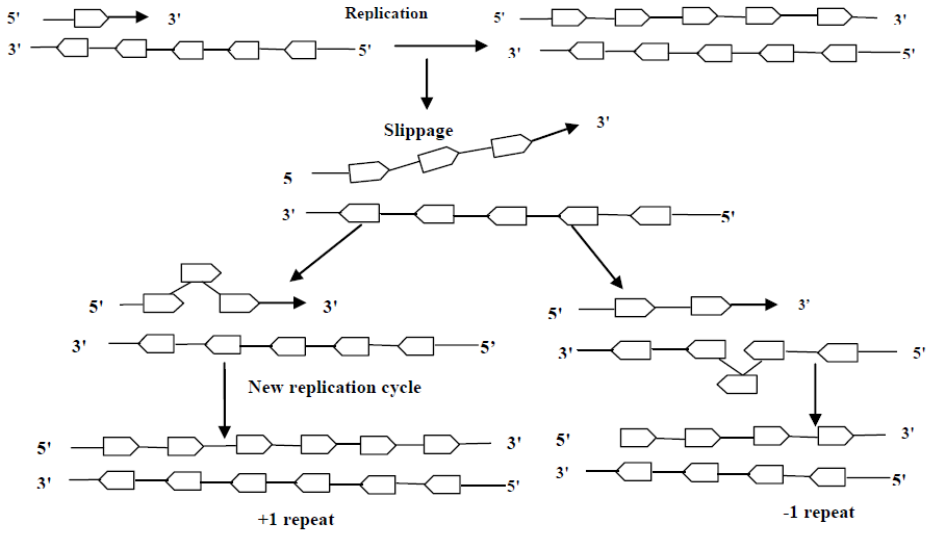


Figure 2.4 Slippage during DNA replication. Assuming that in the original DNA molecule there were 5 repeats of the motif, symbolized by a box, slippage leads to the formation of new alleles with 6 and 4 repeats, depending on the strand containing the polymerase error (Reproduced from Goldstein and Schlottrer, 1999).

Recombination

Recombination could potentially change the SSR length by unequal crossing over (Figure 2.5) or by gene conversion (Brohele and Ellegren, 1999; Jakupciak and Wells, 2000). Unequal exchanges in combination with random genetic drift and selection can have a strong effect on the accumulation of tandem-repetitive sequences in the genome (Charlesworth *et al.*, 1994). Non-reciprocal recombination (gene conversion) play significant role in destabilization of tandem repeats for both micro- and mini-satellite (Jakupciak and Wells, 2000; Richard and Paques, 2000).

Interaction of replication slippage and recombination

A very strong interaction was found between mean repeat length and SSR locus distance from the centromere, impacting the number of alleles and variation in repeat size at SSR loci, and microsatellite diversity in wild emmer wheat (Li *et al.*, 2003). The interaction of slippage and recombination, which may happen in heteroduplex DNA tracts, could also affect SSR stability (Li *et al.*, 2002).

2.6.4 Classification of SSRs

Depending upon the arrangement of nucleotides within the repeat motifs, Weber (1990) used the terms, *viz*., perfect, imperfect and compound to classify microsatellites, whereas Wang *et al.* (2009) coined the terms *viz*., simple perfect, simple imperfect, compound perfect and compound imperfect (Table 2.3). The dinucleotide repeats are most common in many species but are much less frequent in coding region than in non-coding regions (Li *et al.*, 2002; Wang *et al.*, 1994). In many species, exons have more triplet SSRs than other repeats (Morgante *et al.*, 2002; Li *et al.*, 2004). In plants, the most frequent triplet is AAG (Li *et al.*, 2004), although in cereals, most common triplet is CCG (Cordeiro *et al.*, 2001; Varshney *et al.*, 2002; Thiel *et al.*, 2003). In general, frequency of microsatellites is inversely related to the genome size in plants, but the percentage of repetitive DNA appeared to remain constant in coding regions (Morgante *et al.*, 2002). The genomic distribution, evolutionary dynamics, biological functions, and practical utility have been the objective of many studies, as summarized in several review articles (Tautz and Schlotterer, 1994;Jarne and Lagoda, 1996; Schlotterer, 1998; Chambers and

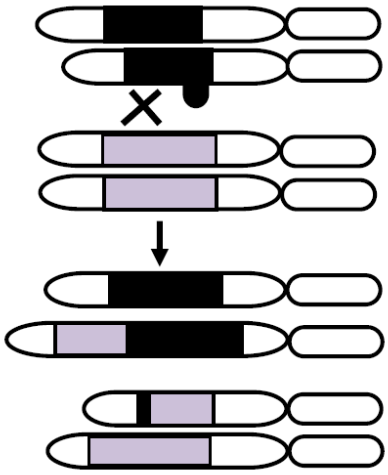


Figure 2.5 Unequal crossing-over between homologous chromosomes. Black and blue regions correspond to microsatellite repeat sequences (Reproduced from Oliveira *et al*., 2006).

Table 2.3 Classificationof SSR markers (Adopted and mdified from Kalia *et al*., 2011).

Types	Example/Notation
(a) Based on the number of nucleotides per repeat	
Mononucleotide	$(A)_n$
Dinucleotide	$(CA)_n$
Trinucleotide	$(CGT)_n$
Tetranucleotide	$(CAGA)_n$
Pentanucleotide	$(AAATT)_n$
Hexanucleotide	$(CTTTAA)_n$
(b) Based on the arrangement of nucleotides in the repeat motifs	
Perfect or simple perfect	$(CA)_n$
Imperfect or simple imperfect	$(AAC)_n$ ACT $(AAC)_{n+1}$
Compound or compound perfect	$(CA)_n$ $(GA)_n$
Interrupted or compound imperfect	$(CCA)_n$ TT $(CGA)_{n+1}$
(c) Based on the location of SSRs in genome	
Nuclear:	nuSSR
Genic	EST-SSR
Genomic	g-SSR
Chloroplastic	cpSSR
Mitochondrial	mtSSR

MacAvoy, 2000; Li *et al.,* 2002; Dieringer and Schlotterer, 2003; Ellegren and Hans, 2004; Oliveira *et al.,* 2006; Subirana and Messeguer, 2008; Sun *et al.,* 2009).

In a perfect microsatellite, the repeat sequence is not interrupted by any other base (e.g., TATATATATATATATA). Whereas, in an imperfect microsatellite, there is one or two bases, between the repeated motifs, which do not match the motif sequence (e.g., TATATATACTATATA). In case of an interrupted microsatellite, there is a small sequence different form the motif sequence (e.g., TATATATATATACGGTGTGTGTGT). A compound microsatellite sequence is composed of two or more distinctive sequence repeats adjacent to each other (e.g., TATATATATAGTGTGTGTGT). Microsatellites (SSRs) can also be classified based based on number of nucleotides per repeat and the location of SSR in genome (Table 2.3

2.6.5 Frequency and distribution of SSRs

Even though SSRs constitute a large fraction of non-coding DNA, several reports have shown that SSRs are also located in transcribed regions of genomes, including protein coding genes and expressed sequence tags (ESTs); repeat number of SSRs in coding regions are comparatively low (Morgante *et al.,* 2002; Li *et al.,* 2004). For instance, in cereals (maize, wheat, barley, Sorghum, and rice), only 1.5-7.5% SSRs are available in ESTs (Kantety *et al.,* 2002; Thiel *et al.,* 2003).

2.6.6 Advantage and disadvantages of SSR marker system

2.6.6.1 Advantages

SSR markers have many advantages over the other marker systems and a few of them are described below.

High reproducibility: it is the most important in genetic analysis. Though SSR analysis is simpler and requires only a small amount of sample DNA with no demand for quality; and, as there is no involvement of restriction-digestion with enzymes; the reproducibility of the SSR profile is as robust as that of RFLPs. Therefore, SSRs are potential to be used for analysis of specimens that are mummified, dry, in fossilized form in the wild, or even contaminated (Manen *et al.,* 2003; Boder *et al.,* 2006).

Hyper-variability

Very high allelic variations can be detected with SSR markers even among very closely related genotypes. A literature survey by Powell *et al.* (1996) showed that diversity indices varied from 0.29 to 0.95 with number of alleles per marker locus in the range of 1 to 37 in major crop species. For instance, when a panel of 61 soybean genotypes were assessed for their diversity using SSRs and RFLPs, level of genetic diversity detected by SSRs was almost two times higher compared to that revealed by RFLPs (Morgante *et al*., 1994).

In a comparative study by Powell *et al.* (1996) on comparison among AFLP, RFLP, SSR, and RAPD marker systems with respect to their suitability for assessing germplasm diversity, SSRs revealed the highest expected heterozygosity**,** while the highest effective multiplex ratio was observed with AFLP.

Co-dominancy

The third advantage is the co-dominat nature of SSR polymorphisms. Although SSR profiles encounter homoplasious bands, different sized amplicons obtained from same set of primers essentially reflect the amplification of different alleles of the marker locus coming from the homologous position. Multiple amplicons produced by RAPD and AFLP markers become allelic bands only if confirmed after converting these markers into STS markers. The co-dominant nature of microsatellite markers (SSRs) makes them an ideal marker system for analyzing parentage in hybrids and segregating F_2 population (Scott *et al.,* 2000; Slavov *et al.,* 2005).

Abundance and distribution

The SSRs are shown to be highly abundant and distributed throughout the genomes (Wang *et al.,* 1994; Toth *et al.,* 2000; Varshney *et al.,* 2005). One of the major difficulties with RAPD or AFLP markers is that they are concentrated in specific regions of the chromosomes as opposed to SSRs which are evenly distribute throughout the chromosome (Vuylsteke *et al.,* 1999; Kwon *et al.,* 2006). *In silico* investigation by Varshney *et al.* (2002) for frequency and distribution of microsatellites in ESTs of some cereal species like barley, maize, oats, rice, rye, and wheat showed that the frequencies of SSRs were: 1/7.5 Kb in

barley, 1/7.5 Kb in maize, 1/6.2 Kb in wheat, 1/5.5 kb each in rye and sorghum, and 1/3.9 Kb in rice. Gupta *et al*. (2003) recorded the average density of SSRs per 9.2 kb of EST sequence of wheat. The genomic distribution of SSRs revealed that 27% were present in the non-genic regions, whereas 73% were also present in the putative genic regions with 26% in 5'UTRs, 25% in introns, 16% in 3' UTRs and 6% in exons (Sharma and Chauhan, 2011).

Genomic location

Microsatellite (SSR) markers are mostly found in the non-repetitive DNA (Varshney *et al.,* 2005; Morgante *et al.,* 2002; Andersen and Liberstedt, 2003) comprising either the transcribed region (genic SSRs) or the non- transcribed region (genomic SSRs). Using the sequence information of ESTs or cDNAs, genic SSR marker system can be developed and used for QTL analysis (Buerstmayr *et al.,* 2002; Breseghello and Sorrells, 2006; Zeng *et al.,* 2009), characterizing gene-function (Ronning *et al.,* 2003), and for gene tagging through association studies (Szalma *et al.,* 2005; Shin *et al.,* 2006; Crossa *et al.,* 2007).

2.6.6.2 Disadvantages

Microsatellite (SSR markers) have the following disadvantages: (i) the requirement of a known sequence to be amplified (Weising *et al.,* 2005); (ii) Developing new microsatellites are expensive and time consuming (Coates and Byrne, 2005); and (iii) the phenomenon of null-alleles, which are non-amplifying alleles, appears frequently. Null-alleles leave no bands in case of homozygosity, but in heterozygosity it gives one band. This will interfere and complicate reading of data, since it will be registered as a homozygote individual when actually being a heterozygote (Idrees and Irshad, 2014). To reduce error due to null alleles, population studies should contain many diverse SSR primers so that different multiple microsatellite loci are investigated (Weising *et al.,* 2005).

2.7 Application of SSR markers for assessment of genetic and evolutionary relationships

Microsatellites have become markers of choice for an array of applications in plants due to their hypervariable nature and extensive genome coverage. There are many applications of microsatellites in plants, a few can be categorized as: (1) genome mapping; (2) cultivar

identification and marker-assisted selection; (3) genetic diversity analysis and phylogenetic relationship; (4) population and evolutionary studies, etc., (Figure 2.6).

2.7.1 Genetic diversity and phylogenetic relationships

Genetic diversity refers to any variation in nucleotides, genes, chromosomes or whole genome of an organism. Genetic diversity can be assessed at different levels within a species or among species. Phylogenetic relationships reflect the relatedness of a group of species based on a calculated genetic distance (sequence conservation or diversification in their evolutionary history). SSRs often constitute a powerful marker system for revealing interspecific, or intra-specific, phylogenetic relationship. For instance, the genetic diversity and phylogenetic relationship from germplasm collection such as a temperate bamboo collection (Barkley *et al.,* 2005), a citrus variety collection (Barkley *et al.,* 2006) and a cultivated and wild peanut collection (Barkley *et al.,* 2007; Cuc *et al.,* 2008) have been assessed with SSR markers.

2.7.2 Population and evolutionary studies

Studies of plant evolution were traditionally based on taxonomic, phenotypic and cytogenetic data (such as morphological and karyotype). Microsatellite markers can be used to determine the population structure within and among the natural population and /or identify the potential progenitors. The development of organelle-specific SSR markers (i.e. cpSSR and mtSSR) had a profound impact on the determination of structure and variation within a natural population as well as phylogenetic relationships. The uni-parental mode of inheritance, conserved gene order and lack of heteroplasmy and recombination of organelle genomes make them an attractive tool for evolutionary studies, mainly patterns of migration, population histories and levels of differentiation (Provan *et al.,* 2001). However, ESTs are also being used for such analysis because in such studies, one actually looks at the evolution of functional genes (Joshi *et al.,* 1979).

Evolution of genetic diversity and phylogenetic relationships has resulted in identification of some mis-classified accessions that were re-classified. Genetic diversity assessment and phylogenetic relationship construction will provide important information for choosing parental lines for breeding programs, classification of plant germplasm accessions and further curation and acquisition of new plant germplasm accessions (Wang *et al.,* 2009).

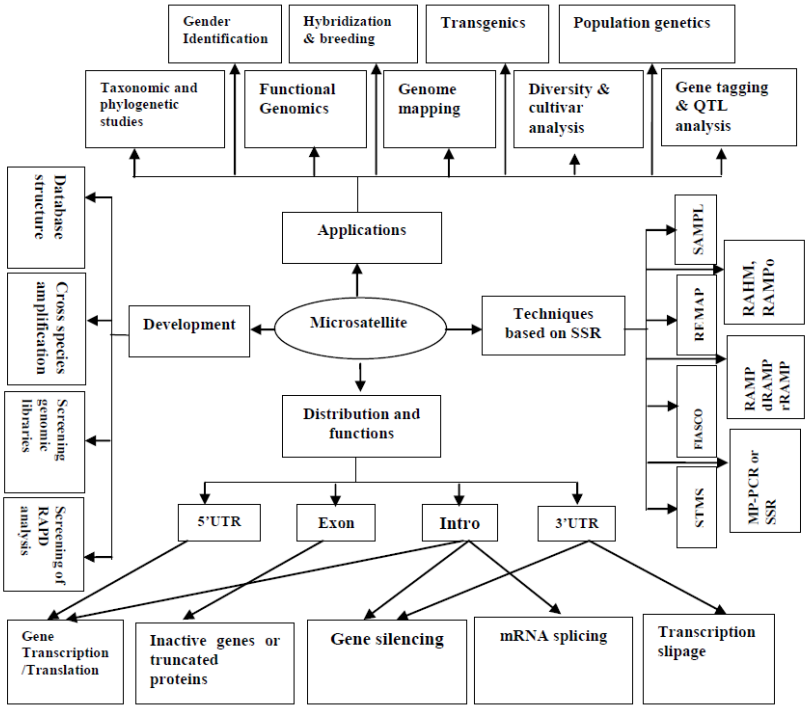


Figure 2.6 Development, distribution, functions and applications of microsatellites (Reproduced from Kalia *et al*., 2011).

2.7.3 Marker-trait association studies

Association mapping, which refers to significant association of a molecular marker with a phenotypic trait, is especially useful for implementing marker assisted selection for quantitative traits in plant breeding programs (Breseghello and Sorrells, 2006).

Once mapped, microsatellite markers could be employed in tagging several individual traits that are particularly important in a breeding programs. Microsatellite (SSR) markers are successfully deployed for association studies in important crop species. Association mapping of SSRs in cotton (*Gossypium hirsutum* L.) germplasm for *Verticillium* wilt resistance was conducted by Zhao *et al*. (2014). Xiao *et al.* (2013) studied the associated markers with drought tolerance at vegetative stage and examined the pattern of linkage disequilibrium (LD) in a diverse and stress adaptated rice panel containing 184 rice germplasm accessions with 141 polymorphic SSR markers that were nearly evenly distributed at 3 mb bin on the 12 rice chromosomes. Hu *et al.* (2014) identified yield-enhancing QTL and conducted association mapping with 85 SSR markers in wild soybeans accessions (*Glycine soja*). Laido *et al.* (2014) utilized SSR and 970 DArT markers for linkage disequilibrium and and genome-wide association mapping in *Triticum turgidum* ssp, a tetraploid wheat.

2.8 Application of molecular markers in wheat genetic diversity

Several types of molecular markers such as RAPD, AFLP, ISSR, SSR, and SNP etc., have been deployed by various researchers for studying genetic diversity, population structure and phylogeny in wheat and its close relatives. The appropriateness of the type of marker to be used in a particular study is to be judged after considering purpose of the study to be undertaken and characteristics of marker system such as easy availability, frequent occurrence in genome, highly polymorphic nature, high reproducibility, Mendelian inheritance, easy and fast assay, selective neutral behavior, and not having pleiotropic and epistatic effect (Weising *et al.*, 1995).

The detailed survey of relevant literatures on molecular marker applications in wheat is described in the following subsections.

2.8.1 Hybridization-based DNA markers

Mori *et al.* (1995) deployed RFLP markers for phylogenetic analysis of *T. araraticum* and *T. dicoccoides*, and found that *T. araraticum* originated in the recent past compared to *T. dicoccoides* and that they had a diplyletic origin. In another study based on RFLP analysis, Mori *et al.* (1997) suggested that the earliest domesticated tetraploid wheat was *T. dicoccum*, while Li *et al*, (1999) concluded that long term introgression between cultivated emmer and its wild progenitor gradually lead to the domestication of emmer wheat.

Since wheat has high proportion of repetitive DNA, there are inherent problems associated with RFLP procedure such as low levels of polymorphism (Ranjekar *et al.*, 1976; Flavell and Smith, 1976). Therefore, polymerase chain reaction (PCR) based markers such as RAPD emerged as an alternative marker system.

2.8.2 PCR-based DNA markers

2.8.2.1 Randomly amplified polymorphic DNA (RAPD)

Germplasm collections of several crop species, including wheat, have been effectively managed for monitoring genetic erosion, removal of duplicates, and estimation of diversity, using RAPD markers (Virk *et al.*, 1995). Since RAPDs are amenable to automation, efficient and rapid, polymorphism in wheat can be effectively detected by using them (Vierling and Nguyen, 1992; Joshi and Nguyen, 1993; Sun *et al.*, 1998). Using Langdon durum wheat disomic substitution lines, specific chromosomes and chromosome-arms have been assigned 11 RAPD markers by Wang *et al.* (1995).

2.8.2.2 AP-PCR Markers

Arbitarily primed-PCR was used by Pujar *et al.* (1999) to assess diversity of Indian tetraploid wheat genotypes.

2.8.2.3 Inter-simple sequence repeats (ISSR)

ISSRs have been used for detection of polymorphism (Nagaoka and Ogihara, 1997) and in genetic mapping of wheat (Kojima *et al*, 1998). They have been used to study diversity among the Indian tetraploid wheat genotypes (Pujar *et al.*, 2002). ISSRs have also been

used to identify markers associated with seed size, yellow berry tolerance (Ammiraju, *et al.*, 2001), and grain protein content (Dholakia *et al.*, 2001) in wheat.

2.8.2.4 AFLP markers

Maccaferri *et al.*, (2007) used 14 AFLP primer combinations, alongwith 38 simple sequence repeat (SSR) markers, to assess genetic diversity in 58 durum wheat accessions. Among them, 43 were composed of old and modern Italian durum cultivars, while 15 were exotic (5 French, 4 from US, 3 Turisian and 3 cultivars from CIMMYT, Mexico). The SSR and AFLP similarity matrices were highly correlated (r=0.74) in this study, suggesting good genome coverage and accurate representation of genetic diversity by both the marker systems.

2.8.2.5 Microsatellite (SSR)

The use of SSRs is a common way to examine and detect polymorphisms in wheat. Simple sequence repeats are very informative due to their co-dominant nature. Additionally, they contain more allelic diversity per marker (Chao, 2009), and they have higher mutation rates than other marker types (Thuillet *et al.*, 2002). SSRs continue to be the marker of choice in wheat, either alone or in combination with other types of markers.

SSR markers were described in hexaploid wheat during 1995 to 1997 (Devos *et al.*, 1995; Roder *et al.*, 1995; Bryan *et al.*, 1997). By analyzing divergence of *T. dicoccoides* accessions highly resistant to yellow rust using SSRs, Fahima *et al.* (1998) showed that diversity in the wild species can be detected using a relatively small number of microsatellite markers. Li *et al.* (2000) amplified microsatellite loci in a panel of Istraelian *T. dicoccoides* in their genetic diversity study and observed that microsatellite divergence was influenced by natural selection under aridity stress. In similar studies, wheat microsatellite markers were used by Dograr *et al.* (2000) to study genetic relatedness among Turkish durum wheat genotypes, and by Hammer *et al.* (2002) to distinguish between diploid species.

They are present in both coding and non-coding regions and are usually characterized by a high degree of length variation (Gupta and Varshney 2000; Zane *et al.*, 2002). The variability of microsatellite sequences in genome is not based on point

mutations but on the variation in the number of the simple sequence repeat motifs (Kumar, 2012). Such variation occurs approximately 10 times more frequently, and originates through processes such as slippage during replication or unequal crossing over (Hancock, 1999).

Akram *et al.* (2012) analysed 40 *Triticum* germplasm using 34 SSR markers and the analysis grouped them into seven units. Their results showed that the *Triticum durum* lines T_2 and T_3 were placed in group A, whereas their counterpart T_1 was quite distinct. Among T_4 to T_9 that were clustered in Group B, T_4 to T_7 were of *Triticum sphaerococcum*, whereas the other two were of *Triticum aestivum* type. *T. aestivum* cultivar C-248 also appeared to be distinct and could not be grouped with any other cultivar. It was suggested that C-248 can be used for enhancing diversity in *T. aestivum* and T_1 in *T. durum.*

2.9 Recent studies on diversity, population structure and phylogeny in wheat

2.9.1 Morphological and phenological traits

Market-driven intensive breeding for higher crop performance and increased crop uniformity often leads to unwanted loss of genetic diversity called genetic erosion (Baranger *et al.,* 2004). In addition, self-pollinating crops such as wheat end-up with increased homozygosity, leading to loss of genetic variation. Efforts are being made worldwide to assay phenotypic diversity in local landraces to mine useful traits as well as to assess exploitable diversity, both in wheat and other crops: for instance, Uttaranchal wheat landraces (Routray *et al.,* 2007), rice landraces (Dikshit *et al.,* 2014); Morocco durum wheat landraces (Aghaee *et al.,* 2010; Zarkti *et al.,* 2012); Indian mustard genotypes (Yadav *et al.,* 2013); Iranian bread wheat genotypes (Sabaghnia *et al.,* 2014); Indian wheat varieties (Malik *et al.,* 2014); Ethiopian durum wheat landraces (Mengistu *et al.,* 2015), Indian wheat cultivars (Mishra *et al.,* 2015), and Sorghum genotypes (Sinha and Kumaravadivel, 2016).

Descriptive statistics, Carl Pearson's coefficient and phenotypic diversity index analyses are effective methods of understanding degree of polymorphism and distribution of studied traits in the evaluated set of genotypes (Sabaghnia *et al.,* 2014). Principal component analysis (PCA) and cluster analysis are widely used tools among multivariate techniques. They help for morphological characterization of genotypes and their grouping

on genetic similarity basis (Peeters and Martinelli, 1989; Mohammadi and Prasanna, 2003; Sabaghnia *et al*., 2014).

Wheat agro-morphological traits critically contributing to genetic variability among genotypes, and their grouping according to genetic similarity, can be comprehended (Rachovska *et al*., 2002) with GT Biplot that is extensively used to display graphical grouping of accessions (Aghaee *et al*., 2010; Malik *et al*., 2014). Since core collection is a small subset of accessions that represents the whole variability-spectrum of a species, detecting duplications within germplasm collections plays a very critical role; as avoiding of unnecessary removal is as important as that of unnecessary duplication; which can be ensured only after quantifying the extent of genetic diversity among accessions (Zarkti *et al*., 2012).

Little attempts have been made in past to identify genetic potential of Indian wheat accessions. The earlier studies were carried out with less number of genotypes, which is inadequate to depict an approximate picture of wheat genetic resources of India. The following are the reports of recent studies on morphology-based diversity studies in wheat, and they are summarized in Table 2.4.

Zahariev *et al*. (2003) analysed a collection of 52 *Aegilops* accessions belonging to three tetraploids and self-pollinated species (*Ae*. *cylindrica* Host, *Ae*. *geniculata* Roth and *Ae*. *Neglecta* Req. ex Bertol.), sampled from different eco-geographical regions of Bulgaria. They evaluated them during three successive years based on morphological descriptors and agronomical characters. They analysed morphological diversity using a set of thirteen quantitative traits and a significant difference between populations was noted for most morphological traits, the highest variation being observed in the three species for thousand kernel weight, grain weight per spike, plant height and earliness. They did not observe any strong relation between the morphology of the studied accessions and their eco-geographical region of origin, which was obviously due to marked differences between populations from the same region. The most important characters contributing to the eco-geographical differentiation were those related to vegetative growth in *Ae. Cylindrica*, and grain production and earliness in *Ae*. *geniculata*. Considerable variability was found among and within species in response to bioric and abiotic stresses. Promising accessions

Table 2.4 List of wheat diversity studies based on agro-morphological and phenological traits, and references.

S. No.	Species/ Cultivars	Morphological traits	Reference
1.	*Ae*. *cylindrica* Host, *Ae*. *genic ulata* Roth and *Ae*. *neglecta* Req. ex Bertol; Fifty-two accessions	Thirteen morphological traits including thousand kernels weight, grain weight per spike, plant height and earliness.	Zahariev *et al*. (2003)
2.	Duram wheat (*Triticum duram*); Hundred and twenty-two entries	Eight morphological characters including awns, glume pubescence.	Tesfaye *et al*. (2008)
3.	Bread wheat (*Triticum aestivum* L.); Sixty-seven advance genotypes.	Twenty-one traits, *viz*., coleoptiles, anthocyanin colouration, plant growth habit, foliage colour, flag leaf anthocyanin colouration of a uricles, flag leaf, hairs on auricles, flag leaf attitude, time of ear emergence, days to maturity, waxiness of sheath, flag leaf waxiness of blade, ear: waxiness, waxiness of neck, plant length (excluding awns/scurs), ear shape in profile, ear length (excluding awns/scurs), ear density, presence of awns/scurs, awn length, 100-kernels weight and number of grains per spike	Maity and Das (2015)
4.	Bread wheat (*Triticum aestivum* L.); Thirty-six winter wheat cultivars of Iran.	Seventeen morphological traits, *viz*., stem diameter, plant height, leaf number at flowering, flag leaf length, flag leaf width, tiller number, internode length, peduncle length, spike length, floret number per spikelet, grain number per spike, awn length, grain diameter, grain length, days to flowering, thousand kernels weight and grain size.	Khodadadi *et al*. (2011)

Contd…

Table 2.4 Contd**…**

S. No.	Species/ Cultivars	Morphological traits	Reference
5.	Bread wheat (*Triticum aestivum* L.); Fifty-six genotypes	Eighteen morphological traits, *viz*., stem diameter, plant height, leaf number, leaf length, leaf width, tiller number, internode, peduncle length, floret number, spikelet number, grain number, awn length, grain diameter, grain length, number of days to flowering, thousand seed weight and grain yield	Sabaghnia *et al*. (2014)
6.	Bread wheat (*Triticum aestivum* L.); Forty-six genotypes	Three morphological traits, *viz*., plant height, spike characteristics and grain Morphology.	Ayala *et al.* (2015)
7.	Bread wheat (*Triticum aestivum)*; Fifteen genotypes	Eight morphological traits, *viz*., leaf area, days to heading, plant height, biomass, spike length, number of kernels per spike,1000 kernel weight and grain yield.	Khaled and Hamam (2015)
8.	Bread wheat (*Triticum aestivum* L.); Seven wheat varieties	Seven morphological traits, *viz*., plant height, days to heading, days to maturity, number of tillers per plant, peduncle length, grain weight per spike and grain yield per plant.	Khan *et al.* (2015)
9.	Bread wheat (*Triticum aestivum* L.); Twenty-three genotypes along with 2 checks	Nine morphological traits, *viz*., early vigor, leaf firing, days to heading, plant height, number of tillers/m, chlorophyll content Index (CCI) at 3 stages and grain yield.	Mishra *et al*. (2015)
10.	Bread wheat (*Triticum aestivum* L.); Twenty genotypes	Twelve morphological traits viz., days of heading, days to maturity, tiller number, plant height, spike length, 1000-kernel weight, grains per spike, grain yield, biomass yield, spike weight per plant, number of spike per plant and grain weight per spike.	Tilahun *et al.* (2015)
11.	Durum wheat (*Triticum turgidum*); Thirty-five genotypes (Kunduru landraces)	Three morphological traits, *viz*., 1000-kernel weight, grain protein content and yellow pigment.	Alsaleh *et al.* (2016)

Contd…

Table 2.4 Contd…

S. No.	Species/ Cultivars	Morphological traits	Reference
12.	Bread wheat; Twenty-four wheat cultivars and five wild progenitors of wheat	Five morphological traits, *viz*., grain number per spike, plant height, spikelet number per spike, spike length and 1000- grain weight.	Erayman *et al.* (2016)
13.	Durum wheat (*Triticum turgidum*)	Ten morphological traits, *viz*., Days to booting, days to flowering, days to maturity, plant height, number of effective tillers per plant, spike length, number of seeds per spike, biomass, grain yield per plant and 1000 -grain weight.	Mengistu *et al.* (2015)
14.	Spring wheat (*Triticum aestivum* L.)	Six morpho-physiological traits, *viz*., early ground coverage, canopy temperature at grain filling stage, peduncle length, relative water content, 1000-grain weight and grain yield per plant.	Rahman *et al*. (2016)
15.	Bread wheat (*Triticum aestivum* L.) Eighteen genotypes	Seven morphological traits viz., grain yield, biological yield, spikes/sq.m, harvest index, grain filling rate, protein yield and peduncle length.	Salehi *et al*. (2016)
16.	Bread wheat (*Triticum aestivum* L.); Sixteen varieties	Eight morphological traits, *viz*., plant height, days to flowering, days to maturity, effective tillers per plant, spike length, grains per spike, 1000-kernels weight and grain yield per plant were.	Sharma (2016)

were identified and being presently used to introgress potentially useful traits into wheat.

Tesfaye *et al.* (2008) assessed genetic diversity of seven tetraploid wheat populations comprised of 122 entries collected from Shoa and Gojem administrative Regions of Ethiopia, using eight morphological characters. According to their report, the awnness trait was fixed in all the accessions and glume hairiness had the lowest level of diversity for all the other characters. There was no relationship between the genetic diversity of genotypes and their origin. Further, they observed that only glume hairiness showed significant differences among accessions. Their analysis of variance of diversity for individual characters indicated that most of the variation was due to differences among districts rather than among accessions within their district-level origin which was further confirmed by a hierarchical analysis of variance of diversity. Therefore, they suggested that during future collection trips, one should cover more of different areas rather than collecting more samples from similar areas.

Imran *et al.* (2016) assessed genetic diversity in wheat synthetic double haploids for yield and drought related traits using factor and cluster analyses and showed variance among wheat genotypes in response to drought, and yield related traits.

Sarfraz *et al.* (2016) studied the diversity of 176 recombinant inbred lines derived from a cross between two durum wheat cultivars, and found a significant difference among durum wheat cultivars for all traits. Based on the estimated components of variation of morphological traits and yield in wheat populations, they discovered that an enormous aggregate of variability resides in the genotypes that can eventually contribute to the genetic improvement of wheat.

Based on agromorphological diversity studies, Donmez *et al.* (2001) showed that spike length is positively correlated with shoot biomass, straw biomass per plant, harvest index and grain yield. Schuler *et al.* (1994) reported that spike length also showed a strong correlation with spikelet density.

Maity and Das (2015) analysed the morphological trait-based diversity among 67 advance genotypes of bread wheat and characterized them according to different

morphologica traits. The evaluation of genotypes for twenty-one traits (Table 2.4) indicated that morphological traits were uniform and stable.

Wheat breeders are more concerned about study of inheritance for yield and associated traits (Khan *et al.*, 2015). Moreover, employed agromorphological traits and molecular markers complement each other and impart greater resolution to understand genetic diversity of plant genetic resources (Sharma *et al*., 2016). The success of any breeding programme depends on the presence of spectrum of genetic variations in existing germplasm collection (Sharma, 2016).

Khodadadi *et al.* (2011) studied 36 winter wheat (*Triticum aestivum* L.) genotypes of Iran for their genetic diversity and reported that most traits, other than days to emergence and days to heading, were statistically significant among different genotypes. According to their report, cluster analysis categorized the genotypes into seven groups. They obtained statistically significant differences among different genotypes except for emergence time and heading time. The study revealed that the first five components in Principal Component Analysis (PCA), explained over 97% of genetic variation. Considerable amount of variation existed among the genotypes for each character. Grain yield and number of tillers per plant recorded high coefficients of variation (CV). In their study, the genotype G55 recorded the highest thousand seed weight (45.57g) while the highest seed yield (6936.3 Kg ha^{-1}) was exhibited by genotype 4. Based on their study, it was recommended that identified differences in genetic component of traits can be utilized by wheat breeders as a new source of variation in constructing their breeding materials, designing their breeding schemes, choosing parents in crossing programs, and implementing selection strategies for genetic improvement of wheat.

Sabaghnia *et al.* (2014) analyzed agro-morphological variations for eighteen traits *viz*., plant height, stem diameter, leaf length, leaf width, leaf number, internodal length, tiller number, peduncle length, number of spikelets, number of florets, number of grains, awn length, seed size, seed length, days to flowering, thousand kernel weight and grain yield, smong 56 bread wheats (*Triticum aestivum*); and observed that there existed a considerable amount of variation among wheat genotypes for each investigated trait.

Mengistu *et al.* (2015) studied the agro-morphological diversity in a panel of 274 Ethiopian durum wheat accessions, where they analysed phenotypic diversity using Shannon–Weaver diversity (H′) index and multivariate analysis. While all the investigated traits had diversity, gulme pubescence showed no polymorphism for estimated H′. They reported that degree of seed shriveling had the highest (0.86) H′, possibly suggesting the differential response of genotypes to water deficit during seed-filling stages. Further, they observed the highest value (0.72) for the Bale district, while SNNP district recorded lowest H′ value (0.44). Altitudes 1600–2000 m above MSL and >3000 m above MSL exhibited the highest (0.76) and lowest (0.62) H′ values, respectively. These substantial variations of genetic diversity, attributed to districts of origin and altitude of collection sites, were observed through principal components analysis (PCA). Based on cluster analysis, genotypes included in their study were grouped into three clusters. Based on these results, they concluded that since Ethiopian durum wheat landraces exhibited considerable spectrum of agro-morphological diversity among districts of origin- as well as altitude-based populations, it can be effectively exploited for genetic improvement of wheat crop particularly with respect to yield and terminal drought tolerance.

Mishra *et al.* (2015) employed principal component analysis (PCA) and genotype by trait biplot (GT-biplot) analysis to investigate a total of 23 bread wheat (*Triticum aestivum*) genotypes, including 2 checks, with respect to their genetic diversity for 9 agro-morphological and physiological characters, namely, plant height, early vigor, days to leaf firing, days to heading, number of tillers per square meter, chlorophyll content index (CCI) at 3 stages of growth and grain yield. They reported a total of 5 principal components explaining 90.5% of total variation. It wass found that grain yield was positively correlated with tillers per square meter and early vigor, while negatively related to days to leaf firing. Through further analysis, apart from reporting a positive association among CCIs (Chlorophyll Content Indices) measured at 3 stages of growth, they found that yield made a 180-degree angle with days to leaf firing indicating that they were opposite in genotype ordering. Based on the results of their analyses, they suggested that considerable genetic variability in tested genotypes offers an excellent opportunity to exploit for wheat genetic improvement through hybridizing the genotypes belonging to distant clusters.

Malik *et al.* (2014) evaluated 258 wheat varieties, developed in last fifty years (1961-2010), in six different agro-climatic wheat growing regions of India, for distinctness, uniformity and stability (DUS) using 20 agro-morphological traits; and identified plant height, waxiness on plant parts, and growth habit as prominent morphological determinants for genetic diversity in Indian wheat. They assessed genetic variability among these varieties using GT biplot and ordination analysis. The first two principal component analysis (PCA) explained 36.9% of the total multivariate morphological variation, while four distinct groups detected by GT biplot analysis of tested Indian varieties. All the six different agro-climatic regions exhibited almost similar patterns of descriptor states and, therefore, selection criteria for high yielding genotypes that similar descriptors pattern, independent of climate and geographic regions, can be used for Indian wheat. Highly variable DUS descriptors identified in their study is useful to evaluate wheat genotypes at agro-morphological level as well as to make core collections of Indian wheat.

Zarkti *et al.* (2012) analyzed agro-morphological diversity in a panel of 467 Moroccan durum wheat (*Triticum turgidum* L. var. *durum*) accessions with the help of traits such as plant height, days to tillering, days to emergence, days to booting, days to head emergence, days to flowering, spike shape, spike density, days to physiological maturity and thousand kernel weight. Univariate and multivariate results indicated that 1000-kernel weight had the highest coefficient of variation of 15.72%, while 15.15% was recorded for plant height. Further, through two-dimensional principal coordinates analysis (2D-PCOA), they reported that apart from explaining 52% of the total variance, it grouped the tested accessions into three main groups: the shorter and early maturing type with 81 accessions, the taller and moderately late type having 154 accessions, and the taller and late maturing 232 accessions. Spike characters, based on the frequency distribution results, revealed that 54% were pyramidal, while dense spike constituted 42%. Based on their study, they suggested that their results augment the efforts of exploration, conservation and utilization of Moraccan landraces for genetic improvement of durum wheat.

Though morphological, biochemical and DNA markers can be deployed for diversity studies (Mohammadi and Prasanna, 2003; Sudre *et al.*, 2007; Goncalves *et al.*, 2009), the DNA markers are advantageous over the other type as they are not affected by environmental factors and gene interactions. In addition, since they involved advanced

methods of multiplexing and high-resolution genotyping (Binneck *et al.,* 2002; Antonio *et al.,* 2004; Saker *et al.,* 2005; Goncalves *et al.*, 2008; Souza *et al*, 2008), they are useful for detailed study of genetic diversity involving large number of genotypes or accessions.

Different types of molecular markers such as RAPD, AFLP, ISSR, SSR, SNP, DArT, DNA sequence analysis, etc., have been used by various researchers for studying genetic diversity, population structure and phylogeny in wheat and its close relatives, as summarized in the Table 2.5.

2.9.2 Microsatellite (SSR) markers

In a study conducted by Chen et al., (2012) to assess genetic diversity and population structure in a set of 90 Chinese winter wheat accessions, amplifying 269 SSR markers loci, distributed throughout the wheat genome, it was found that there was a total of 1358 alleles ranging from 2 to 10 alleles per locus and a mean genetic richness of 5.05. The average genetic diversity index was 0.60, with three genomes having values in the range of 0.05 to 0.86. Based on ANOVA they found that B and D genomes had the highest (0.63) and the lowest (0.56) genetic diversity, there were three subgroups based on STRUCTURE, UPGMA cluster, and principal coordinate analyses. In addition, STRUCTURE analysis showed that model-based population was positively correlated to some extent with geographic eco-type.

To understand the status of genetic diversity among 319 Indian wheat varieties, Arora *et al*. (2014) deployed 16 SSR markers, and DARwin (Version 5.0) and STRUCTURE (Version 2.3.3) analyses were employed to infer genotyping data using Jaccob's similarity model and Bayesian analysis, respectively. The results of their finding showed that there was a consistency in grouping obtained from both the methods and, therefore, they were competent to evaluate the distribution of genetic variability among different centers of wheat breeding in India; the genetic diversity among the groups (Ht; 0.38888) was higher than within groups (Hs; 0.3130), demonstrating that there existed less genetic variability among sub-populations (Gst; 0.19); there was a relatively high Gene flow (Nm) value, 2.0654; and, the genetic diversity was mainly present within the population, 81% as inferred from POPGENE (Version 1.32) and 78.5% as revealed from AMOVA. Based on these results, Arora *et al*. (2014) suggested that their findings are useful

Table 2.5 List of recent studies on wheat diversity using DNA markers.

S. No.	Wheat genotypes (Number)	Marker		Reference
		Type	Number	
1	Bread wheat varieties (56)	EST-SSR (Microsatellite)	32	Leigh *et al.* (2003)
4	Tibetan and Chinese Wheat (52)	SSR	206	Wang *et al*. (2007)
5	varieties (52)	SSR	48	Salem *et al*. (2008)
6	*Triticum aestivum* (4)	Lr34 sequence		Wicker *et al*. (2009)
7	*T. aestivum* (21)	SSR	32	Emon *et al*. (2010)
8	Ancestral *Triticum* and *Aegilops* species (20)	EST-SSR		Gadaleta *et al*. (2011)
9	*T. aestivum* cultivars (8)	ISSR & SSR (labelled)	5 & 2	Zhu *et al*. (2011)
10	Chinese Winter Wheat (90)	SSR	269	Chen *et al*. (2012)
11	*T. aestivum* (70)	Microsatellite	60	Jamalirad *et al*. (2012)
12	Wheat Cultivars (30)	ISSR & Marpho* traits	10 20	Najaphy *et al*. (2012)
13	Wheat genotypes (30)	SSRs	24	Spanic *et al*. (2012)
14	Wheat (*T. aestivum*) (20)	SSR	34	Sehgal *et al*. (2012)
15	(Tetraploid & hexaploid) Wheat (45)	SSR	11	Abouzied *et al*. (2013)

Contd..

Table 2.5 Contd…

S. No.	Wheat genotypes (Number)	Marker		Reference
		Type	Number	
16	Turkish bread wheat	SSR		Akfirat *et al*. (2013)
17	Argentinean bread wheat (*T. aestivum*) (102)	Biochemical and molecular	38	Leonardo *et al.* (2013)
18	Durum germplasm collection (150)	SNP	146	Ren *et al*. (2013)
19	Hexapolid bread wheat (*T. aestivum*) (30)	Microsatellite	37	Saedouie-Nasab *et al*. (2013)
20	Indian wheat varieties(319)	SSR	16	Arora *et al*. (2014)
21	European Bread wheat (49)	DArT	1849	Nielsen *et al*. (2014)
22	(Hexapolid & Tetrapolid)	SNP	90,000	Wang *et al*. (2014)
23	Tetraploid & Hexaploid Wheat (95)	RAPD ISSR	27	Khan *et al*. (2015)
24	Wheat (T. aestivum) (57)	SSR	50	Kumar *et al*. (2016a)
25	Wheat (*T. aestivum*) (54)	SSR	39	Kumar *et al*. (2016c)
26	Ethiopian durum Wheat (287)	SNP Am*. traits	30,155 & 10	Mengistu *et al*. (2016)
27	(*T. aestivum*) (60)	Microsatellite	60	Mwale *et al*. (2016)
28	Durum Wheat & Mediterranean landraces (172)	SSR & Morpho. traits	44 14	Sariano *et al*. (2016)

*Note: Morpho: Morphological, Am: Agro-morphological

in deciding effective parental combinations so that progenies will have maximum possible genetic variability.

Genetic diversity among 45 wheat cultivars comprising of two populations, 15 hexaploid (*Triticum aestvium*) and 30 tetraploid (*Triticum durum*), was studied by Abouzied *et al.* (2013) using 11 SSR markers. A total of 3840 amplicons observed on 6% denaturing polyacrylamide gel electrophoresis (PAGE) were utilized for molecular analyses; the Nei's genetic diversity (H) and the Shannon's Information index (I) for the two population were 0.2827 and 0.4533 with standard deviations ± 0.0699 and ± 0.0852, respectively, as inferred by POPGENE (Version 1.32); genetic diversity among populations (Hs=0.2761) was 97.7% of the total genetic diversity (HT=0.2827) and the proportion of the total genetic diversity within population that is attributed to population differentiation was low (Gst=0.0233) within the population; R-value was equal to 0.9048 ($P<0.0001$), as revealed by analysis of similarities (ANOSIM), indicating that all the most similar samples of genotypes were within the same population; the wheat genotypes from the four distinct regions grouped into two main clusters, durum wheat varieties and bread wheat varieties, as inferred by dendrogram and reiterated by the principal coordinate analysis (PCoA); and, finally, their results demonstrated that there existed a moderately considerable level of genetic diversity, and a little genetic differentiation, among the two populations.

Mwale *et al*. (2016) investigated genetic diversity of 60 wheat cultivars using 60 SSR marker loci and found that there was a total of 276 polymorphic alleles across 48 marker loci, with an average of 5.7 alleles per locus; a total of 12 loci were not amplified across 60 cultivars; subgenome-A had the highest genetic diversity followed by subgenome-B, and subgenome-D had the lowest; as revealed by Shannon index and polymorphic information content (PIC)-based diversity measures, inferred from allelic richness of each locus as well as genetic richness of each subgenome. Four clusters comprised of 3, 7, 9 and 41 genotypes each; with genetic distance between clusters varying from 0.56 to 0.87, with most cultivars showing high diversity between genetic distances of 0.65 and 0.75. Based on their findings, it was concluded that the established genetic

relationships and four clusters are useful to breeders for breeding new disease resistant cultivars and their rational deployment in wheat production.

In a study by Wang *et al.* (2007) to analyze genetic diversity in 52 accessions from three unique Chinese wheat varieties, using 206 SSR marker loci, it was found that a total of 488472 and 308 allelic variants were detected in 31 Yunnan, 15 Tibetan and 6 Xinjiang wheat accessions with an average of PIC values 0.2764, 0.3082, and 0.1944, respectively. Substantial variations in allelic richness were detected by SSR markers in all the 21 homologous chromosomes, the 7 homoeologous groups, and the three subgenomes (A, B, and D) in Yunnan, Tibetan, and Xinjiang wheats; among the seven homoeologous groups, the lowest and highest allelic polymorphisms were observed in groups 6 and 3, respectively; among 21 homologous chromosomes 3B and 1D had the highest and lowest allelic polymorphisms, respectively; among three subgenomes, B showed the highest, A the intermediate, and D the lowest allelic polymorphism; genetic diversity was much higher within Yunnan and Tibetan wheat than within Xinjiang wheat, but between Yunnan and Tibetan GD value was much lower than those between: Yunnan and Xinjiang wheat; and Tibetan and Xinjiang wheat; whereas, Yunnan and Tibetan wheats were closely related than were Yunnan and Xinjiang wheats, or Tibetan and Xinjiang wheats.

2.9.2.1 EST-SSR vs Genomic-SSR

In a study by Varshney *et al*. (2007) combining 4 AFLP markers with 16 EST-SSRs and 15 EST-SNPs to analyze diversity among 43 wild, 35 cultivated and 12 elite barley lines; the SSRs showed the most polymorphism (average PIC, 0.593) and allelic richness of 8 per locus. Based on the results, they suggested that SSR markers are more suitable for diversity analysis and fingerprinting.

Leigh *et al.* (2003) analyzed the impact and performance of EST-SSR (20) vis-à-vis genomic-SSR (12) in 56 varieties of bread wheat, and found that the PIC values for genic markers were generally lower than those computed for genomic SSRs.

Gadaleta *et al.* (2011) compared the usefulness of EST-SSRs and genomic-SSRs in phylogenetic analysis of wheat. They reported that when investigated for their transferability and level of DNA polymorphism (number of alleles/ locus) using distinct tetraploid (*Triticum* spp) and diploid (*Aegilops* species) ancestors, differences were found

in both the markers. Their phylogenetic trees derived from EST-SSR and genomic-SSR resembled cytogenetic analyses-based trees.

2.9.2.2 Allelic diversity of microsatellite markers

In a study, Kumar *et al*. (2016a) deployed 50 SSR marker loci to evaluate genetic diversity of the fifty-seven wheat (*Triticum aestivum* L.) genotypes and found that there was a high degree of genetic polymorphism (70% among genotypes); a total of 114 alleles were polymorphic with an average of 2.71 ranging from 2 to 6 per locus, and the average genetic diversity was 0.465 with a range of 0.037- 0.892.

In another study involving 54 wheat genotypes (41 Indian origin and 13 exotic) and 39 SSR markers, Kumar *et al.* (2016c) reported a total of 112 polymorphic alleles ranging from 1-7 with an average of 2.87 per marker locus; PIC values ranging from 0.03 to 0.49 with a mean of 0.29, indicating lower level of genetic diversity. Broadly four clusters (A, D, B and D) were revealed by UPGMA-based cluster analysis, wherein seven further sub-clusters were found in the Cluster D that included maximum (43) number of genotypes.

2.9.2.3 Fluorescence-labelled microsatellite markers

In a genetic discrimination study by Zhu *et al.* (2011) of 8 closely related wheat cultivars, 5 ISSR markers and 2 fluorescently labelled SSR markers were deployed and it was found that there was a total of 43 distinct reproducible bands of which 29 (67.44%) were polymorphic; the number of polymorphic bands per marker ranged from 3 to 8 with an average of 4.8 per marker; one ISSR primer ‘UBC-849’ was able to identify all the eight wheat cultivars; 2 fluorescently labelled SSR markers loci amplified a total of 29 alleles with an average of 14.5 per locus in 8 wheat cultivars, and all were polymorphic alleles; cultivars were discriminated based on numbers and sizes of alleles; and, since two fluorescence-labeled TP-M13-SSR primers together distinguished all the eight closely related wheat cultivars, a combination of them was successfully utilized to fingerprint 8 closely related wheat cultivars. Based on findings of their study, it was suggested that for fingerprinting and identification of wheat cultivars, fluorescence-labeled SSR marker are highly useful, as they are amenable for high-throughput and high accuracy genotyping.

Twenty-four fluorescence-labelled microsatellite (SSR) markers were deployed by Spanic *et al.* (2012) to analyze molecular genetic diversity of 30 wheat genotypes and it was found that there were 1 to 14 polymorphic markers per locus with a mean of 8.44 alleles per marker locus; the subgenome-A had the highest number of 7.2 alleles per locus, followed by subgenomes B and D having 5.9 and 5.0, respectively; and, chromosome 7 had the highest number 9.5 of alleles, while chromosomes 3 and 4 had the lowest number of 5.0 and 5.3 alleles, respectively.

2.9.2.4 SSR-based phylogenetic study

Jamalirad *et al.* (2012) amplified 60 SSR marker loci to assess genetic diversity and phylogenetic relationship among 70 bread wheat genotypes and recorded that there were a total of 309 polymorphic alleles with an average of 9.26 allele per marker locus; Nei gene diversity for the 42 SSR marker loci varied from 0.4 to 0.91 with an average of 0.74; polymorphism information content (PIC) value ranged from 0.365, for the barc87 and the second locus of barc165, to 0.902, for the xgwm213, with a mean of 0.688; neighbour-joinig algorithms-based cluster analyses performed using distance coefficient matrix demonstrated that phylogenetically most related genotypes (based on pedigree information), were grouped in the same cluster; the Rogers distance-based dendrogram demonstrated the high concordance with available pedigree information of genotypes; and, the evidence was further corroborated with the results of principal components analysis (PCA).

2.9.3 Single nucleotide polymorphism (SNP) markers

The following studies demonstrate that single nucleotide polymorphism (SNP) markers are effective in deciphering genetic diversity in wheat.

In *Triticum turgidum* (subsp. durum), a key durum wheat crop across the world but yet to be exploited, breeders face the challenges of limited allelic diversity in elite pool making it difficult to breed the crop for adaptation to evolving environmental threats. Therefore, extensive molecular and phenotypic characterization of unexplored material may throw light on novel loci useful to international crop breeding through quantitative trait loci (QTL) mapping. In a genomic survey work carried out by Mengistu *et al.* (2016) to ascertain the diversity, structure, and genome-specific variations present among 287

Ethiopian wheat germplasm accessions, utilizing 30155 SNP markers and 10 agronomic traits, it was found that Ethiopian germplasm entries were unique among Mediterranean durum wheat accessions, indicating their importance and need for exploration; and, there were several loci underpinning traits of agronomic importance, some corroborating previous reports and the others describing new promising genomic regions. Based their findings it was concluded that the panels deployed in this study are useful to integrate QTL maps that set an initial platform facilitating untapped diversity in the Ethiopian lines to be exploited for wheat improvement.

In a study undertaken by Ren *et al.* (2013), 150 worldwide durum wheat accessions were analyzed for genetic diversity and population structure among them deploying 946 polymorphic SNP markers covering the whole of the tetraploid wheat genome. Based on the results, it was deciphered that several factors; including large differences in prevalent environmental conditions, breeding techniques used at the period of varietal release, and man-mediated gene flows; greatly impacted the genetic structure. Modern cultivars released during the early years of Green-Revolution period, old cultivars and landraces showed loss of genetic diversity, whereas more diversity was found in cultivars released during the Post-Green Revolution. North America, South America, and Europe had the richest genetic diversity and Middle East showed the moderate levels, as inferred from a comparative assessment among the 10 mega eco-geographical regions of the wheat world.

Since understanding of marker–trait associations, genomic diversity, and inferring ancestral relationships among individuals in populations requires a powerful tool such as high-density single nucleotide polymorphism (SNP) genotyping arrays, Wang *et al*. (2014) deployed genetic diversity assessment in allohexaploid and allotetraploid wheat populations to validate the usefulness of 90000 gene-associated SNPs in genotyping array.They included in the array most of genome-wide distributed SNPs common to populations of diverse geographical origin. In fact, sequence divergence and/or deletions at the target SNP-site results in multiple clusters having low signal-intensity and, in addition, because of polyploid nature of the wheat the SNP data-set was a complex one. Yet, they could achieve accurate genotype-calling utilizing density-based spatial clustering, a model-free clustering algorithm. the distribution of presence-absence variation (PAV) in wheat populations was inferred from assays capable of detecting low-

intensity clusters. A combination of eight mapping populations were deployed to map a total of 46977 SNPs from the wheat 90K array; and, finally it was concluded that they developed an array and cluster-identification algorithms that constitute an invaluable genomic resource for inferring detailed haplotype-structure which is crucial in studying genetic diversity and genetic-basis of trait variation in polyploid and complex genome of wheat.

2.9.4 Diversity array technology (DArT)

Wenzl *et al.* (2006) integrated the genotyping data generated by DArT with those obtained from restriction enzyme fragment polymorphism (RFLP), sequence tagged site (STS) and/or simple sequence repeat (SSR) and generated a consensus map; using several purpose-built perl scripts and JoinMap 3.0 software; that comprised of 2935 loci (2085 DArT, 850 other loci) spanning 1161 cM and having a total of 1629 'bins' (unique loci), with an average inter-bin distance of 0.7 ± 1.0 cM (median, 0.3 cM); single DArT assay covered greater than 98% of the map. The loci could be covered with a single DArT assay; and, consensus as well as component maps showed the same sequence of marker arrangement of loci.

In a Diversity Array Technology (DArT) study by Nielsen *et al*. (2014) conducted to ascertain diversity and characterize a population of 94 European bread wheats (*Triticum aestivum* L.), a total of 1849 markers was found to be polymorphic out of 7,000 tested markers, and were used for analyzing population structure; where two major subgroups of wheat varieties, GrI and GrII, were inferred by the STRUCTURE software, and the same was further confirmed by principal component analysis (PCA); two subgroups were largely separated according to their origin: varieties from Southern and Eastern Europe belonged to GrI, whereas most of the modern varieties from Western and Northern Europe constituted GrII. The markers located near reduced height 8 (*Rht8*) on chromosome 2D locus formed a large bulk of the markers that mostly contributed to the genetic separation of the subgroups; and, the results suggested that breeding for the *Rht8* had the impact on subgroup separation.

2.9.5 DNA sequence-based diversity

To study intra-species genomic diversity in wheat (*Triticum aestivum*) and barley (*Hordeum vulgare*) that have repetitive, large, polyploid and complex genomes; and, whose varieties have evolved in agro-ecosystems since the time of domestication; large genomic sequences at the *Lr34* locus of the wheat varieties: Chinese Spring, Renan, and Glenlea; diploid wheat *Aegilops tauschii*;and, along with the barley loci *Vrs1* and *Rym4* of its varieties: Morex, Cebada Capa, and Haruna Nijo; were comopared by Wicker *et al.* (2009) and molecular dating deciphered that while there were post-domestication gene flow events from wild barley relatives, it was rare to absent in the D genome of hexaploid; and, based on such observations, it was concluded that wheat D-genome haplotypes diverged only a few thousand years ago whereas it was more than 500000 years ago that some barley and *Ae. tauschii* haplotypes diverged.

2.9.6 Combination of more than one type of markers

2.9.6.1 RAPD & SSRs

Ninety-five tetraploid and hexaploid wheat genotypes, belonging to India and Turkey, were studied by Khan *et al*. (2015) using combined analysis with RAPD and ISSR markers, and found that there were 177 polymorphic alleles. Similar groupings of the genotypes were found in both the dendrogram and two-dimensional scatter plot. They reported two clusters of Turkish hexaploid varieties of which one showed close association with hexaploid and the other with tetraploid varieties from India. High (77%) genetic variation within Indian as well as Turkish populations was evident from AMOVA. The wheat genotypes showed distinct clustering according to geographic origin as well as ploidy, as elucidated by STRUCTURE analysis.

2.9.6.2 Biochemical and molecular markers

Leonardo *et al*. (2013), in their work on analysis of genetic structure in a set of 102 Argentinean bread wheat (*Triticum aestivum* L.) elite cultivars, utilized 38 biochemical and molecular markers that were distributed throughout 18 wheat chromosomes and belonged to a wide range of functional category from neutral though closely-linked to functional genes. The results of their study showed that model-based clustering methods allocated the

genotypes among three subpopulations that largely corresponded to origin of the accessions.

2.9.6.3 Molecular markers and morphological traits

Salem *et al.* (2008) evaluated the genetic diversity of 7 wheat varieties using 9 morphological traits and 48 SSR marker loci, and found that there was a total of 48 polymorphic alleles, ranging from 2 to 7 alleles per locus with an average of 3.2; polymorphism information content (PIC) values ranged from 0.278, for the xgwm95, to 0.816, for the xgwm437, with an average of 0.548; both morphological and SSR marker data could differentiate the genotypes; morphological trait-based average genetic diversity (23.49 with a range of 8.51-38.46) was higher than that based on SSR markers (0.53 with a range of 0.42-0.63); and, the diversity among varieties was analysed by using a dendrogram that was generated from genetic dissimilarity values between genotypes computed by the SSR marker genotyping data. It was suggested that the information on grouping of the genotypes, based on morphological and SSR marker data, will be useful for wheat breeders to plan crosses for positive traits; and, that the wheat microsatellite primers can be used to distinguish the genotypes and estimate their genetic diversity.

Najaphy *et al.* (2012) combined 20 morphological traits, including agro-morphological and phenological, with DNA-marker (10 ISSR markers) data to study genetic diversity of 30 wheat cultivars and advanced breeding lines. They found that most of the 20 phenotypic traits showed significant variations and revealed genetic differences among the wheat genotypes; the genotypes were clustered into four groups as assigned by WARD method; fifty-nine out of 86 (80.2%) ISSR alleles were polymorphic; PIC values for majority of the markers ranged from 0.21 to 0.23 (indicating diversity of accessions and high-informativeness nature of the markers); molecular marker-derived dendrogram and principal coordinate (PCoA) analyses showed grouping with some minor deviations clustering of the genotypes into five groups; even though both morphological traits and molecular markers revealed a good level of genetic variability among the genotypes, there was no correlation between them. It was suggested that information on genetic diversity among 30 wheat genotypes as revealed by 10 ISSR and 20 morphological traits would be useful in traditional and molecular breeding programs.

In a study by Soriano *et al.* (2016) that involved analyses of genetic structure among 172 durum wheat landraces, utilizing both 14 agro-morphological traits (including yield, physiological and phenological traits) and 44 SSR marker loci data, it was found that there was a total of 448 polymorphic alleles, of which 226 were rare alleles; there was an average of 10 alleles per locus; all the alleles were fixed in 59% of the marker loci; total genetic diversity (HT) was 0.708 and the genetic differentiation (GST) value was 0.1730; landraces from the western Balkans and Egypt had the highest genetic diversity (HT =0.6979) whereas the modern cultivars had the lowest (HT =0.4835); genetic variability within the subpopulations was higher (83%) than that among the populations (17%), as inferred by AMOVA; clusters elucidated by dendrogram analysis matched with that inferred by STRUCTURE analysis with the latter improving the classification of accessions that had large admixture; and, STRUCTURE software allocated 90.1% of the genotypes in five subpopulations: one including all modern cultivars, and the four containing landrace related to their geographic origins: eastern Mediterranean, eastern Balkans and Turkey, western Balkans and Egypt, and western Mediterranean; landraces belonging to eastern Balkans and Turkey that were allocated within a single subpopulation, based on molecular data were separated by phenotypic data-based dendrogram into two branches, suggesting different origins for these landraces; and, finally, it was observed that morphological data-based population structure had similarity with that derived from DNA marker data, indicating that they were connected by geographic origin of the landraces.

2.9.7 Practical applications of findings of diversity studies

Emon *et al*. (2010), in their study to identify and characterize boron efficient varieties in a panel of 21 diverse wheat (*Triticum aestivum* L.) genotypes, deployed 32 SSR markers and reported a total of 234 polymorphic alleles with an average of 7.3 alleles per locus, ranging from 3 to 11 alleles per locus. The PIC values ranged from 0.562 to 0.873 with an average of 0.776. Overall genetic diversity across the 32 loci and 21 genotypes ranged from 0.637 to 0.884 with a mean of 0.804. Cluster analyses, NJ tree, UPGMA, and PCO, identified a similar pattern of variations. The study found that INIA-66 and BAW-1086 were the most boron-efficient genotypes and useful for developing boron-efficient varieties.

Sehgal *et al.* (2012) performed molecular characterization and genetic diversity of 20 wheat genotypes, using 34 SSR marker loci, and reported that genotypes Shalimar-86 and Chakwal-86 showed the maximum genetic diversity (0.98 dissimilarity index) with SH-02 and Ufaq, respectively, and the lowest genetic diversity (dissimilarity index, 0.74) was observed in Chakwal-50 and Bhakar; SSR makers could distinguish and characterize all the genotypes for rust resistance genes; and, therefore, SSR markers deployed in their study are useful for genetic diversity analysis required for achieving the saturation of different genomic regions through further research. In addition, identification of rust resistant genes in Pakistani wheat germplasm can accelerate the pyramiding of different rust resistant genes in wheat breeding programs.

In a study conducted by Sardouie-Nasab *et al.* (2013), 37 SSR marker loci, associated with QTLs for salinity tolerance, were amplified in 30 promising lines of hexaploid bread wheat (*Triticum aestivum* L.). It was found that there was a total of 438 polymorphic alleles with a mean of 11.84 per locus; the marker xgwm312 had alleles ranging from 2 to 20; xgwm445 and xgwm312 had gene diversity indices ranging from 0.66 to 0.94 and polymorphism information content (PIC) values in the range of 0.64 to 0.93, respectively; since the marker xgwm312 had the highest PIC value, it was regarded as the best marker for genetic diversity analysis with respect to salinity tolerance; the genotypes were assigned into 3 distinct clusters by UPGMA-based dendrogram, and they had distinct reaction to salinity; all the genotypes showed a wide range of genomic diversity among them; this pattern of genetic diversity was corroborated by principal coordinates analysis (PCoA); and, finally, it was suggested that identified genotypes are useful as prime candidates in wheat breeding program aiming to improve salinity tolerance in wheat. In another similar study Akfirat *et al.* (2013) amplified 230 SSR marker loci and found xgwm382 as a molecular diagnostic marker for yellow rust resistance.

2.10 Studies on practical application of SSR markers used in the present study

Various studies related to wheat microsatellite (SSR) markers and associated or linked genes/traits is summarized in Table 2.6. Five microsatellite markers, *viz.*, xgwm257, xgwm268, xgwm458, Xgwm183 and BARC1121 have no reports on association/linkage with any genes/traits.

Table 2.6 SSR markers, their chromosomal locations, associated/linked genes/traits and references.

S. No.	SSR Marker	Chr. Location	Associated with/ Linked to		Reference
			Trait	**QTL/Gene**	
1	Xgwm136	1 AS	Tiller inhibition	*Tin* gene	Kumar *et al*. (2015)
2	Xgwm497	1 AL	Cell membrane stability	Drought tolerance QTL *QCMSb2AC*	Malik *et al*. (2015)
3	Xgwm550	1 BS	Resistance to bacterial leaf streak (BLS)	BLS resistant QTL	Kandel *et al*. (2015)
4	Xgwm268	1 BL	Information not available		-
5	Xgwm458	1 DS	Information not available		-
6	Xgwm642	1 DL	Sedimentation volume (SED)	*QSed.Z86-1D.a* QTL	Kunert *et al*. (2007)
7	Xgwm448	2 AS	Number of seeds per spike	*QKNPS-DH-2A* QTL	Zhang *et al*. (2016);
			Thousand kernel weight	*Qtl5-8*;*Qtl12-14* and *Qtl16*	Wang (2012)
8	Xgwm445	2 AL	Resistance to multiple leaf sot diseases.	*QSb.bhu-2A QTL*	Gurung *et al*. (2014)
9	Xgwm257	2 BS	Information not available		-
10	Xgwm47	2 BL	Resistance to stripe rust	Putative QTL	Badakhshan *et al*. (2008)
11	Xgwm484	2 DS	Days to heading	Putative QTL	Heidari *et al*. (2012)
12	Xgwm349	2 DL	Susceptibility on Recognition of the *Parastagonospora nodorum* Necrotrophic Effector SnTox7	*Snn7* gene	Shi *et al*. (2015)

*Note: Chr = Chromosomal;

Contd..

Table 2.6 Contd…

S. No.	SSR Marker	Chr. Location	Associated with/ Linked to		Reference
			Trait	**QTL/Gene**	
13	Xgwm2	3 AS	Productive tillers per plant; grain yield per plant	Putative QTL	Gupta *et al*. (2015)
14	Xgwm494	3 AL	Flag leaf length to width ratio	*Qflwr.acs–3A.3* QTL	Yang *et al*. (2016)
15	Xgwm533	3 BS	Resistance to Fusarium head blight	*Fhb1* QTL	Anderson *et al*. (2001); Buerstmayr *et al*. (2002)
16	Xgwm181	3 BL	Spike harvest index	Putative QTL	Maryam *et al.* (2009)
17	Xgwm183	3 DS	Information not available		-
18	Xgwm3	3 DL	Germination Index (GI)	Putative QTL	Kim *et al*. (2014)
19	Xgwm165	4 AS	Stripe rust resistance	*QYr.caas-4DL* QTL	Ren *et al*. (2012); Chen (2013)
20	Xgwm637	4 AL	Thousand kernel weight (TKW)	A putative QTL	Zhang *et al*. (2013)
21	Xgwm368	4 BS	Resistance to disease caused by *Pratylenchus neglectus*	*QRlnn.lrc-4B.1* QTL	Zwart *et al*. (2008)
22	Xgwm513	4 BL	Dwarfing effect	*Rht-B2* gene	Jin *et al*. (2013)
23	Xgwm194	4 DS	sterile spikelets at the basal part of the spike (SSPN), Spike weight (SW), and grain weight (GW)	Three putative QTLs	Trkulja *et al*. (2012)
24	Xgwm624	4 DL	Adult plant resistance to stem rust	An allele of *Sr* gene	Zhang *et al*. (2014)

*Note: Chr = Chromosomal;

Contd..

Table 2.6 Contd…

S. No.	SSR Marker	Chr. Location	Associated with/ Linked to		Reference
			Trait	**QTL/Gene**	
25	Xgwm443	5 AS	Adult plant resistance to stem rust	*Lr52* gene	Tar *et al*. (2008); Hiebert *et al*. (2005)
26	Xgwm666	5 AL	Thousand kernel weight (TKW)	A putative QTL	Zhang *et al*. (2013)
27	Xgwm544	5 BS	Glutamine synthetase activity in leaf	*Q.GSA-5B* QTL	Li *et al*. (2015)
28	Xgwm335	5 BL	*Septoria tritici* blotch (STB) disease rsistance	*Stb1* gene	Adhikari *et al*. (2004)
29	Xgwm190	5 DS	Yellow berry (YB) tolerance	Putative QTL	Ammiraju *et al*. (2002)
30	Xgwm182	5 DL	Frost resistance	Putative QTL	Halaieva *et al*. (2013)
31	Xgwm459	6 AS	Stripe rust resistance	Putative minor QTL	Hao *et al*. (2011)
32	Xgwm169	6 AL	Stripe rust resistance	*YrQ*3 QTL	Chen (2013)
33	Xgwm361	6 BS	Gran protein content	*Gpc-B1* gene	Vishwakarma *et al*. (2014)
34	Xgwm219	6 BL	Stem water-soluble carbohydrates (WSC)	Putative QTL	Li *et al*. (2015)
35	Xgdm141	6 DS	Resistance to Hessian fly	H13 gene	Liu *et al*. (2005)
36	BARC1121	6 DL	Information not available	-	-

*Note: Chr = Chromosomal;

Contd..

Table 2.6 Contd…

S. No.	SSR Marker	Chr. Location	Associated with/ Linked to		Reference
			Trait	**QTL/Gene**	
37	Xgwm635	7 AS	Powdery mildew resistance	*MlNCD1* gene	Maxwell *et al*. (2012)
38	Xgwm282	7 AL	Thousand kernel weight (TKW)	Putative QTL	Kuzmanović *et al*. (2014); Huang *et al*. (2004)
39	Xgwm400	7 BS	Thousand kernel weight (TKW)	Putative QTL	Hai *et al*. (2008); Borner *et al*. (2002); Quarrie *et al*. (2006)
40	Xgwm131	7 BL	Fertility restoration against *Triticum timopheevii* cytoplasm	*Rf3* gene	Zhou *et al*. (2005)
41	Xgwm44	7 DS	Russian wheat aphid resistance	*Dn6* gene	Liu *et al*. (2002)
42	Xgwm37	7 DL	Stripe rust resistance	*YrY201* gene	Zhang *et al*. (2008)

*Note: Chr = Chromosomal;

Kumar *et al*. (2015) reported that a SSR marker xgwm136 is detected to be associated with tiller inhibition (*tin*) gene, known to reduce the number of tillers in a wheat plant. They observed nearly perfect co-segregation of the marker and the number of tillers per square meter area and suggested that the marker can be used in marker assisted selections (MAS) for the number of tillers in wheat.

Quantitative trait loci (QTLs) related to drought tolerance in wheat can be manipulated and detected using SSR markers for improving drought tolerance, in the face of global challenges that water shortage is causing a major constraint to crop production in many parts of the world. Malik *et al*, (2015) deployed microsatellite markers for mapping QTLs for photosynthesis, cell membrane stability and relative water content in wheat using F_2 population-derived from the intra-specific cross: Chakwal-86 (drought tolerant cultivar) x 6544-6 (drought sensitive genotype). It was found that QTL QCMSb2AC, confering cell membrane stability, was located at 4.4 cM form xgwm 497 with a LOD score of 2.7.

Kandel *et al.* (2015) conducted a study to map quantitative trait loci (QTL) controlling resistance to bacterial leaf streak (BLS; caused by *Xanthomonas campestris* pv. *translucens*) in spring wheat using an identity by descent (IBD) mapping approach and observed that xgwm550 was significantly associated with resistance under field conditions.

Kunert *et al.* (2007) utilized advanced backcross QTL (AB-QTL) strategy to locate quantitative trait loci (QTLs) for baking quality traits in two BC_2F_3 populations of winter wheat, derieved from a cross: *T. dicoccoides* X *T. tauschii* and found that the greatest effect of a putative QTL QSed.Z86-1D.a, located at xgwm642, explained 32.8% of the genetic variance for sedimentation volume (SED).

Zhang *et al*. (2016) detected that xgwm448 is linked to a QTL, QKNPS-DH-2A, governing the number of seeds per spike (NSS) which; along with the spike number per m^2 (SN), and thousand-kernel weight (TKW); forms the three main components determining wheat productivity. Kernel weight is a key factor influencing wheat processing quality, which is determined by quantitative trait loci(QTL). Wang *et al.* (2012) used recombinant inbred lines (RIL) population derived from a cross: 6044X1-35 and showed that xgwm448 was linked to eight of the total sixteen QTL of thousand kernel weight detected in different seed filling periods at 10 days interval.

Gurung *et al.* (2014) reported that QTL, *QSb.bhu-2A*, effective against major leaf spot diseases of wheat, was found be located some where between the two markers, *xbarc353* and *xgwm445, with a* marker interval of 37.4 cM.

Stripe rust is an important wheat disease worldwide and resistance to stripe rust is often not controlled by race-specific resistance genes. Badakhshan *et al.* (2008) used 329 $F_{2:3}$ families from a cross, between highly resistant winter type cv. MV17 and highly susceptible spring type local cv. Bolani, to identify quantitative trait loci (QTL) affecting resistance to stripe rust. The population was evaluated for adult plant resistance to pathotype 134E134A+ of *Puccinia striiformis*, in the field condition. A set of 56 SSR markers, polymorphic between parental genotypes, were screened in the population and linkage map was constructed using these markers. Based on composite interval mapping (CIM), xgwm47 was found to be significantly linked to QTL on 2B chromosome having 8% effect with dominance.

Based on their study using doubled haploid, Heidari *et al.* (2012) reported that QTL associated with days to heading was linked to a microsatellite markers xgwm484 located on chromosome 2D.

Parastagonospora nodorum is a necrotrophic fungal pathogen that causes the *Septoria nodorum* blotch (SNB) disease in wheat. The fungus produces necrotrophic effectors (NEs), SnTox7, that is recognized by products of host gene *Snn7*. Shi *et al.* (2015) reported that the *Snn7* gene governs sensitivity to SnTox7 and was linked to microsatellite marker xgwm349 at distances of 1.8 cM.

Gupta *et al.* (2015) reported that two markers were detected to have association with QTL for number of productive tillers/plant with significance at 5% level and one marker xgwm2 was detected to have significantly been associated with QTL for grain yield/plant, at 5% level, mapped on chromosome 3A.

Yang *et al.* (2016) reported that a QTL, Qflwr.acs–3A.3, governing flag leaf length to width ratio was tightly linked to microsatellite marker xgwm494 on long arm of chromosome 3A.

Anderson *et al*. (2001) positioned *Fhb1* between the markers xgwm493 and xgwm533 with an approximate genetic distance of 7 cM. Almost the same genetic distance (8 cM) between these two markers was also stated by Buerstmayr *et al*. (2002).

Maryam *et al.* (2009) reported that xgwm181-3B marker was found to be associated with spike harvest index. xgwm3-3D Kim *et al.* (2014) reported that markers xgwm3-3D was found to be associated with germination index. Ren *et al.* (2012) as well as Chen (2013) reported that xgwm165 is detected to have been associated with one of the QTL "QYr.caas-4DL" confering resistance to stripe rust disease in wheat. Zhang *et al.* (2013) reported that microsatellite locus *xgwm637* was detected to be associated with a putative QTL governing thousand kernel weight (TKW).

Zwart *et al.* (2008) reported that xgwm368 is detected to be associated with QTL *QRlnn.lrc-4B.*1 that confers resistance to disease caused by *Pratylenchus neglectus.* Jin *et al.* (2013) reported that microsatellite loci xgwm513 was closely linked to dwarfing gene *Rht-B2* suggesting that the gene was mapped to the 4B long arm in genotype Huaai 01. Trkulja *et al*. (2012) reported that xgwm 194 was detected to be associated with sterile spikelets at the basal part of the spike (SSPN), Spike weight (SW), and grain weight (GW) with LOD scores of 2.929, 2.511 and 2.575, respectively.

Zhang *et al.* (2014) reported that xgwm624 on chromosome 4DL was detected to be associated with an allele of stem rust resistance gene *Sr* located on 4 DL. Tar *et al.* (2008) as well as Hiebert *et al*. (2005) reported that xgwm443 was the closest marker linked to leaf rust resistance gene *Lr52* located on 5AS. Zhang *et al.* (2013) reported that microsatellite loci xgwm666 was detected to be associated with a putative QTL, on chromosome 5AL, governing thousand kernel weight (TKW). Li *et al.* (2015) reported that microsatellite marker xgwm544 was detected to be associated with a QTL, *Q.GSA-5B*, on chromosome 5B contributing to glutamine synthetase activity of wheat leaves at different grain filling stages.

Adhikari *et al.* (2004) reported that Microsatellite loci xgwm335 was proximal (7.4 cM) to *Stb1* gene governing resistance to *Septoria tritici* blotch (STB) disease. Yellow berry (YB) is a serious seed disorder in wheat, caused by deficiency of nitrogen in the soil. YB adversely affects the grain protein content (GPC) and therefore bread making quality in bread wheat. Ammiraju *et al.* (2002) studied the inheritance and DNA markers associated with YB tolerance, using a recombinant inbred line (RIL) population of 113 individuals developed by making a cross between Rye Selection111 (RS111), highly resistant to YB, and Chinese Spring (CS), a susceptible parent. Phenotyping of this population for YB incidence indicated that, at least one major

gene/QTL and a few minor genes govern the tolerance to YB. Microsatellite marker xgwm190 on chromosome 5D was detected to be associated with YB tolerance.

Halaieva *et al.* (2013) performed analysis of frost resistance and association of microsatellite of the group-5 chromosomes on parental varieties and recombinant-inbred lines F_7 derived from a cross: *Luzanovka odesskaya* X *Odesskaya krasnokolosaya*, and reported that xgwm182-5D microsatellite loci was associated with the level of frost resistance of the lines.

Stripe rust is an important disease of soft red winter wheat caused by *Puccinia striiformis* f. sp. *tritici*. Hao *et al*. (2011) performed association studies using a set of 178 recombinant inbred lines (RILs) developed using single-seed descent from a cross: Pioneer 26R61 (resistant) X AGS 2000 (susceptible). Microsatellite marker xgwm459 loci was detected to be associated with a minor QTL located on chromosome arm 6AS.

Chen (2013) reported that microsatellite marker loci xgwm169 is detected to be associated with one of the QTL *YrQ*3, confering resistance to stripe rust disease in wheat. High grain protein content (GPC) in wheat has been a major trait of interest for breeders since it has enormous end use potential. Vishwakarma *et al.* (2014) reported that microsatellite marker locus *xgwm169* is linked to a gene *Gpc-B1*governing high grain protein content.

Wheat grain yield is drastically affected by environmental conditions like drought, unfavorable especially for photosynthesis, during the grain filling stage that can be buffered by stem water-soluble carbohydrates (WSC). Li *et al.* (2015) conducted an association analysis using 262 winter wheat accessions and 209 genome-wide SSR markers, based on a mixed linear model (MLM), and reported that xgwm219 is detected to be associated with QTL contributing to stem water-soluble carbohydrates (WSC).

Maxwell *et al*. (2012) genetically characterized a novel *Aegilops tauschii* Coss.-derived wheat powdery mildew resistance gene *MlNCD1* present in the germplasm line NC96BGTD1 using F_2- and F_4-derived lines from a cross: NC96BGTD1 × 'Saluda' and found that it was a monogenic trait and was linked to SSR marker locus *xgwm635* at 5.5 cM distance.

Kuzmanovic *et al.* (2014) reported that SSR locus *xgwm282* is associated with QTL governing thousand kernel weight (TKW) corresponding to the onother study reported earlier by Huang *et al*, (2004)). Hai *et al.* (2008) found that QTL governing

thousand kernel weight (TKW) on chromosome 7BS was closely linked to the marker xgwm400, as described before by Borner *et al.* (2002). It was corroborated by the reports of Quarrie *et al.* (2006) that a QTL, corresponding to the one identified by Hai *et al*. (2008) on chromosome 7BS, was significantly associated with GPC and TKW.

Exploitation of heterosis is an important way to improve yield and quality for many crops. Hybrid rice and hybrid maize contributed to enhanced productivity with the discovery of male sterility system. Zhou *et al*. (2005) reported that xgwm131 is closely linked to a major fertility restoration gene, *Rf3*, against *Triticum timopheevii* cytoplasm, paving way for possibility of exploiting heterosis in *Triticum aestivum*.

Liu *et al*. (2002) reported that microsatellite marker locus *xgwm44* is linked to a Russian wheat aphid (RWA) resistance gene *Dn6* in coupling phase on the short arm of chromosome 7D at 14.6 cM distance. Zhang *et al*. (2008) reported that microsatellite marker xgwm37 locus is linked to a dominant wheat stripe rust resistance gene *YrY201* with a genetic distance of 5.8 cM.

Gene *H13* is a dominant gene with highly stable level of resistance against many Hessian fly biotypes. Using a population of 192 $F_{2:3}$ families, Liu *et al. (*2005) identified that the gene is located on the distal part of chromosome arm 6DS flanked by markers xcfd42 and xgdm141 that can be utilized for breeding Hessian fly resistance.

This information on the practical utility of microsatellite markers that are deployed in the present study helps to appreciate the practical importance of the genetic relatedness and population structure deciphered from a panel of 618 Indian wheat landraces.

CHAPTER 3

MATERIALS & METHODS

3. MATERIALS AND METHODS

In order to achieve the objectives of the current study, the experiments of the present investigation were carried out at the laboratories of ICAR-National Research Centre on Plant Biotechnology; and in the experimental fields of ICAR-Indian Agricultural Research Institute, PUSA Campus, New Delhi-110 012, India. The experiments were conducted during 2011-2014. The materials and methods used for conducting various experiments, in the present study, are described in the following sections.

3.1 Materials

3.1.1 Plant materials used for study of diversity, population structure and phylogeny

The plant material of the present study comprised of 618 Indian wheat landrace accessions (Table 3.1); kindly supplied by late Dr. S.K. Mishra, ICAR-National Bureau of Plant Genetic Resources (NBPGR), New Delhi, with details available in the wheat passport database accessible at the official website of ICAR-NBPGR (www.nbpgr.ernet.in). Among these accessions, 597 were bread wheat (hexaploid) and 21 were macaroni or durum (tetraploid) wheat. The selected accessions represented almost all the wheat growing regions of India (Figure 3.1). Geopolitical State from which wheat landraces were collected is given Table 3.2 for 108 accessions (89 hexaploids and 19 diploids). These accessions were grown at the experimental farm (Figure 3.2) of ICAR-Indian Agricultural Research Institute, New Delhi (latitude: 80°38'23"N, longitude:77°09'27"E, and with an elevation of 228.61m above msl) for two consecutive cropping seasons (2011-12 and 2012-13), following standard agronomic practices.

3.1.2 Wheat DNA used for study of diversity, population structure and phylogeny

The DNA was extracted from 618 Indian wheat accessions raised under laboratory conditions in germination trays (Figure 3.3) using the seeds that were stored at 4°C after harvesting from the plants raised under field conditions during the crop season: 2012-13. The germination trays, filled with sterile soilrite mix, were used to raise fresh and healthy seedlings. Approximately 2 gram of young leaf tissue was harvested from 15-20 days old seedlings and used for DNA extraction.

Table 3.1 List of 618 Indian wheat accessions identified by their indigenous collection (IC) numbers used in the present study.

S.No.	Accession	S.No.	Accession	S.No.	Accession	S.No.	Accession	S.No.	Accession
1	IC_10453	48	IC_35071	95	IC_53212	142	IC_53221	189	IC_53228
2	IC_10454	49	IC_47478	96	IC_53212	143	IC_53222	190	IC_53228
3	IC_10454	50	IC_53205	97	IC_53213	144	IC_53222	191	IC_53228
4	IC_10455	51	IC_53205	98	IC_53213	145	IC_53222	192	IC_53228
5	IC_10456	52	IC_53205	99	IC_53213	146	IC_53222	193	IC_53229
6	IC_10456	53	IC_53205	100	IC_53213	147	IC_53222	194	IC_53229
7	IC_10456	54	IC_53206	101	IC_53213	148	IC_53222	195	IC_53229
8	IC_10457	55	IC_53206	102	IC_53213	149	IC_53222	196	IC_53229
9	IC_10457	56	IC_53206	103	IC_53213	150	IC_53223	197	IC_53230
10	IC_11872	57	IC_53206	104	IC_53213	151	IC_53223	198	IC_53230
11	IC_11872	58	IC_53206	105	IC_53213	152	IC_53223	199	IC_53231
12	IC_11872	59	IC_53206	106	IC_53214	153	IC_53223	200	IC_53231
13	IC_11872	60	IC_53207	107	IC_53214	154	IC_53223	201	IC_53240
14	IC_11873	61	IC_53207	108	IC_53214	155	IC_53223	202	IC_53247
15	IC_11873	62	IC_53207	109	IC_53214	156	IC_53224	203	IC_53247
16	IC_11873	63	IC_53207	110	IC_53214	157	IC_53224	204	IC_53247
17	IC_11873	64	IC_53207	111	IC_53214	158	IC_53224	205	IC_53247
18	IC_11873	65	IC_53207	112	IC_53214	159	IC_53224	206	IC_53247
19	IC_11873	66	IC_53208	113	IC_53214	160	IC_53224	207	IC_53248
20	IC_11874	67	IC_53208	114	IC_53214	161	IC_53224	208	IC_53248
21	IC_11874	68	IC_53208	115	IC_53214	162	IC_53224	209	IC_53248
22	IC_11875	69	IC_53208	116	IC_53215	163	IC_53224	210	IC_53248
23	IC_11876	70	IC_53208	117	IC_53215	164	IC_53224	211	IC_53248
24	IC_13884	71	IC_53208	118	IC_53215	165	IC_53224	212	IC_53248
25	IC_13884	72	IC_53208	119	IC_53215	166	IC_53225	213	IC_53248
26	IC_13889	73	IC_53209	120	IC_53215	167	IC_53225	214	IC_53248
27	IC_13889	74	IC_53209	121	IC_53217	168	IC_53225	215	IC_53249
28	IC_13889	75	IC_53209	122	IC_53218	169	IC_53225	216	IC_53249
29	IC_13889	76	IC_53209	123	IC_53218	170	IC_53225	217	IC_53249
30	IC_13889	77	IC_53209	124	IC_53218	171	IC_53225	218	IC_53250
31	IC_21054	78	IC_53209	125	IC_53218	172	IC_53225	219	IC_53250
32	IC_21214	79	IC_53209	126	IC_53218	173	IC_53226	220	IC_53250
33	IC_21214	80	IC_53209	127	IC_53218	174	IC_53226	221	IC_53269
34	IC_21217	81	IC_53209	128	IC_53218	175	IC_53226	222	IC_53269
35	IC_21218	82	IC_53210	129	IC_53218	176	IC_53226	223	IC_53269
36	IC_28664	83	IC_53210	130	IC_53218	177	IC_53226	224	IC_53270
37	IC_28926	84	IC_53210	131	IC_53220	178	IC_53226	225	IC_53270
38	IC_28958	85	IC_53210	132	IC_53220	179	IC_53226	226	IC_53270
39	IC_28969	86	IC_53210	133	IC_53220	180	IC_53227	227	IC_53270
40	IC_28974	87	IC_53211	134	IC_53220	181	IC_53227	228	IC_53270
41	IC_29035	88	IC_53211	135	IC_53220	182	IC_53227	229	IC_53270
42	IC_30234	89	IC_53211	136	IC_53221	183	IC_53227	230	IC_53271
43	IC_30276	90	IC_53211	137	IC_53221	184	IC_53227	231	IC_53272
44	IC_31496	91	IC_53212	138	IC_53221	185	IC_53227	232	IC_53272
45	IC_32520	92	IC_53212	139	IC_53221	186	IC_53227	233	IC_53272
46	IC_35069	93	IC_53212	140	IC_53221	187	IC_53228	234	IC_53272
47	IC_35070	94	IC_53212	141	IC_53221	188	IC_53228	235	IC_53272

Contd…

Table 3.1 Continued…

S.No.	Accession	S.No.	Accession	S.No.	Accession	S.No.	Accession	S.No.	Accession
236	IC_532729	283	IC_533954	330	IC_534582	377	IC_534778	424	IC_59544
237	IC_532736	284	IC_533962	331	IC_534605	378	IC_534786	425	IC_59612
238	IC_532737	285	IC_533963	332	IC_534609	379	IC_534787	426	IC_64238
239	IC_532738	286	IC_533964	333	IC_534688	380	IC_534792	427	IC_73198
240	IC_532768	287	IC_533965	334	IC_534690	381	IC_534794	428	IC_73201
241	IC_532773	288	IC_533966	335	IC_534697	382	IC_534798	429	IC_73205
242	IC_532775	289	IC_533970	336	IC_534720	383	IC_534802	430	IC_73207
243	IC_532779	290	IC_533971	337	IC_534721	384	IC_534805	431	IC_73210
244	IC_532784	291	IC_534072	338	IC_534723	385	IC_534806	432	IC_73215
245	IC_532787	292	IC_534123	339	IC_534736	386	IC_534808	433	IC_73232
246	IC_532790	293	IC_534189	340	IC_534737	387	IC_534811	434	IC_73493
247	IC_532794	294	IC_534190	341	IC_534739	388	IC_534814	435	IC_75327
248	IC_532805	295	IC_534192	342	IC_534741	389	IC_534819	436	IC_75328
249	IC_532807	296	IC_534193	343	IC_534742	390	IC_534820	437	IC_75333
250	IC_532811	297	IC_534197	344	IC_534743	391	IC_534822	438	IC_75339
251	IC_532813	298	IC_534223	345	IC_534744	392	IC_534823	439	IC_75351
252	IC_532815	299	IC_534235	346	IC_534745	393	IC_534854	440	IC_78096
253	IC_532821	300	IC_534322	347	IC_534747	394	IC_534855	441	IC_78707
254	IC_532832	301	IC_534323	348	IC_534748	395	IC_534857	442	IC_78728
255	IC_532833	302	IC_534334	349	IC_534749	396	IC_534858	443	IC_78715
256	IC_532835	303	IC_534363	350	IC_534750	397	IC_534859	444	IC_78824
257	IC_532837	304	IC_534419	351	IC_534751	398	IC_534864	445	IC_78828
258	IC_532839	305	IC_534430	352	IC_534752	399	IC_534870	446	IC_78832
259	IC_532841	306	IC_534432	353	IC_534754	400	IC_534871	447	IC_78836
260	IC_532842	307	IC_534435	354	IC_534755	401	IC_534873	448	IC_78837
261	IC_532847	308	IC_534451	355	IC_534756	402	IC_534883	449	IC_78839
262	IC_532849	309	IC_534455	356	IC_534757	403	IC_534884	450	IC_78843
263	IC_532851	310	IC_534480	357	IC_534758	404	IC_534885	451	IC_78854
264	IC_532853	311	IC_534481	358	IC_534759	405	IC_534886	452	IC_78859
265	IC_532855	312	IC_534509	359	IC_534760	406	IC_534887	453	IC_78860
266	IC_532857	313	IC_534524	360	IC_534761	407	IC_55507	454	IC_78862
267	IC_532863	314	IC_534534	361	IC_534762	408	IC_55578	455	IC_78865
268	IC_532868	315	IC_534543	362	IC_534763	409	IC_55593	456	IC_78869
269	IC_532872	316	IC_534549	363	IC_534764	410	IC_55617	457	IC_78872
270	IC_532880	317	IC_534553	364	IC_534765	411	IC_55636	458	IC_78877
271	IC_532886	318	IC_534554	365	IC_534766	412	IC_55652	459	IC_78888
272	IC_532887	319	IC_534555	366	IC_534767	413	IC_55656	460	IC_78889
273	IC_532891	320	IC_534556	367	IC_534768	414	IC_55657	461	IC_78890
274	IC_532905	321	IC_534557	368	IC_534769	415	IC_55659	462	IC_78891
275	IC_532910	322	IC_534560	369	IC_534770	416	IC_55664	463	IC_78895
276	IC_532923	323	IC_534561	370	IC_534771	417	IC_55685	464	IC_78897
277	IC_532930	324	IC_534562	371	IC_534772	418	IC_57844	465	IC_78899
278	IC_532934	325	IC_534564	372	IC_534773	419	IC_57983	466	IC_78901
279	IC_532935	326	IC_534565	373	IC_534774	420	IC_57998	467	IC_78905
280	IC_532936	327	IC_534566	374	IC_534775	421	IC_59179	468	IC_78908
281	IC_532937	328	IC_534568	375	IC_534776	422	IC_59191	469	IC_78911
282	IC_533953	329	IC_534572	376	IC_534777	423	IC_59511	470	IC_78912

Contd…

Table 3.1 Continued…

S.No.	Accession	S.No.	Accession	S.No.	Accession	S.No.	Accession
471	IC_78914	518	IC_79063	565	IC_82179	612	IC_82421
472	IC_78915	519	IC_79065	566	IC_82180	613	IC_82425
473	IC_78918	520	IC_79066	567	IC_82181	614	IC_82426
474	IC_78919	521	IC_79067	568	IC_82185	615	IC_82432
475	IC_78920	522	IC_79068	569	IC_82187	616	IC_82440
476	IC_78923	523	IC_79077	570	IC_82189	617	IC_82221
477	IC_78924	524	IC_79079	571	IC_82190	618	IC_82555
478	IC_78925	525	IC_79080	572	IC_82192		
479	IC_78927	526	IC_79083	573	IC_82193		
480	IC_78928	527	IC_79085	574	IC_82195		
481	IC_78930	528	IC_79090	575	IC_82197		
482	IC_78931	529	IC_79091	576	IC_82198		
483	IC_78933	530	IC_79095	577	IC_82199		
484	IC_78935	531	IC_79100	578	IC_82200		
485	IC_78936	532	IC_79102	579	IC_82204		
486	IC_78937	533	IC_79106	580	IC_82206		
487	IC_78940	534	IC_79107	581	IC_82210		
488	IC_78941	535	IC_79108	582	IC_82236		
489	IC_78945	536	IC_79110	583	IC_82247		
490	IC_78948	537	IC_79940	584	IC_82256		
491	IC_78960	538	IC_82116	585	IC_82257		
492	IC_78968	539	IC_82118	586	IC_82259		
493	IC_78981	540	IC_82119	587	IC_82263		
494	IC_78982	541	IC_82125	588	IC_82285		
495	IC_78987	542	IC_82126	589	IC_82286		
496	IC_78990	543	IC_82127	590	IC_82303		
497	IC_78991	544	IC_82130	591	IC_82335		
498	IC_78996	545	IC_82132	592	IC_82338		
499	IC_78999	546	IC_82133	593	IC_82342		
500	IC_79022	547	IC_82134	594	IC_82344		
501	IC_79023	548	IC_82135	595	IC_82367		
502	IC_79026	549	IC_82136	596	IC_82369		
503	IC_79028	550	IC_82145	597	IC_82370		
504	IC_79038	551	IC_82146	598	IC_82371		
505	IC_79039	552	IC_82147	599	IC_82372		
506	IC_79040	553	IC_82156	600	IC_82375		
507	IC_79041	554	IC_82157	601	IC_82377		
508	IC_79042	555	IC_82158	602	IC_82381		
509	IC_79043	556	IC_82159	603	IC_82385		
510	IC_79046	557	IC_82160	604	IC_82386		
511	IC_79047	558	IC_82161	605	IC_82387		
512	IC_79050	559	IC_82163	606	IC_82388		
513	IC_79052	560	IC_82164	607	IC_82393		
514	IC_79053	561	IC_82165	608	IC_82394		
515	IC_79055	562	IC_82167	609	IC_82398		
516	IC_79056	563	IC_82169	610	IC_82400		
517	IC_79062	564	IC_82177	611	IC_82410		

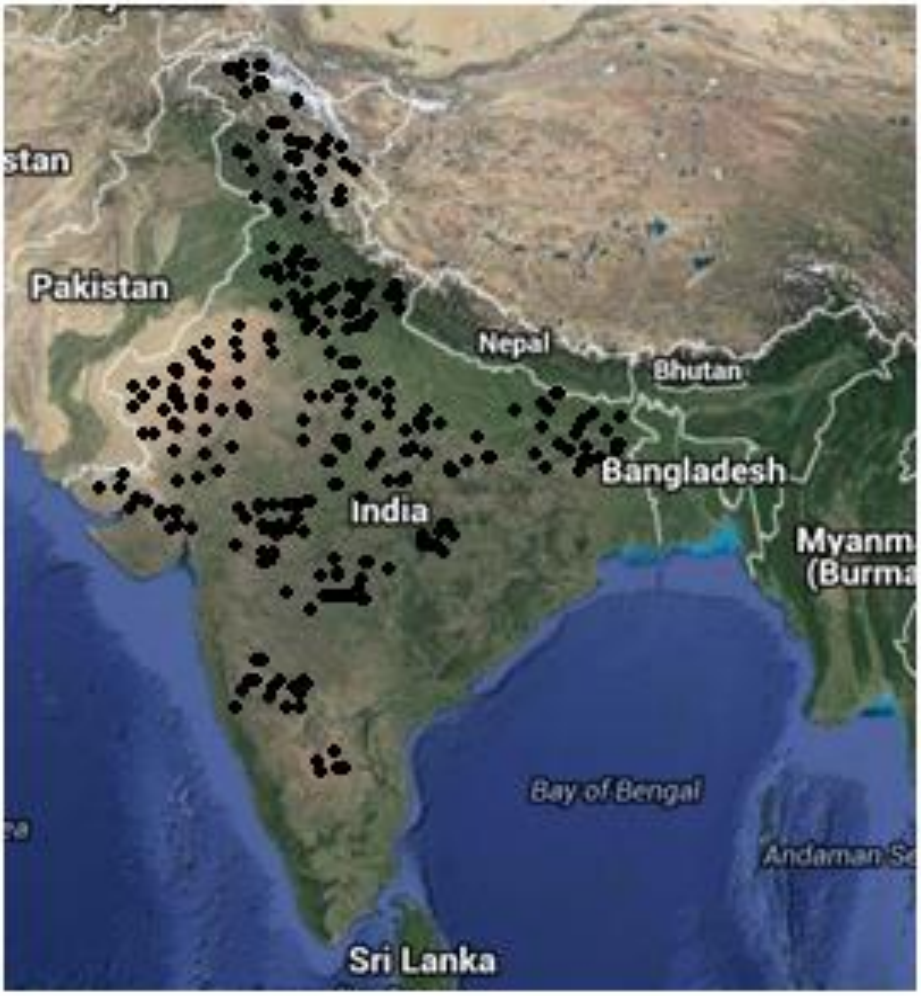


Figure 3.1. Satellite map (not to scale) showing the distribution of sample collection sites of wheat landraces in India. Black dots indicate approximate locations of geographical origins of wheat accessions used in the present study.

Table 3.2 List of 108 Indian wheat landraces with passport information obtained from online database of ICAR-NBPGR, New Delhi

S.No.	Accession No.	Crop	Scientific Name	Cultivar Name	Biological Status	Source (District)	Source (State)
1	IC_212179	Wild wheat	*Triticum durum*	Unknown	Wild	Un known	Gujarat
2	IC_210548	Wheat	*Triticum aestivum*	Unknown	Others	Un known	Himachal Pradesh
3	IC_78981	Wheat	*Triticum aestivum*	Kankoo	Wild	Un known	Himachal Pradesh
4	IC_78982	Wheat	*Triticum aestivum*	Wheat/Dharmauri	Wild	Un known	Himachal Pradesh
5	IC_78987	Wheat	*Triticum aestivum*	Wheat/Kankoo	Wild	Un known	Himachal Pradesh
6	IC_78990	Wheat	*Triticum aestivum*	Wheat/Kankoo	Wild	Un known	Himachal Pradesh
7	IC_78991	Wheat	*Triticum aestivum*	Wheat/Dharmauri	Wild	Un known	Himachal Pradesh
8	IC_78996	Wheat	*Triticum aestivum*	Wheat/Kankoo	Others	Un known	Himachal Pradesh
9	IC_78999	Wheat	*Triticum aestivum*	Wheat/Dharmauri	Wild	Un known	Himachal Pradesh
10	IC_79062	Wheat	*Triticum aestivum*	Wheat/Dharmauri	Wild	Un known	Himachal Pradesh
11	IC_79066	Wheat	*Triticum aestivum*	Unknown	Wild	Un known	Himachal Pradesh
12	IC_79077	Wheat	*Triticum aestivum*	Unknown	Wild	Un known	Himachal Pradesh
13	IC_79079	Wheat	*Triticum aestivum*	Kankoo	Wild	Un known	Himachal Pradesh
14	IC_79090	Wheat	*Triticum aestivum*	Wheat/Jau	Wild	Un known	Himachal Pradesh
15	IC_79091	Wheat	*Triticum aestivum*	Wheat/Dharmauri Kankoo(Mix)	Wild	Un known	Himachal Pradesh
16	IC_79095	Wheat	*Triticum aestivum*	Wheat/Kankoo	Wild	Un known	Himachal Pradesh
17	IC_82303	Wheat	*Triticum aestivum*	Kanak	Others	Un known	Himachal Pradesh

Contd…

Table 3.2 Contd…

S.No.	Accession No.	Crop	Scientific Name	Cultivar Name	Biological Status	Source (District)	Source (State)
18	IC_82385	Wheat	*Triticum aestivum*	Kanak wheat	Others	Un known	Himachal Pradesh
19	IC_82342	Wheat	*Triticum aestivum*	Wheat/Kankoo	Others	Un known	Himachal Pradesh
20	IC_118762	Wheat	*Triticum aestivum*	Unknown	Others	Un known	Karnataka
21	IC_118727	Wheat	*Triticum aestivum*	Unknown	Others	Un known	Maharashtra
22	IC_118729	Wheat	*Triticum aestivum*	Unknown	Others	Un known	Maharashtra
23	IC_118731	Wheat	*Triticum aestivum*	Unknown	Others	Un known	Maharashtra
24	IC_104537	Wheat	*Triticum aestivum*	Unknown	Others	Un known	Rajasthan
25	IC_104543	Wheat	*Triticum aestivum*	Unknown	Others	Un known	Rajasthan
26	IC_104551	Wheat	*Triticum aestivum*	Unknown	Others	Un known	Rajasthan
27	IC_212142	Wheat	*Triticum aestivum*	Unknown	Others	Un known	Rajasthan
28	IC_212145	Wheat	*Triticum aestivum*	Unknown	Others	Un known	Rajasthan
29	IC_212185	Wheat	*Triticum aestivum*	Unknown	Others	Un known	Rajasthan
30	IC_532098	Wheat	*Triticum aestivum*	Unknown	Others	Un known	Uttar Pradesh
31	IC_75327	Macaroni	*Triticum durum*	Unknown	Others	Un known	Uttar Pradesh
32	IC_75328	Macaroni	*Triticum durum*	Unknown	Others	Un known	Uttar Pradesh
33	IC_75333	Macaroni	*Triticum durum*	Unknown	Others	Un known	Uttar Pradesh
34	IC_75339	Wheat	*Triticum aestivum*	Unknown	Others	Un known	Uttar Pradesh
35	IC_75351	Macaroni	*Triticum durum*	Unknown	Others	Un known	Uttar Pradesh
36	IC_78096	Wild wheat	*Triticum aestivum*	PBW-120/-	Others	Un known	Uttar Pradesh

Contd..

Table 3.2 Contd…

S.No.	Accession No.	Crop	Scientific Name	Cultivar Name	Biological Status	Source (District)	Source (State)
37	IC_78707	Macaroni	*Triticum durum*	Anjali-2301/-	Cultivar	Un known	Uttar Pradesh
38	IC_78824	Wild wheat	*Triticum aestivum*	K-3049/-	Others	Un known	Uttar Pradesh
39	IC_78828	Wheat	*Triticum aestivum*	K-3039/-	Others	Un known	Uttar Pradesh
40	IC_78832	Wheat	*Triticum aestivum*	K-3115/-	Others	Un known	Uttar Pradesh
41	IC_78837	Wheat	*Triticum aestivum*	K-2374/-	Others	Un known	Uttar Pradesh
42	IC_78839	Wheat	*Triticum aestivum*	K-3082/-	Others	Un known	Uttar Pradesh
43	IC_78843	Wheat	*Triticum aestivum*	K-3041/-	Others	Un known	Uttar Pradesh
44	IC_78854	Wheat	*Triticum aestivum*	K-3175/-	Wild	Un known	Uttar Pradesh
45	IC_78859	Wheat	*Triticum aestivum*	K-3376/-	Wild	Un known	Uttar Pradesh
46	IC_78860	Wheat	*Triticum aestivum*	K-3055/-	Others	Un known	Uttar Pradesh
47	IC_78862	Macaroni	*Triticum durum*	K-2952/-	Cultivar	Un known	Uttar Pradesh
48	IC_78865	Wild wheat	*Triticum aestivum*	K-3079/-	Others	Un known	Uttar Pradesh
49	IC_78869	Wheat	*Triticum aestivum*	K-3122/-	Others	Un known	Uttar Pradesh
50	IC_78888	Macaroni	*Triticum durum*	K-3319/-	Cultivar	Un known	Uttar Pradesh
51	IC_78889	Wheat	*Triticum aestivum*	K-3077/-	Wild	Un known	Uttar Pradesh
52	IC_78899	Wheat	*Triticum aestivum*	K-3159/-	Others	Un known	Uttar Pradesh
53	IC_78901	Wheat	*Triticum aestivum*	K-3087/-	Wild	Un known	Uttar Pradesh
54	IC_78905	Wheat	*Triticum aestivum*	K-3093 /-	Wild	Un known	Uttar Pradesh
55	IC_78908	Wheat	*Triticum aestivum*	K-3066/-	Wild	Un known	Uttar Pradesh

Contd..

Table 3.2 Contd..

S.No.	Accession No.	Crop	Scientific Name	Cultivar Name	Biological Status	Source (District)	Source (State)
56	IC_78911	Wheat	*Triticum aestivum*	K-3206/-	Wild	Un known	Uttar Pradesh
57	IC_78912	Wheat	*Triticum aestivum*	K-3172/-	Wild	Un known	Uttar Pradesh
58	IC_78914	Wheat	*Triticum aestivum*	K-2995/-	Wild	Un known	Uttar Pradesh
59	IC_78915	Wheat	*Triticum aestivum*	K-3069/-	Wild	Un known	Uttar Pradesh
60	IC_78918	Wheat	*Triticum aestivum*	K-3028/-	Wild	Un known	Uttar Pradesh
61	IC_78919	Wheat	*Triticum aestivum*	K-3142/-	Wild	Un known	Uttar Pradesh
62	IC_78920	Wheat	*Triticum aestivum*	K-2944/-	Wild	Un known	Uttar Pradesh
63	IC_78935	Macaroni	*Triticum durum*	K-2929/-	Cultivar	Un known	Uttar Pradesh
64	IC_78936	Wheat	*Triticum aestivum*	K-2890/-	Wild	Un known	Uttar Pradesh
65	IC_78937	Wheat	*Triticum aestivum*	K-3003/-	Wild	Un known	Uttar Pradesh
66	IC_78945	Wheat	*Triticum aestivum*	K-3011/-	Wild	Un known	Uttar Pradesh
67	IC_78948	Wheat	*Triticum aestivum*	K-3106/-	Wild	Un known	Uttar Pradesh
68	IC_78960	Wheat	*Triticum aestivum*	K-3221/-	Wild	Un known	Uttar Pradesh
69	IC_78968	Wheat	*Triticum aestivum*	K-3124/-	Wild	Un known	Uttar Pradesh
70	IC_79067	Wheat	*Triticum aestivum*	Unknown	Wild	Un known	Uttar Pradesh
71	IC_79080	Wheat	*Triticum aestivum*	Unknown	Wild	Un known	Uttar Pradesh
72	IC_79083	Wheat	*Triticum aestivum*	Unknown	Wild	Un known	Uttar Pradesh
73	IC_79085	Wheat	*Triticum aestivum*	Unknown	Wild	Un known	Uttar Pradesh
74	IC_28664	Wheat	*Triticum aestivum*	Unknown	Others	Patan	Gujarat

Contd…

Table 3.2 Contd..

S.No.	Accession No.	Crop	Scientific Name	Cultivar Name	Biological Status	Source (District)	Source (State)
75	IC_30284	Wheat	*Triticum aestivum*	Unknown	Others	Leh	Jammu & Kashmir
76	IC_30276	Wheat	*Triticum aestivum*	Unknown	Others	Leh	Jammu & Kashmir
77	IC_57998	Macaroni	*Triticum durum*	Unknown	Others	Leh	Jammu & Kashmir
78	IC_35069	Macaroni	*Triticum durum*	Unknown	Others	Dharwad	Karnataka
79	IC_35071	Macaroni	*Triticum durum*	Unknown	Others	Haveri	Karnataka
80	IC_31496	Wheat	*Triticum aestivum*	Unknown	Others	Khargone	Madhya Pradesh
81	IC_32520	Macaroni	*Triticum durum*	Unknown	Others	Guna	Madhya Pradesh
82	IC_73198	Wild wheat	*Triticum durum*	Gulab	Cultivar	Nashik	Maharashtra
83	IC_73201	Wild wheat	*Triticum durum*	Jay/	Cultivar	Nashik	Maharashtra
84	IC_73205	Wild wheat	*Triticum durum*	Boxi. 288-18	Cultivar	Nashik	Maharashtra
85	IC_73207	Wild wheat	*Triticum durum*	Vijay	Cultivar	Nashik	Maharashtra
86	IC_73210	Wild wheat	*Triticum aestivum*	MotiI(HD-1949)	cultivar	Nashik	Maharashtra
87	IC_73215	Wheat	*Triticum aestivum*	N-59/-	Others	Nashik	Maharashtra
88	IC_28926	Wheat	*Triticum aestivum*	Unknown	Others	Barmer	Rajasthan
89	IC_28969	Wheat	*Triticum aestivum*	Unknown	Others	Pali	Rajasthan
90	IC_28974	Wheat	*Triticum aestivum*	Unknown	Others	Pali	Rajasthan
91	IC_29035	Macaroni	*Triticum durum*	Unknown	Others	Tonk	Rajasthan
92	IC_73493	Macaroni	*Triticum durum*	Unknown	Others	Sri Ganganagar	Rajasthan
93	IC_82126	Wheat	*Triticum aestivum*	Unknown	Cultivar	Jalor	Rajasthan

Contd…

Table 3.2 Contd..

S.No.	Accession No.	Crop	Scientific Name	Cultivar Name	Biological Status	Source (District)	Source (State)
94	IC_82127	Wheat	*Triticum aestivum*	Unknown	Cultivar	Jalor	Rajasthan
95	IC_82133	Wheat	*Triticum aestivum*	Unknown	Cultivar	Barmer	Rajasthan
96	IC_82134	Wheat	*Triticum aestivum*	Unknown	Cultivar	Jalor	Rajasthan
97	IC_82135	Wheat	*Triticum aestivum*	Unknown	Cultivar	Jalor	Rajasthan
98	IC_82136	Wheat	*Triticum aestivum*	Unknown	Cultivar	Jalor	Rajasthan
99	IC_82145	Wheat	*Triticum aestivum*	Unknown	Cultivar	Pali	Rajasthan
100	IC_82146	Wheat	*Triticum aestivum*	Unknown	Cultivar	Pali	Rajasthan
101	IC_82147	Wheat	*Triticum aestivum*	Unknown	Cultivar	Rajsamand	Rajasthan
102	IC_82156	Wheat	*Triticum aestivum*	Unknown	Cultivar	Udaipur	Rajasthan
103	IC_82157	Wheat	*Triticum aestivum*	Unknown	Cultivar	Udaipur	Rajasthan
104	IC_82158	Wheat	*Triticum aestivum*	Unknown	Cultivar	Udaipur	Rajasthan
105	IC_82159	Wheat	*Triticum aestivum*	Unknown	Cultivar	Udaipur	Rajasthan
106	IC_82160	Wheat	*Triticum aestivum*	Unknown	Cultivar	Udaipur	Rajasthan
107	IC_82161	Wheat	*Triticum aestivum*	Unknown	Cultivar	Udaipur	Rajasthan
108	IC_82180	Wheat	*Triticum aestivum*	Unknown	Others	Udaipur	Rajasthan

Figure 3.2 An aerial view of a part of experimental plot used for raising Indian wheat accessions in field condition. Genetic variability for morphological and phenological traits including days to maturity, plant height, growth habits, etc., are illustrated here.

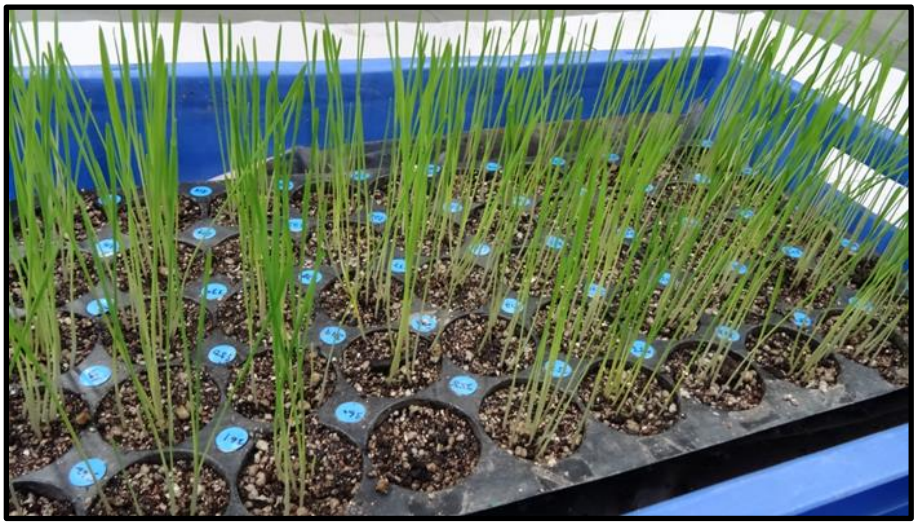


Figure 3.3 Fifteen-days old seedlings raised under laboratory condition for DNA isolation.

3.1.3 Chemicals and reagents used for DNA extraction and gel electrophoresis

1 M Tris-Cl (pH 8.0)

121 g Tris-Cl base was dissolved in 800 ml double distilled H_2O, and pH was adjusted using concentrated HCl to 8.0. Volume was brought up to 1000 ml and solution was autoclaved.

0.5 M EDTA (pH 8.0)

186 g di-sodium salt of EDTA was dissolved in 800 ml double distilled H_2O, and pH was adjusted to 8.0 using NaOH pellets. Volume was brought up to 1000 ml and solution was autoclaved.

5 M NaCl

292.2 g of NaCl was dissolved in 800 ml double distilled H_2O. Volume was adjusted up to 1000 ml, and the solution was sterilized by autoclaving.

DNA extraction buffer (1000 ml)

1M Tris-Cl (pH 8.0)	100.0 ml	(100 mM)
0.5 M EDTA (pH 8.0)	40.0 ml	(20 mM)
5 M NaCl	280.0 ml	(1.4 M)
CTAB	20 g	(2%)
β-Mercaptoethanol	2000 µl	(0.2%).

7.5 M Ammonium acetate (pH 5.5)

578.1 g ammonium acetate was dissolved in 800 ml of deionized or distilled water. The final volume was made up to 1000 ml and the solution was filter-sterilized using 0.2µm filter. The pH was adjusted to 5.5 using glacial acetic acid.

3 M Sodium acetate (pH 5.2)

408 g sodium acetate was dissolved in 800 ml of deionized or distilled water and the pH was adjusted to 5.2 by adding glacial acetic acid and monitoring with a pH meter. Water was added to bring the total volume of solution to 1000 ml.

RNase A (10 mg/ml)

10 mg RNase A was dissolved in 1 ml autoclaved double distilled H_2O, vortexed well and boiled in water bath at 100^oC for 10-15 minutes, cooled and stored at -20 ^oC.

Phenol Chloroform Isoamyl alcohol (PCI) solution

Ready-made Phenol-Chlorophorm-Isoamyl alcohol (25:24:1) v/v was obtained from Amresco company.

Chloroform: Isoamyl alcohol

Ready-made Chloroform-Isoamyl alcohol (24:1) v/v was obtained from Amresco ompany.

Tris-EDTA buffer 10:1(TE)

Tris (10 mM): EDTA (1 mM) buffer was prepared in 100 ml as follows:

1 mM Tris-Cl, pH 8.0	1 ml
0.5 M EDTA pH 8.0	0.2 ml
Autoclaved double distilled H_2O	98.8 ml

Ethidium bromide (EtBr) Stock solution

Prepared as 10 mg/ml concentration by dissolving 1 g ethidium bromide in 100 ml H_2O by stirring for several hours; stored in a dark bottle at room temperature.

50 x TBE buffer

242 g of Tris base, 57.1 ml glacial acetic acid, and 100 ml of 0.5 M EDTA, (pH 8.0), were dissolved in 800 ml H_2O and made the volume up to 1000 ml.

5 x TBE buffer

54 g of Tris base, 27.5 g of boric acid, and 20 ml of 0.5 M EDTA, (pH 8.0), were dissolved in 800 ml H_2O, volume was made up to 1000 ml.

3.1.4 DNA markers used for study of diversity, population structure and phylogeny

Genome-wide distributed 42 microsatellite (SSR) markers, covering one each on short- and long-arm of each of the 21 homoeologous chromosomes, were selected based

on the bin map reported by Somers *et al*. (2004) and the primer sequences were obtained from the GrainGenes database (https://wheat.pw.usda.gov/GG3/). Initially 126 SSR markers, three each on the short- and long arm of each of twenty-one homoeologous covering A, B and D genomes were assessed for single locus amplification and polymorphism. The markers which showed at least two band sizes (alleles), consistently in two or more among randomly sampled 20 accessions from a set of 618 Indian wheat landrace accessions, were considered for further study. Finally, 42 polymorphic SSR markers with genome wide coverage, one each on the short- and long arm of each of twenty-one homoeologous chromosomes covering A, B and D genomes were fluorescence labelled and deployed for molecular diversity analysis (Table 3.3, Table 3.4 and Table 3.5).

3.2. Methods

3.2.1 Genomic DNA extraction

Leaf tissue for DNA isolation was collected from 15-20 days old seedlings grown under laboratory conditions in 96-well germination trays using sterile soilrite media, and approximately 2 gram of young leaf tissue was harvested from 15-20 days old seedlings and was used for DNA extraction. DNA was isolated using modified CTAB method (Murray & Thompson, 1980; Doyle and Doyle, 1990) with few modifications. About 2 gm young and tender leaves from fresh, healthy and young (15-20 days old) seedlings was collected and ground to a fine powder with liquid nitrogen using a pre-cooled mortar and pestle. About 120 mg of PVP (Polyvinyl pyrrolidine) (2%) was added during the process to avoid phenol formation. The powder, without being allowed to thaw, was transferred to 50 ml polypropylene tubes containing 25 ml of pre-warmed (65℃) DNA extraction buffer (2% w/v CTAB, 1.4M NaCl, 20mM EDTA, 100mM Tris-Cl, pH8.0 and 0.2% of freshly added β-mercaptoethanol). Suspension was incubated at 65℃, in a water-bath, for 90 minutes with intermittent shaking at 10-minute interval. The tubes were allowed to cool to room temperature. The contents of the tubes were emulsified with an equal volume (10 ml) of chloroform-isoamyl alcohol (24:1) for 5 min by gentle inversion and the content was then subjected to differential centrifugation: 3000 rpm (Heraeus Megafuge 16R, Thermo Scientific, USA) for two minutes followed by 10000 rpm for 8 minutes at room temperature. Following centrifugation, the separated aqueous phase was removed with a

Table 3.3 Details of the primers covering subgenome-A and fluorescent dyes used for labelling them.

S. No.	Primer Name	Location	Fluorescent dye /colour	Repeat unit	Primer nucleotide sequence		T_a in ^oC
1	Xgwm136	1 AS	6-FAM (Blue)	$(CT)_{58}$	Forward:	5' GACAGCACCTTGCCCTTTG 3'	60
					Reverse	5' CATCGGCAACATGCTCATC 3'	
2	Xgwm497	1 AL	6-FAM (Blue)	$(GT)_{29}$ imp	Forward:	5' GTAGTGAAGACAAGGGCATT 3'	55
					Reverse	5' CCGAAAGTTGGGTGATATAC 3'	
3	Xgwm448	2 AS	6-FAM (Blue)	$(GA)_{29}$	Forward:	5' AAACCATATTGGGAGGAAAGG 3'	60
					Reverse	5' CACATGGCATCACATTTGTG 3'	
4	Xgwm445	2 AL	6-FAM (Blue)	$(CT)_{19}$	Forward:	5' TTTGTTGGGGGTTAGGATTAG 3'	55
					Reverse	5' CCTTAACACTTGCTGGTAGTGA 3'	
5	Xgwm2	3 AS	VIC (Green)	$(CA)_{18}$	Forward:	5' CTGCAAGCCTGTGATCAACT 3'	50
					Reverse	5' CATTCTCAAATGATCGAACA 3'	
6	Xgwm494	3 AL	VIC (Green)	$(CA)_{13}$	Forward:	5' ATTGAACAGGAAGACATCAGGG 3'	60
					Reverse	5' TGCTGCTGGTCTCTGTGC 3'	
7	Xgwm165	4 AS	VIC (Green)	$(GA)_{20}$	Forward:	5' TGCAGTGGTCAGATGTTTCC 3'	60
					Reverse	5' CTTTTCTTTCAGATTGCGCC 3'	
8	Xgwm637	4 AL	PET (Red)	$(CA)_{18}$	Forward:	5' AAAGAGGTCTGCCGCTAACA 3'	60
					Reverse	5' TATACGGTTTTGTGAGGGGG 3'	
9	Xgwm443	5 AS	PET (Red)	$(CA)_{20}(GA)_{22}$	Forward:	5' GGGTCTTCATCCGGAACTCT 3'	50
					Reverse	5' CCATGATTTATAAATTCCACC 3'	

Note: T_a: Annealing temperature

Contd…

Table 3.3 Contd…

S. No.	Primer Name	Location	Fluorescent dye /colour	Repeat unit	Primer nucleotide sequence		T_a in ^oC
10	Xgwm666	5 AL	PET (Red)	$(CA)_{13}$	Forward:	5' GCACCCACATCTTCGACC 3'	60
					Reverse	5' TGCTGCTGGTCTCTGTGC 3'	
11	Xgwm459	6 AS	NED (Yellow)	$GA)_{>28}$	Forward:	5' ATGGAGTGGTCACACTTTGAA 3'	
					Reverse	5' AGCTTCTCTGACCAACTTCTCG 3'	55
12	Xgwm169	6 AL	NED (Yellow)	$(GA)_{23}$	Forward:	5' ACCACTGCAGAGAACACATACG 3'	
					Reverse	5' GTGCTCTGCTCTAAGTGTGGG 3'	60
13	Xgwm635	7AS	NED (Yellow)	$(CA)_{10}(GA)_{14}$	Forward:	5' TTCCTCACTGTAAGGGCGTT 3'	
					Reverse	5' CAGCCTTAGCCTTGGCG 3'	60
14	Xgwm282	7 AL	NED (Yellow)	$(GA)_{38}$	Forward:	5' TTGGCCGTGTAAGGCAG 3'	
					Reverse	5' TCTCATTCACACACAACACTAGC 3'	55

Note: T_a: Annealing temperature

Table 3.4 Details of the primers covering subgenome-B and fluorescent dyes used for labelling them.

S. No.	Primer Name	Location	Fluorescent dye (colour)	Repeat unit	Primer nucleotide sequences of the marker		T_a in ^oC
1	Xgwm550	1 BS	PET (Red)	$(CT)_8(GT)_{18}$	Forward:	5' CCCACAAGAACCTTTGAAGA 3'	55
					Reverse	5' CATTGTGTGTGCAAGGCAC 3'	
2	Xgwm268	1 BL	6-FAM (Blue)	$(GA)_{17}TA(GA)_{27}$	Forward:	5' AGGGGATATGTTGTCACTCCA 3'	55
					Reverse	5' TTATGTGATTGCGTACGTACCC 3'	
3	Xgwm257	2 BS	6-FAM (Blue)	$(GT)_{30}$	Forward:	5' AGAGTGCATGGTGGGACG 3'	60
					Reverse	5' CCAAGACGATGCTGAAGTCA 3'	
4	Xgwm47	2 BL	6-FAM (Blue)	$(CT)_7TT(CT)_{16}$	Forward:	5' TTGCTACCATGCATGACCAT 3'	60
					Reverse	5' TTCACCTCGATTGAGGTCCT 3'	
5	Xgwm533	3 BS	VIC (Green)	$(CT)_{18}(CA)_{20}$	Forward:	5' AAGGCGAATCAAACGGAATA 3'	60
					Reverse	5' GTTGCTTTAGGGGAAAAGCC 3'	
6	Xgwm181	3 BL	VIC (Green)	$(GA)_{28}$	Forward:	5' TCATTGGTAATGAGGAGAGA 3'	50
					Reverse	5' GAACCATTCATGTGCATGTC 3'	
7	Xgwm368	4 BS	PET (Red)	$(AT)_{25}$	Forward:	5' CCATTTCACCTAATGCCTGC 3'	60
					Reverse	5' AATAAAACCATGAGCTCACTTGC 3'	
8	Xgwm513	4 BL	VIC (Green)	$(CA)_{12}$	Forward:	5' ATCCGTAGCACCTACTGGTCA 3'	60
					Reverse	5' GGTCTGTTCATGCCACATTG 3'	

Note: T_a: Annealing temperature

Contd…

Table 3.4 Contd…

S. No.	Primer Name	Location	Fluorescent dye (colour)	Repeat unit	Primer nucleotide sequences of the marker		T_a in ^oC
9	Xgwm544	5 BS	PET (Red)	$(CT)_{12}(ATCT)_5$ $(CT)_{16}$	Forward:	5' TAGAATTCTTTATGGGGTCTGC 3'	55
					Reverse	5' AGGATTCCAATCCTTCAAAATT 3'	
10	Xgwm335	5 BL	PET (Red)	$(GA)_{14}(GCGT)_3$	Forward:	5' CGTACTCCACTCCACACGG 3'	55
					Reverse	5' CGGTCCAAGTGCTACCTTTC 3'	
11	Xgwm361	6 BS	VIC (Green)	$(GA)_{20}$imp	Forward:	5' GTAACTTGTTGCCAAAGGGG 3'	
					Reverse	5' ACAAAGTGGCAAAAGGAGACA 3'	60
12	Xgwm219	6 BL	NED (Yellow)	$(GA)_{35}$imp	Forward:	5' GATGAGCGACACCTAGCCTC 3'	
					Reverse	5' GGGGTCCGAGTCCACAAC 3'	60
13	Xgwm400	7 BS	NED (Yellow)	$(CA)_{21}$	Forward:	5' GTGCTGCCACCACTTGC 3'	
					Reverse	5' TGTAGGCACTGCTTGGGAG 3'	60
14	Xgwm131	7 BL	NED (Yellow)	$(CT)_{22}$	Forward:	5' AATCCCCACCGATTCTTCTC 3'	
					Reverse	5' AGTTCGTGGGTCTCTGATGG 3'	60

Note: T_a: Annealing temperature

Table 3.5 Details of the primers covering subgenome-D and fluorescent dyes used for labelling them.

S. No.	Primer Name	Location	Fluorescent dye /colour	Repeat unit		Primer nucleotide sequence	T_a in ^oC
1	Xgwm458	1 DS	NED (Yellow)	$(CA)_{13}$	Forward:	5' AATGGCAATTGGAAGACATAGC 3'	60
					Reverse	5' TTCGCAATGTTGATTTGGC 3'	
2	Xgwm642	1 DL	6-FAM (Blue)	$(GT)_{14}$	Forward:	5' ACGGCGAGAAGGTGCTC 3'	60
					Reverse	5' CATGAAAGGCAAGTTCGTCA 3'	
3	Xgwm484	2 DS	VIC (Green)	$(CT)_{29}$	Forward:	5' ACATCGCTCTTCACAAACCC 3'	55
					Reverse	5' AGTTCCGGTCATGGCTAGG 3'	
4	Xgwm349	2 DL	VIC (Green)	$(GA)_{34}$	Forward:	5' GGCTTCCAGAAAACAACAGG 3'	55
					Reverse	5' ATCGGTGCGTACCATCCTAC 3'	
5	Xgwm183	3 DS	VIC (Green)	$(GA)_{21}(N)_{51}(C)_{25}$	Forward:	5' GTCTTCCCATCTCGCAAGAG 3'	55
					Reverse	5' CTCGACTCCCATGTGGATG 3'	
6	Xgwm3	3 DL	VIC (Green)	$(CA)_{18}$	Forward:	5' GCAGCGGCACTGGTACATTT 3'	55
					Reverse	5' AATATCGCATCACTATCCCA 3'	
7	Xgwm194	4 DS	PET (Red)	$(CT)_{32}$imp	Forward:	5' GATCTGCTCTACTCTCCTCC 3'	50
					Reverse	5' CGACGCAGAACTTAAACAAG 3'	
8	Xgwm624	4 DL	PET (Red)	$(GT)_{26}$	Forward:	5' TTGATATTAAATCTCTCTATGTG 3'	50
					Reverse	5' AATTTTATTTGAGCTATGCG 3'	

Note: T_a: Annealing temperature

Contd…

Table 3.5 Contd…

S. No.	Primer Name	Location	Fluorescent dye /colour	Repeat unit	Primer nucleotide sequence		T_a in ^oC
9	Xgwm190	5 DS	PET (Red)	$(CT)_{22}$	Forward:	5' GTGCTTGCTGAGCTATGAGTC 3'	
					Reverse	5' GTGCCACGTGGTACCTTTG 3'	60
10	Xgwm182	5 DL	PET (Red)	$(CT)_{18}$	Forward:	5' TGATGTAGTGAGCCCATAGGC 3'	
					Reverse:	5' TTGCACACAGCCAAATAAGG 3'	60
11	Xgdm141	6 DS	NED (Yellow)	$(CT)_{22}$	Forward:	5' ATGGAGACCATGGACCAGAG 3'	
					Reverse	5' GGCGGTGTTCCTATGCC 3'	60
12	Xbarc1121	6 DL	NED (Yellow)	$(ATT)_{22}$	Forward:	5' GCGAGCAAACTGATCCCAAAAAG 3'	
					Reverse	5' TATCGGTGAGTACGCCAAAAACA 3'	52
13	Xgwm44	7 DS	6-FAM (Blue)	$(GA)_{28}$	Forward:	5' GTTGAGCTTTTCAGTTCGGC 3'	
					Reverse	5' ACTGGCATCCACTGAGCTG 3'	60
14	Xgwm37	7 DL	6-FAM (Blue)	$(AG)_8GG(AG)_{21}$	Forward:	5' ACTTCATTGTTGATCTTGCATG 3'	
					Reverse	5' CGACGAATTCCCAGCTAAAC 3'	60

Note: T_a: Annealing temperature

wide-bore pipette to a clean oak-ridge tube and 0.6 volume of chilled isopropanol (isopropyl alcohol) and 0.1 volume of ammonium acetate were added to the content in tube, followed by quick and gentle inversion, and incubated at 4˚ C for overnight, or, alternatively, at -20^{0}C for 30 min, for coagulation of DNA. If clump of precipitate appears, the DNA was spooled using a bent Pasteur pipette and transferred into 2.5 ml micro-centrifuge tube and centrifuged at 10000 rpm for 5 minutes at room temperature and the supernatant was discarded; otherwise, the supernatant was subjected to centrifugation at 10000 rpm for 10 minutes and the supernatant was discarded carefully: taking care not to dislodge the pellet. Further, the DNA pellet was washed twice with 70% ethanol by brief centrifugation, air dried, and dissolved in appropriate volume of TE buffer (100 mM Tris and 50 mM EDTA, pH 8.0), depending on the DNA pellet size. Samples were stored at 4^{0}C for complete dissolution.

3.2.1.1 Purification of DNA

Major contaminants of crude DNA preparation are RNA, protein and polysaccharides. Inclusion of CTAB in DNA extraction buffer helps in elimination of polysaccharides to a large extent. The RNA is removed by treating the sample with DNase free RNase (10 mg/ml) denatured at 70^{0}C. Proteins including RNase can be removed by treatment with phenol: chloroform: isoamylalcohol (25:24:1). For purification, crude DNA solution was transferred to 2 ml microcentrifuge tubes and the following protocol was used: (1) RNaseA @ 5µl/gm of leaf tissue was added to each DNA sample followed by incubation at 37^{0}C for one hour; (2) after RNase treatment the samples were purified by adding equal volume of phenol: chloroform: isoamyl alcohol (25:24:1) followed by mixing gently by inverting for about 5 minutes and centrifuged for 10 min at 10000 rpm; (3) the aqueous phase was transferred to 1.5ml eppendorf and Equal volume of chloroform-isoamyl alcohol (24:1) was added to it, mixed gently for 5 minutes and centrifuge at 10000 rpm for 5 minutes; (4) the aqueous phase was transferred to fresh eppendorf tube and 0.1 volume of 3M sodium acetate (pH 5.2) and 2 volume of chilled absolute ethanol were added and mixed properly followed by incubation at -20^{0}C for 30 minutes and centrifugation at 10000 rpm for 5 minutes; (5) the supernatant was discarded and the DNA pellet was washed for 5 min with 70% ethanol twice, air dried properly till even traces of

ethanol were removed completely and DNA pellet was dissolved finally in appropriate volume of TE buffer depending on DNA pellet size.

3.2.1.2 Quality assessment and quantification of genomic DNA

DNA was initially quantified using gel quantification method followed by nanodrop spectrophotometer and finally precise concentration was determined using varioscan (Varioscan Flash, Thermo Fisher Scientific, USA).

For quantification, 1µl of DNA samples, along with known quantity of λ uncut DNA (100ng, 250 ng) were loaded on 0.8% agarose gel. The electrophoresis was carried out at 70 volts for 30 min. The concentration of the extracted DNA was estimated by visual comparison of intensity of fluorescence in the band of sample DNA with that of λ-bacteriophage DNA. DNA was stained with ethidium bromide and observed under UV. The sample DNA was diluted with TE buffer such that the final concentration of DNA was approximately 30ng/µl. For further quantification, a spectrophotometer (Nanodrop, Thermo Scientific, USA) was also used. Purity was assessed based on ratio between absorbance at 230/260 nm wavelength and the acceptable absorbance ratio (A260/A280) is 1.8; while DNA concentration was calculated assuming that absorbance value of 1 (O.D.) at 260nm corresponds to 50ng/ml of double stranded DNA. According to the concentration of DNA, samples were diluted, and again loaded on 0.8% gel till all the samples finally reached around 50 ng/µl in a uniform manner. DNA samples were stored in in 96 well deep-well plates in duplicates: one set was kept at -80^{o}C for long term storage and the working set was stored at -20^{o}C.

3.2.2 Amplification of SSR marker loci using polymerase chain reaction (PCR)

3.2.2.1 Confirmatory amplification

The specified primers, described in Table 3.3, Table 3.4 and Table 3.5, were diluted to an initial concentration of 1 nm/µl and further diluted to working concentration of 5 pm/µl. For validation and optimization of PCR condition, the initial PCR amplification was carried out in 25 µl reaction volume containing 10mM Tris- HCl (pH 8.0), 50 mM KCl, 2.0 mM $MgCl_2$, 100 µM of each dNTP (MBI Fermentas, Germany), 1.0 unit Taq DNA

polymerase (Bangalore GeneI Pvt Ltd, India), 5 pm of each primer, and 30-50 ng of genomic DNA.

3.2.2.2 Final amplification

For high resolution genotyping of SSR maker loci using DNA analyser (ABI, USA), the final PCR reactions were carried out in 10µl reaction mixture containing 50ng/µl of genomic DNA, 0.25 µl of 50mM $MgCl_2$, 0.25 µl of 10mM dNTPs, 1U of Taq polymerase (0.25 µl from 5U/ µl) and 0.5 µl of 5 mM primer in 1X reaction buffer. The amplification reactions were performed in the Eppendorf Master cycler with an initial denaturation for 3 minutes at 94ºC, then 35 cycles: 1minute denaturation at 94ºC; 1 minute annealing at 52-60ºC (primer dependent) (Table 3.3, Table 3.4 and Table 3.5); 2 minute extension at 72ºC; and, final extension was carried out at 72ºC for 5 minutes and products were stored at 4˚C until electrophoresis.

Amplification reactions were performed using thermocycler PTC-225 (DNA Engine Tetrad, Gradient Cycler, MJ Research, Biorad, USA) with the following thermal profile: initial denaturation at 94ºC for 4 min, followed by 35 cycles at 94ºC for 1 min (denaturation), 60ºC for 1 min (primer annealing) and 72ºC for 2 min (primer extension) with a final extension at 72ºC for 10 min. PCR thermal profile used for amplifying SSR loci is given below:

Steps	Temperature (º C)	Duration (min.)	Cycles	Activity
1	94	4	1	Denaturation
2	94	1	↑	Denaturation
3	52-60 (primer specific)	1	40	Annealing
4	72	2	↓	Polymerization
5	72	10	1	Polymerization
6	4	∞	1	Storage

3.2.2.3 Electrophoretic separation of amplicons according to size

The PCR products of SSR markers were first separated in 4% metaphor agarose gel using 1x TBE buffer. Metaphor agarose was added to 1X TBE buffer and boiled using microwave oven to dissolve it in the buffer. It was cooled to about 50º C by swirling the flask and ethidium bromide (10 mg/ml) was added and mixed properly before pouring on to the gel plate fixed in the gel caster carrying combs fixed to the provided-notch such that

it gets inserted into the formed gel creating wells for loading the PCR products. Gel was allowed to solidify for about an hour, combs were removed carefully and gel was kept submerged in an electrophoresis buffer tank containing 1X TBE buffer. To each PCR product 6X gel loading dye was added, mixed well and loaded in the wells of the gel. 100-bp DNA ladder (MBI, Fermentas, Germany) was used as molecular weight marker. The gel was run for at least 3 ½ h at 100 volts using power pack 1000 (BIO-RAD, USA). The gel was visualized under UV trans-illuminator in a gel documentation system (Flurochem Gel Documentation System).

3.2.2.4 Visualization of PCR products

Gels were visualized under UV and photographed by using gel documentation system (Molecular Image™, Gel Doc™ XR+ with ImageLlab™ software, Biorad, USA).

3.2.3 PCR amplification of fluorescent labelled SSRs

3.2.3.1 Fluorescent labeling of SSRs

In order to achieve high resolution in genotyping SSR locus, alleles of SSR loci were resolved using automated ABI capillary electrophoresis. To facilitate multiplexing of PCR products and detection of locus-specific signals in automated capillary electrophoresis, 5-prime-end of forward primer, of each of the forty-two SSR markers was directly labelled with one of the four dyes each carrying one of the four different coloured fluorochromes, namely, FAM, VIC, NED and PET (Applied Biosystems, USA) provided in Table 3.3, Table 3.4, Table 3.5, and Appendix A1.

3.2.3.2 Amplification and confirmation of polymorphism on agarose gel

The PCR amplification was carried out in 10µl reaction mixtures that contained 10 g of genomic DNA, 1X PCR master mix (AmpliTaq Gold®, Applied Biosystems, USA), 0.1 µl (5pmol/µl) of forward primer (labeled with either 6- FAM, VIC, NED, or PET), and 0.3µl (5pmol/µl) reverse primer). PCR amplification reactions were performed using thermocycler PTC-225 (DNA Engine Tetrad, Gradient Cycler, MJ Research, Biorad, USA) with thermal profile given in Table 3.5. After confirmation of PCR amplification on agarose gel, PCR amplicons were separated using capillary electrophoresis on DNA analyser (ABI 3730xl, USA).

3.2.4 Automated capillary electrophoresis (ABI 3730xl) and allele scoring

The SSR amplified PCR product was also separated on ABI 3730xl DNA analyzer. The ABI DNA analyzer not only has the ability to perform sequencing by capillary electrophoresis but also can perform a variety of additional DNA analysis applications based on the sizing and intensity of fluorescently labelled DNA fragments. Collectively, these applications are referred to as "fragment analysis". For the fragment analyses following steps were involved:

3.2.4.1 Determining optimum level of dilution

Since the sensitivity of fluorescence-detector is much below the signal quantity of PCR-amplicons, "blow-out" of peak-heights may occur. To avoid this, SSR-amplicons were subjected to suitable recommended level of dilution. Thus, the optimal dilution for the SSR amplicons obtained in this study was determined initially through trial-and-error method. Dilutions were carried out in two-step procedure: initially, in the first step, amplified products to be multiplexed were added together into a single well in 96-well dilution plate, and, in the second step, mixed PCR-products from the first step were combined with size-standard (LIZ® 600) into a single well in the 96-well optical plate (3730 plate, ABI, USA). Initially dilution experiments started with 1:20 in the first step and 1:10 in the second, achieving the final dilution level of 1:200 dilution.

3.2.4.2 Preparation of sample

PCR-products were confirmed of their amplification on the 4% using gel electrophoresis. From each sample of PCR products to be multiplexed, 2 µL aliquot was transferred to 96-well plate and diluted with 40 µL nuclease-free water. Homogeneous mixture of 8.85 µL HiDi and 0.15 µL ladder (LIZ® 600) was prepared freshly and 8 µL of it was added to 2 µL of multiplexed PCR product taken in fresh 96-well optical plate (3730 plate, Applied Biosystems, USA) and covered with adhesive optical covers (Applied Biosystems, USA) followed by heating at 95° C for 5 minutes and kept at 4° C till fed to ABI DNA analyzer (3730xl, Applied Biosystems, USA) for capillary gel electrophoretic separation of amplicons according to their sizes in terms of base-pairs (bp). This cocktail of 10 µL for each sample was directly loaded to ABI DNA analyzer 3730xl set to support

multiplexing of different dye colors (i.e. mixing the amplification products labeled with different dyes in the same well).

3.2.4.3 Capillary electrophoresis

Both the dye-labelled post-PCR-multiplexed samples and freshly prepared cock-tail of recommended size standard (GeneScan™ 600 LIZ®) fragments were co-injected to the fragment analyser for capillary gel electrophoretic separation of amplicons resolving them according to their size in terms of base pairs (bp). The separation of different sized fragments occurs taking into account not only size and but also its net negative charge contributed by each base of the fragment. As the PCR-amplicons move across the polymer gel filled in the capillary and encounter the laser window, the fluorophores present in the fluorescent labelled SSR marker-primed-amplicons in the sample produce the fluorescence, the signal of which is detected by integrated optical detection unit in the instrument. The run-data was checked for the quality and suitable corrections were incorporated whereever necessary, keeping-in-mind the purpose of study, and the re-runs were performed till the desired quality was achieved. The final output, which was not readily-readable in terms of allele sizes, was stored in compact discs for subsequent processing.

Since the instrument is integrated with only data collection software, the data out-put file is not readily-readable. Therefore, the out-put file was used as input for the recommended data analysis software (Gene-Mapper®, ABI, USA).

3.2.4.4 Microsatellite (SSR) marker allele scoring using software

Since the output file of DNA analyzer instrument was not readily-readable, secondary data analysis software ‘Gene-Mapper® V.4.0’ (ABI, USA) was used for the computational analysis of the gel capillary electrophoretic data to obtain allele-size information in form of chromatogram. An over-view of procedural steps followed in setting-up of the software is provided in Appendix A2, parameters are described in Appendix A3, and step-wise protocol is given in Appendix A4. The quality analysis parameters are summarized in Appendix A5. The threshold-level of dye-signal intensity was set at 400. The peaks in the chromatogram having dye-signal intensity above 400 along the y-axis but having base discrete form those of neighboring peaks along the x-axis were

considered as alleles. The numbers at the bottom of each peak indicate the corresponding allele sizes in base-pairs. Different colored peaks correspond to different SSR loci used in the post-PCR multiplexing. The final output was generated and stored as excel-sheet having the matrix of alleles, detected in 618 Indian wheat accessions, of 42 SSR marker loci. The data was formatted as per the required input-file-format of the bioinformatic softwares, or tools, discussed in the subsequent sections.

3.2.5 Statistical analysis of SSR data

For each marker locus, SSR allelic composition was determined in the genotypes using the information on allelic data for SSR loci obtained from ABI DNA analyzer and were arranged in a matrix of SSR loci versus Indian wheat landrace accessions. The SSR allelic data-set was subjected to various statistical analyses using different bioinformatic softwares.

3.2.5.1 Marker polymorphism analysis

The data were converted into required input file format. The allelic data were subjected to Power Marker software (Liu and Muse, 2005) to calculate number of alleles, observed heterozygosity (Ho), gene diversity or expected heterozygosity (He), major allele frequency and polymorphic information content (PIC) value. The PIC value was calculated following the method reported by Botstein *et al*. (1980) using the formula:

$$PIC = 1 - \left[\sum_{i=1}^{n} Pi^2\right] - \left[\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} 2Pi^2 Pj^2\right]$$

Where P_i and P_j are the frequencies of i^{th} and j^{th} allele.

3.2.5.2 Analysis of variance (ANOVA)

Analysis of variance was performed to understand molecular genetic diversity using the genotyping data obtained from 42 SSR marker loci amplified across a panel of 618 Indian wheat accessions. The software tool PowerMarker (Version 3.0) was used and procedure according the user manual (Liu, 2017) was followed.

3.2.5.2.1 Single-locus case control test

Since the prior information on population was not known, single locus case control test (Nielsen and Weir, 1999) based on the multi-allelic trend test (Slager and Schaid, 2001) was used as it has the same degrees of freedom as allele test but remains valid even with the violation of Hardey-weinberg equation (HWE)'s assumptions.

For rare alleles, the total count is < 5 in the 100 genotype-sample, permutation-based distance test was performed to calculate the p-values. Prevosti *et al*. (1975) distance method of permutation was used to compute the distance based on genotype frequencies.

3.2.5.2.2 Expected p-value estimation

Significant correlation between two matrices was confirmed using the Mantel test (Mantel, 1967) which assumes a null hypothesis that for the same objects in two matrices, A and B, distance in one matrix is independent of that in the other matrix. Since the matrices of the present investigation was large, having a total data points of 25,956 (42 loci X 618 genotypes), asymptotic results were applied to obtain an exact p-value.

3.2.5.3 Cluster analysis

Cluster analysis was performed according to the unweighted pair-group method with arithmetic average (UPGMA) with the Nei (Roussel *et al*., 2004) similarity index using Power Marker (Version 3.25) (Anderson *et al*., 1993; Liu and Muse, 2005) and the dendrogram was constructed using MEGA 4.0 software (Tamura *et al.,* 2007).

Genetic distance matrix-derived UPGMA tree was used for cluster analysis performed to ascertain diversity among the 618 wheat accessions. The genetic distance matrix based on genetic dissimilarity indices was derived from allelic information of evenly distributed 42 SSR loci amplified across the genomes of 618 Indian wheat accessions deploying genetic distance method of Cavalli-Sforza (CS) chord. A simple agglomerative (bottom-up) hierarchical clustering method was used to generate unweighted pair group method with arithmatic mean (UPGMA) based tree.

3.2.5.4 Analysis of molecular variance (AMOVA)

An analysis of molecular variance (AMOVA) procedure was executed employing GenAlEx (version 6.5) and Arlequin softwares, using 1000 permutations (Excoffier *et al*., 1992; Peakall and Smouse, 2012). The data were converted into required input file format. The allelic data were subjected to analysis using the software Arlequin to get information on molecular variance (AMOVA) (Excoffier *et al*., 1992).

3.2.5.5 Principal coordinate analysis (PCoA)

A Principal Co-ordinate Analysis (PCoA) was carried out on the distance matrix (D) to visualize the genetic interrelationships among the accessions in three-dimensional PCoA plots, with resultant scores for samples on the first two components plotted pairwise in each case. Multidimensional scaling (MDS) plots were constructed on the Distance Matrix, D, and the stress values have been calculated before PCoA was performed on the matrix of Dice similarity coefficients (Dice, 1945) using the modules DCENTER and EIGEN of the NTSYS-PC ver. 2.1 (Rohlf, 2000) and the 2D plot was generated using the MOD2D module of the PowerMarker version 3.25 (Anderson *et al*., 1993).

3.2.5.6 Analysis of model-based population structure

The data were converted into required input file format. The allelic data were subjected to analysis using the software STRUCTURE version 2.3.3 (Pritchard *et al*. 2000). The number of presumed population (*K*) was set from 2 to 10 with admixture model,

3.2.6 Field evaluation, purification and multiplication of wheat accessions

In the crop season 2011-12, 1500 Indian wheat accessions were raised uder field condition and the recommended package of practice for irrigated wheat was followed. Among 1500 accessions, 1350 were germinated. Based on morphological observation and rejection of grassy-type, wild-relatives, diploids, redundant and admixtures, 618 accessions were selected for further study. In the crop season: 2012-13, 618 accessions were raised in field conditions following recommended package of practice.

3.2.7 Field experimental design

In crop seasons 2011-12 and 2013-13, each accession was raised in field using randomized block design (RBD) with two replications. Each accession was raised in a row of one meter with 30 cm row-to-row and 10 cm plant-to-plant spacing in each plot.

3.2.8 Recording of data and scoring of agro-morphological and phenological traits

Observations were recorded on randomly tagged five representative individual plants, of each accession in each plot (i.e., each of two replications), at appropriate stage of crop growth, when the character under study had full expression. These traits were recorded as per the descriptors of IPGRI (IBPGR, 1985), now called biodiversity international. The phenological data were observed according to the scale of Zadoks *et al*.(1974). Further, recorded data for twenty-three agro-morphological and phenological traits, some qualitative and others quantitative, were grouped according to different states of each trait and were assigned numerical codes in the following manner:

A. Agro-morphological traits

Growth habit (Categorical with assigned values)

Growth habit (GRHB) was recorded based on the visual observation of the angle subtended by tillers with respect to the main stem. Tillers subtending by ≤45 degree angle were recorded as erect, those subtending by >45 but ≤90 degree were described as semi-spreading, and those with ≥90 degree were classified as spreading. The class of the traits were assigned numerical codes: 3 = erect, 5 = semi spreading, and 7 = spreading.

Plant height (cm, categorical with assigned values)

Plant height (PLH) was recorded at maturity, measuring in centimetre (cm) from the ground to the tip of spike, excluding awns (if any). The trait was grouped into following classes: 1 = very short (≤75 cm), 3 = short (>75 to 90 cm), 5 = medium (>90 to 105 cm), 7 = long (>105 to 120 cm), and 9 = very long (≥ 120 cm).

Tillering capacity (Numerical made categorized with assigned values)

Tillering capacity (TILC) was recorded based on the subjective judgement of number of tillers per plant, during tillering but before booting, at a given plant density (30

X 10 cm spacing). The trait was grouped into following classes: 3 = low (≤ 5 tillers), 5 =intermediate (>5 to 10 tillers), and 7 = high (≥ 10 tillers).

B. Inflorescence characteristics

Spikelength (cm, made categorical with assigned values)

Spike length (SPL) was recorded, based on a matured spike, measuring in centimetre (cm) from base of the spike to the tip excluding awns (if any) and the data were grouped into the following classes: 1 = very short (≤6 cm), 3 = short (>6 to 8 cm), 5 = medium (>8 to 10 cm), 7 = long (>10 to 12 cm), and 9 = very long (>12 cm).

Number of spikelets per spike (Numerical, made categorical with assigned values)

Number of spikelets per spike (NSLS) was recorded based on the average number of spikelets per spike from five typical spikes selected from each accession. The categories of the traits are: 3 = few (≤ 15), 5 = medium (>15 to 20), and 7 = very high (≥20).

Number of florets per spikelet (Numerical, made categorical with assigned values)

Number of florets per spikelet (NFSL) was recorded based on the average number of spikelets obtained from the central portion of the five typical spikes selected from each accession. The classes of the traits are: 1 = low (1), 3 = medium (3), 5 = high (5), and 7 = very high (7).

Number of seeds per spikelet (Numerical, made categorical with assigned values)

Number of seeds per spikelet (NSLS) was recorded based on the average number of seeds from a spikelet obtained from the central portion of the five typical spikes selected from each accession. The data was grouped into the following classes: 1 = very low (1), 3 = low (2), 5 = medium (3), 7 = high (4), and 9 = very high (5).

Maximum number of seeds per floret (Numerical, made categorical with assigned values)

Maximum number of seeds per floret (MNSF) was recorded based on the maximum number of seeds found in atleast one floret of a spikelet obtained from the central portion of the five typical spikes selected from each accession. The data was grouped into the following classes: 3 = low (1), 5 = medium (2), and 7 = high (3).

Number of seeds per spike (Numerical, made categorical with assigned values)

Number of seeds per spike (NSS) was recorded by counting the actual number of seeds in five typical spikes selected from each accession. The data were grouped into the following classes: 1 = very low (≤25), 3 = low (>25 to 50), 5 = medium (>50 to 75), 7 = high >75 to 100), and 9 = very high (>100).

Awnness (Categorical with assigned values)

Awnness (AWN) was recorded based on visual observation of tip of the spike and the observation was grouped into the following categories: 1 = awnless, 5 = awnletted (short awns), and 7 = awned (conspicuous awns).

Glume color (Categorical with assigned values)

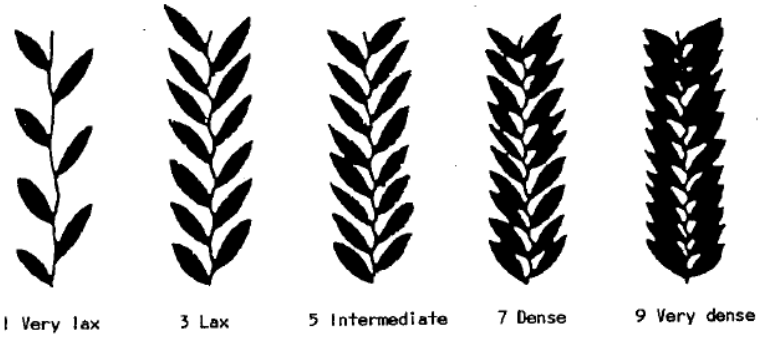
Glume color (GLCL) was recorded based on the visual observation of the outer glume. The following classes were recorded: 1 = white, 5 = red to brown and 9 = purple to black.

Glume hairiness (Categorical with assigned values)

Glume hairiness (GLHR) was recorded based on the visual observation of the outer side of sterile glume. The trait classes were: 1 = absent, 5 = low, and 9 = high.

Spike density (Categorical with assigned values)

Spike density (SPD) was measured based on the visual observation of matured spike and comparing with visual score on a 1-9 scale (IBPGR, 1985). It is important to note that spike density is not same as spike shape. The trait classes were: 1 = very lax, 3 = lax, 5 = intermediate, 7 = dense, and 9 = very dense.



Visual scoring guide for spike density on 1-9 scale (Reproduced from IBPGR, 1985)

Seed color (Categorical with assigned values)

Seed color (SDCL) was recorded based on the visual observation of dehusked seed and the following classes were found: 1 = white, 5 = red, and 9 = purple.

Seed size (Relative: categorized and values assigned)

Seed size (SDZ) was recorded based on their relative size and expressed in terms of qualitative descriptors coded into the following classes: 3 = small, 5 = intermediate, and 7 = large.

Degree of seed shriveling (Categorical with assigned values)

Degree of seed shrivelling (DSHL) was recorded based on the appearance of dry seed after harvest and were grouped into the following classes: 3 = plump, 5 = intermediate, and 7 = shrivelled.

C. Phenological traits

Days to emergence (Numerical, made categorical with assigned values)

Days to emergence (DEM) was recorded based on days from sowing to emergence of fourth leaf in atleast 50% of plants in each accession. The trait classes are: 1 = very early (≤ 12 days), 3 = early (>12 to 15 days), 5 = medium (>15 to 18 days), 7 = late (>18 to 21 days), and 9 = very late (>21 days).

Days to tillering (Numerical, made categorical with assigned values)

Days to tillering (DTIL) was recorded based on days from sowing to formation of atleast one tiller in 50% of plants in each accession. The trait classes are: 3 = early (15 to 20 days), 5 = medium (>20 to 25 days), 7 = late (>25 to 30 days), and 9 = very late (>30 days).

Days to booting (Numerical, made categorical with assigned values)

Days to booting (DBOO) was recorded based on days from sowing to swelling of boot on the main stem in 50% of plants in each accession. The trait classes are: 1 = very early (≤ 50 days), 3 = early (>50 to 60 days), 5 = medium (>60 to 75 days), 7 = late (>75 to 90 days), and 9 = very late (>90 days).

Days to spike emergence (Numerical, made categorical with assigned values)

Days to spike emergence (DSEM) was recorded based on days from sowing to complete appearance of spike from the head in 50% of plants in each accession. The trait classes are: 1 = very early (≤ 60 days), 3 = early (>60 to 75 days), 5 = medium (>75 to 90 days), 7 = late (>90 to 105 days), and 9 = very late (>105 days).

Days to flowering or anthesis (Numerical, made categorical with assigned values)

Days to flowering (DFLO) or anthesis was recorded as days from sowing to 50% flowering of plants in each accession bearing atleast one open flower (anthers visible). The trait was grouped into the following classes: 1 = very early (≤70 days), 3 = early (>70 to 85 days), 5 = medium (>85 to 105 days), 7 = late (>105 to 120 days), and 9 = very late (≥120 days).

Days to physiological maturity (Numerical, made categorical with assigned values)

Days to physiological maturity (DMAT) was recorded based on days from sowing to head loosing green color in five tagged plants in each accession. The trait classes are: 1 = early (90 to 105 days), 3 = medium (>105 to 120 days), 5 = late (>120 to 135 days), and 7 = very late (>135 days).

D. Production traits

Thousand kernels weight (g: categorized and assigned values)

Thousand kernel weight (TKW) in gram (g) was recorded by weighing 1000 manually counted seeds obtained from each of the five tagged plants in each accession. Wherever the seed number was less than 1000, actual numbers of seeds were weighed and TKW was arithmetically derived. The trait classes are: 1 = very low (≤20 g), 3 = low (>20 to 30 g), 5 = medium (>30 to 40 g), 7 = high (>40 to 50 g), and 9 = very high (>50 g).

3.2.9 Morphological diversity analysis

Mean, minimum, maximum, difference between maximum and minimum, standard deviation, coefficient of variation, standard error, skewness and kurtosis were determined for recorded data of each of the quantitative characters, using statistical software SAS, version 9.3 (SAS Institute Inc., USA, 2011). Further, recorded data for both qualitative and

quantitative characters were grouped according to different states of each trait and were assigned numerical codes. Phenotypic diversity index (PDI), denoted by H', was estimated for each character using Shannon-Weaver diversity index, according to the standard procedure (Mengistu *et al*. 2015) and accordingly classified as high (H≥6.0), intermediate (0.4≤H≥0.6) and low (H≤0.4), as described by Eticha *et al*. (2005). Numerically coded data were used for estimating Carl Pearson's coefficients of correlation, factorial analysis, principal component analysis, cluster analysis and Genotype-by-trait (GT) plot using statistical software SASversion 9.3, and SAS procedures (SAS; 2011). Heat map for Pearson's correlation coefficients and minimum spanning tree for genotypes were generated using the GenStat software (11th edition, www.genstat.com). Factorial analysis was performed to know the trait(s) contributing maximum variability. Principal component analysis was employed to examine the percentage contribution of each trait to the total genetic variation. Agglomerative hierarchical clustering was performed with Euclidean distance matrix using Ward's linkage method. GT-biplot was used to graphically display inter-relationship among accessions based on the variability for 23 agro-morphological traits.

3.2.9.1 Statistical analyses

3.2.9.1.1 Univariate analysis

3.2.9.1.1.1 Analysis of quantitative traits

Prior to analysis, the data were standardized to zero mean and unit variance, by subtracting the mean value of the character from mean observations for each trait and subsequently dividing it by respective standard deviation; as various traits were measured on very different scales. Mean, minimum, maximum, difference between maximum and minimum, standard deviation, coefficient of variation, standard error, skewness and kurtosis were determined for recorded data of each of the fifteen quantitative traits, using statistical software SAS, version 9.3 (SAS Institute Inc., USA, 2011).

3.2.9.1.1.2 Combined analysis of quantitative and qualitative traits

Further, recorded data for both qualitative and quantitative characters were grouped according to different states (classes) of each trait and were assigned numerical codes.

Phenotypic diversity index (PDI), denoted by H', was estimated for each character using Shannon-Weaver diversity index, according to the standard procedure (Mengistu *et al*. 2015)

3.2.9.1.2 Multivariate analysis

3.2.9.1.2.1 Correlation analysis

The correlation coefficients of fifteen quantitative traits and eight qualitative traits were used in characterizing the 618 Indian wheat accessions. Data were analysed using Karl Pearson coefficients of correlation, to assess the traits that are associated with each other, both positively as well as negatively, at statistically significant levels, $p<0.05$ and $p<0.01$.

3.2.9.1.2.2 GT-biplot analysis

GT-biplot analysis for visualization of genetic correlation among genotypes and understanding of genotype-by-trait relationships was performed as described by Ma *et al*. (2004) and Yan and Fregeau-Reid, 2008.

3.2.9.1.2.3 Principal component analysis

First, factor analysis was performed in order to reduce a large set of phenotypic traits to a more meaningful smaller set of traits and to know which trait is contributing maximum to the total variability, because genetic improvement depends on the magnitude of genetic variation. A Principal component analysis (PCA) was carried out on the traits obtained after the factor analysis. PCA analysis was used to select the number of principal components showing Eigen value of greater than or equal to one, as described by Mishra *et al.,* 2015.

3.2.9.1.2.4 Cluster analysis

Euclidean distance-based hierarchical clustering using Ward's minimum variance, measure of dissimilarity, method was used to assemble 618 Indian wheat accessions into dendrogram projected on a scale of semi-partial R-quared values, as described by Hoque and Rahman, 2006. Statistically optimal number of clusters in a dendrogram was determined using criteria described by Millign and Cooper (1985). Accordingly, in the present study optimal number of clusters was determined based on the concurrence of these

three criteria: one, Cubic clustering criteria (CCC) curve should have the first peak with sudden fall; two, Pseudo F curve should be declining slowly; and, three, Pseudo T-squared curve should be showing the steady phase.

CHAPTER 4

RESULTS

4. RESULTS

The results of the present study were obtained after the raw data; recorded/observed directly from field- and laboratory experiments; were processed, organized and analyzed by employing computational methods and bioinformatic tools that are advanced and widely used in genomic and genetic diversity studies using DNA markers and morphological traits. These results are presented, illustrated and/or depicted in this chapter in the form of organized table(s), graph(s) and figure(s), respectively, that are systematically numbered and referred to in the text.

4.1 Analysis of genetic diversity of Indian wheat landraces using DNA markers

To study genetic diversity, population structure and phylogeny in a panel of 618 Indian wheat landrace accessions (Table 3.1 and Table 3.2), DNA markers in general, microsatellite or simple sequence repeat (SSR) markers in particular, were deployed. The following sections and sub-sections illustrate and explain the results obtained from various experiments and computational analyses.

4.1.1 Genome-level diversity analyses

4.1.1.1 High resolution genotyping

4.1.1.1.1 Fluorescent-labelled microsatellite (SSR) markers

For achieving high resolution genotyping, each of the 42 SSR markers was labelled with one of the four different colored fluorophore-carrying dyes. The details of fluorescent-labelling and dye colours are given in Table 3.3, Table 3.4, Table 3.5, and Appendix A1. Forty-one markers belonged to "xgwm-series" of wheat microsatellite markers and one was "Barc1121".

4.1.1.1.2 Confirmation of PCR-amplicons

Before injecting the sample for capillary gel electrophoresis, the presence of amplicons of SSR loci was confirmed through 4% agarose gel electrophoresis. As an illustration, the confirmation of PCR-products amplified by SSR markers in 18 Indian wheat landraces is illustrated in Figure 4.1. Four panels correspond to PCR-amplification profiles of four SSR loci primed by four different-color-labelled SSR primers. The presence/absence of the band indicates the presence/absence of PCR-amplicons of respective marker loci. Size of the band corresponds to alleles size of the

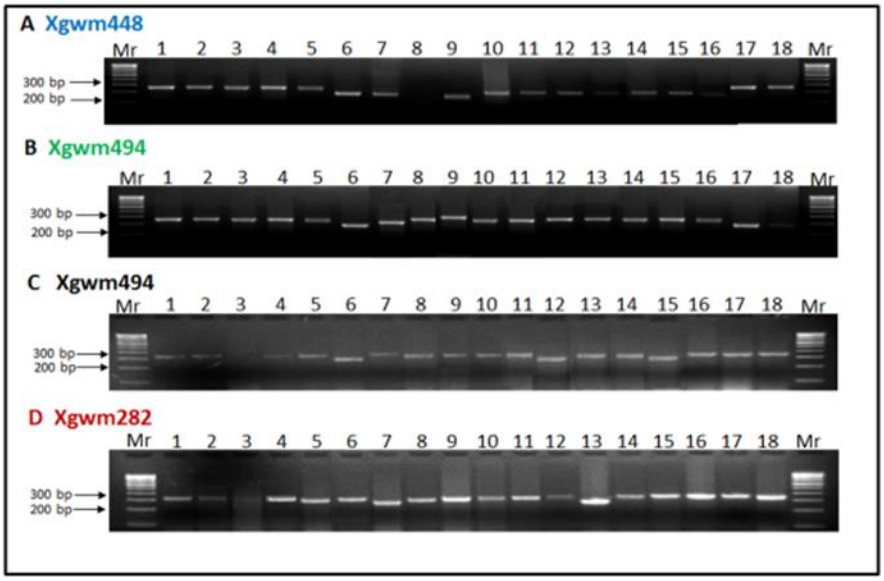


Figure 4.1 Illustration of amplification of SSR marker loci before capillary gel electrophoresis. Lanes 1-18, different Indian wheat landrace accessions listed in Table 3.1; Mr: 100 bp DNA ladder. Different colors of marker names (A, B, C.and D) correspond to respective fluorescent dye colors used for labelling them as detailed in Table 3.3, Table 3.4 and Table 3.5.

corresponding marker locus. The absence of PCR-amplicons corresponds to the presence of null-alleles. Failure of PCR-amplification was ruled-out after the repetition of confirmatory experiments involving positive- and negative- controls at various levels: DNA isolation through PCR to agarose gel electrophoresis.

4.1.1.1.3 High-resolution genotyping

After confirmation of PCR-amplicons, multiplexed amplicons and size standard were co-injected for capillary gel electrophoresis. Size- and charge-based differential movement of different-sized amplicons and size-standards were captured by the camera unit and stored in the output file. Since the output file of DNA analyzer instrument was not readily-readable, secondary data analysis software 'Gene-Mapper ® V.4.0' (ABI, USA) was used to obtain allele-size information in form of chromatogram. For illustration purpose, an example of a chromatogram displaying different allele-sizes amplified by four different colored markers is depicted in Figure 4.2. An over-view of procedural steps followed in data-analysis using the software is provided in Appendix A.2, Appendix A.3, Appendix A.4 and Appendix A.5. The threshold-level of dye-signal intensity was set at 200. Every peak in the chromatogram (Figure 4.2) having dye-signal intensity above 200 along the y-axis and having its base discrete form those of neighboring peaks along the x-axis were considered as alleles. The numbers at the bottom of each peak indicate the respective allele sizes in base-pairs. Different coloured peaks correspond to dye-colours used to label respective SSR markers whose amplicons were pooled in post-PCR multiplexing. The final output was generated and stored in an excel-sheet having the matrix of alleles of 42 SSR marker loci, detected in 618 Indian wheat accessions, The data was arranged as per the required input-file-format of the respective bioinformatic software deployed in the present study.

4.1.1.2 Different type of alleles and their distribution among 42 loci across 618 genotypes

The details of the maximum allele frequency, number of alleles, gene diversity index, heterozygosity, polymorphism information content (PIC) and F-value, derived from one-way analysis of variance (ANOVA), for 42 SSR marker loci among 618 Indian wheat accessions, are presented in Table 4.1. The detailed ANOVA table is given in Appendix A.6. As revealed in Table 4.1, F-value was uniformly one for all the loci and it was considered in error reduction procedure. In addition, heterozygosity was zero

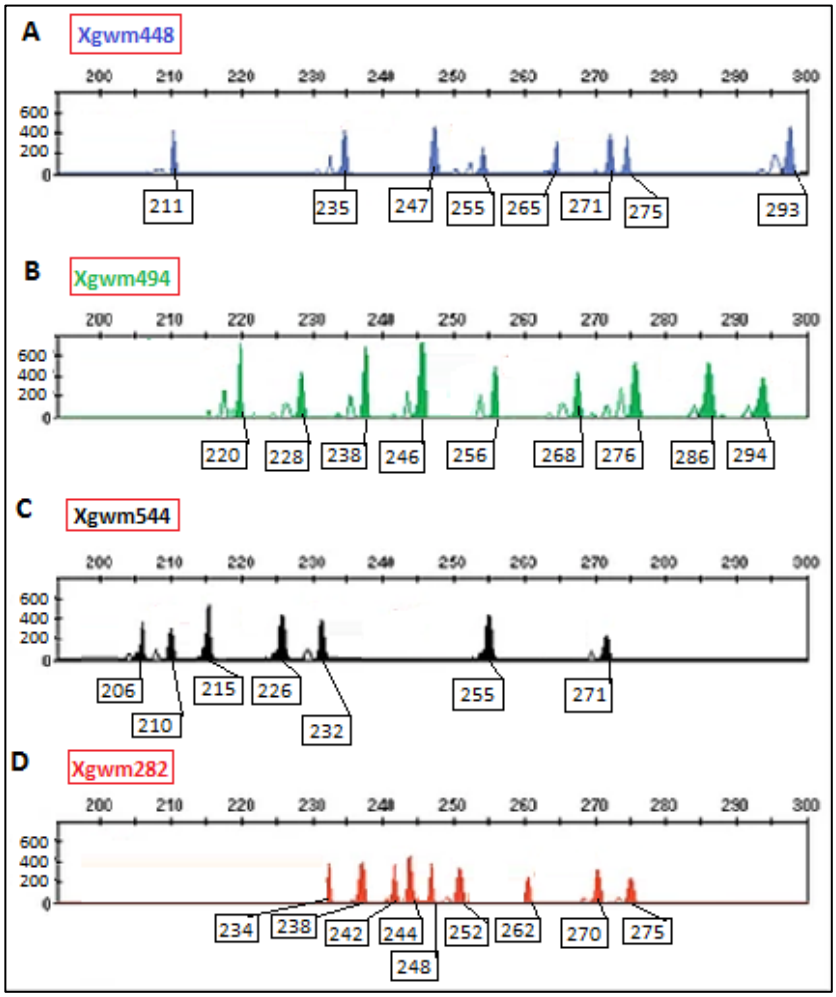


Figure 4.2 Capillary chromatogram illustrating high resolution allele sizing of 4 SSR marker loci. Markers names and colors correspond to those detailed in Table 3.3, Table 3.4 and Table 3.5. Numbers below the colored peak represent the amplicon sizes in base pairs. Amplicon sizes correspond to the first 8, 9, 7 and 9 lanes in panels A, B, C and D, respectively, of Figure 4.1.

Table 4.1 Details of gene diversity derived from analysis of variance (ANOVA) for 42 SSR markers loci across 618 Indian wheat genomes.

S. No.	Chr.*	Marker	MAF (>5%)	No of Alleles	Gene Diversity	Hetero zygosity	PIC*	F
1	1-AS	xgwm136	0.16	26.00	0.91	0.00	0.91	1.00
2	1-AL	xgwm497	0.15	25.00	0.92	0.00	0.91	1.00
3	2-AS	xgwm448	0.17	23.00	0.92	0.00	0.92	1.00
4	2-AL	xgwm445	0.19	22.00	0.92	0.00	0.92	1.00
5	3-AS	xgwm2	0.17	24.00	0.91	0.00	0.91	1.00
6	3-AL	xgwm494	0.11	23.00	0.93	0.00	0.93	1.00
7	4-AS	xgwm165	0.14	20.00	0.92	0.00	0.91	1.00
8	4-AL	Xgwm637	0.26	22.00	0.88	0.00	0.87	1.00
9	5-AS	xgwm443	0.21	23.00	0.90	0.00	0.89	1.00
10	5-AL	Xgwm666	0.19	22.00	0.91	0.00	0.90	1.00
11	6-AS	xgwm459	0.10	23.00	0.93	0.00	0.93	1.00
12	6-AL	xgwm169	0.12	22.00	0.92	0.00	0.91	1.00
13	7-AS	xgwm635	0.26	21.00	0.88	0.00	0.88	1.00
14	7-AL	xgwm282	0.17	24.00	0.92	0.00	0.91	1.00
15	1-BS	xgwm550	0.17	15.00	0.90	0.00	0.90	1.00
16	1-BL	xgwm268	0.19	20.00	0.90	0.00	0.89	1.00
17	2-BS	xgwm257	0.29	16.00	0.86	0.00	0.85	1.00
18	2-BL	xgwm47	0.20	17.00	0.89	0.00	0.88	1.00
19	3-BS	xgwm533	0.18	18.00	0.91	0.00	0.91	1.00
20	3-BL	xgwm181	0.17	18.00	0.89	0.00	0.89	1.00
21	4-BS	xgwm368	0.17	17.00	0.89	0.00	0.88	1.00
22	4-BL	Xgwm513	0.18	18.00	0.90	0.00	0.89	1.00
23	5-BS	xgwm544	0.27	17.00	0.85	0.00	0.84	1.00
24	5-BL	Xgwm335	0.18	18.00	0.90	0.00	0.89	1.00
25	6-BS	Xgwm361	0.14	18.00	0.91	0.00	0.90	1.00
26	6-BL	Xgwm219	0.19	20.00	0.91	0.00	0.90	1.00
27	7-BS	xgwm400	0.14	18.00	0.91	0.00	0.90	1.00
28	7-BL	xgwm131	0.14	20.00	0.92	0.00	0.91	1.00
29	1-DS	Xgwm458	0.17	15.00	0.90	0.00	0.89	1.00
30	1-DL	xgwm642	0.14	16.00	0.91	0.00	0.91	1.00
31	2-DS	xgwm484	0.25	14.00	0.84	0.00	0.82	1.00
32	2-DL	xgwm349	0.25	13.00	0.87	0.00	0.85	1.00
33	3-DS	xgwm183	0.11	16.00	0.92	0.00	0.92	1.00
34	3-DL	xgwm3	0.14	15.00	0.91	0.00	0.90	1.00
35	4-DS	xgwm194	0.17	16.00	0.91	0.00	0.91	1.00
36	4-DL	xgwm624	0.29	15.00	0.87	0.00	0.86	1.00
37	5-DS	xgwm190	0.15	15.00	0.91	0.00	0.90	1.00
38	5-DL	Xgwm182	0.22	13.00	0.89	0.00	0.88	1.00
39	6-DS	Xgdm141	0.15	12.00	0.89	0.00	0.88	1.00
40	6-DL	Barc1121	0.19	13.00	0.90	0.00	0.89	1.00
41	7-DS	Xgwm44	0.18	13.00	0.89	0.00	0.88	1.00
42	7-DL	Xgwm37	0.16	15.00	0.91	0.00	0.90	1.00
		Mean	**0.18**	**18.36**	**0.90**	**0.00**	**0.89**	**1.00**

*Note: Chr.: Chromosomal location, MAF: Major allele frequency, PIC: Polymorphic information content, F: F-value of one-way ANOVA.

for all the loci, indicating that all the 42 SSR marker loci amplified in a panel of 618 genotypes were fixed (i.e., homozygous).

Microsoft excel work book was used for further analysis of data using information on major allele frequency (Table 4.1) to compute number of major-, minor-, unique- and null-alleles for each of the 42 SSR marker loci and these results are presented in Table 4.2. In addition, range of amplicon sizes were also estimated manually from the raw genotyping data and the results are provided in Table 4.2. Forty-two SSR marker loci revealed a total of 771 size-based alleles, with an average of 18.36 per locus, with amplicon-sizes ranging from 104 bp to 422 bp (Table 4.1 and Table 4.2). The highest number of 26 alleles per locus was observed in SSR marker xgwm136 (located on 1-AS) followed by 25 per locus at xgwm497 (1-AL). The minimum number of 12 alleles per locus was recorded in xgdm141 located on 6-DS. The number of major alleles per locus ranged from 5.72 at xgwm637 (4-AL) to 1.76 at xgwm183 (3-DS), with an average of 3.27 alleles per locus. The number of minor alleles per locus was highest (21.84) at xgwm136 (1-AS), followed by 21.25 and 20.7 at xgwm497 (1-AL) and xgwm459 (6-AS), respectively. The lowest average number of 9.75 minor alleles per locus was present at xgwm349 (2-DL), followed by 10.2 and 10.5 at xgdm141 (6-DS) and xgwm484 (2-DS), respectively. The studied set of 42 SSR markers had polymorphism information content (PIC) in the range of 0.82, exhibited by the marker xgwm484 (2-DS), to 0.93 exhibited by each of the two loci: xgwm494 (3-AL) and xgwm459 (6-AS) (Table 4.1). An overall gene diversity, computed based on Shannon-wiener diversity index, revealed by 42 SSR loci among 618 genotypes, was 0.9. The highest gene diversity of 0.93 was recorded for each of the the 2 loci, xgwm494 (3-AL) and xgwm459 (6-AS), followed by 0.92 (exhibited by each of the 8 loci) and 0.91(shown by each of the 12 loci). The lowest gene diversity index of 0.84 was recorded at xgwm484 locus (2-DS), followed by 0.85 in case of xgwm544 (5-BS). Major allele frequency ranged from 0.1, recorded for xgwm459 (6-AS), to 29, recorded at locus xgwm257 (2-BS).

The summary statistics of amplification profile of all the 42 marker loci genotyped across 618 genotypes were manually computed using the information in Table 4.1, Table 4.2 and Appendix A.6, and are presented in Table 4.3. The number and percentage of different category of alleles were computed at three levels: for all genotypes and loci, for each genotype, and for each locus. Though a total of 25956

Table 4.2 Allelic information of 42 SSR markers loci across 618 Indian wheat genomes.

S. No.	Chr.*	Marker Name	Number of total alleles	Range of alleles size (bp)	Number of Major alleles	Number of minor alleles	Number of unique alleles	Number of null alleles
1	1-AS	Xgwm136	26.00	236-321	4.16	21.84	1	0
2	1-AL	Xgwm497	25.00	104-321	3.75	21.25	3	82
3	2-AS	Xgwm448	23.00	201-375	3.91	19.09	2	3
4	2-AL	Xgwm445	22.00	130-326	4.18	17.82	2	22
5	3-AS	Xgwm2	24.00	124-304	4.08	19.92	4	59
6	3-AL	Xgwm494	23.00	138-332	2.53	20.47	2	58
7	4-AS	Xgwm165	20.00	150-320	2.8	17.2	2	26
8	4-AL	Xgwm637	22.00	144-316	5.72	16.28	2	37
9	5-AS	Xgwm443	23.00	150-370	4.83	18.17	4	23
10	5-AL	Xgwm666	22.00	114-300	4.18	17.82	2	30
11	6-AS	Xgwm459	23.00	114-326	2.3	20.7	2	55
12	6-AL	Xgwm169	22.00	164-354	2.64	19.36	3	46
13	7-AS	Xgwm635	21.00	108-302	5.46	15.54	2	54
14	7-AL	Xgwm282	24.00	160-368	4.08	19.92	1	37
15	1-BS	Xgwm550	15.00	150-300	2.55	12.45	0	106
16	1-BL	Xgwm268	20.00	184-350	3.8	16.2	1	64
17	2-BS	Xgwm257	16.00	174-310	4.64	11.36	2	20
18	2-BL	Xgwm47	17.00	128-278	3.4	13.6	1	0
19	3-BS	Xgwm533	18.00	108-282	3.24	14.76	1	25
20	3-BL	Xgwm181	18.00	120-276	3.06	14.94	1	0
21	4-BS	Xgwm368	17.00	128-280	2.89	14.11	1	0
22	4-BL	Xgwm513	18.00	124-292	3.24	14.76	2	32
23	5-BS	Xgwm544	17.00	236-422	4.59	12.41	1	0
24	5-BL	Xgwm335	18.00	128-328	3.24	14.76	2	0
25	6-BS	Xgwm361	18.00	118-308	2.52	15.48	1	34
26	6-BL	Xgwm219	20.00	148-368	3.8	16.2	1	32
27	7-BS	Xgwm400	18.00	134-312	2.52	15.48	2	35
28	7-BL	Xgwm131	20.00	104-304	2.8	17.2	2	17
29	1-DS	Xgwm458	15.00	104-248	2.55	12.45	1	20
30	1-DL	Xgwm642	16.00	214-378	2.24	13.76	1	20
31	2-DS	Xgwm484	14.00	140-278	3.5	10.5	0	20
32	2-DL	Xgwm349	13.00	198-320	3.25	9.75	0	20
33	3-DS	Xgwm183	16.00	114-336	1.76	14.24	1	20
34	3-DL	Xgwm3	15.00	110-252	2.1	12.9	1	20
35	4-DS	Xgwm194	16.00	109-352	2.72	13.28	1	20
36	4-DL	xgwm624	15.00	120-252	4.35	10.65	1	20
37	5-DS	Xgwm190	15.00	218-382	2.25	12.75	0	20
38	5-DL	Xgwm182	13.00	154-292	2.86	10.14	1	20
39	6-DS	Xgdm141	12.00	132-242	1.8	10.2	0	20
40	6-DL	Barc1121	13.00	109-256	2.47	10.53	0	20
41	7-DS	Xgwm44	13.00	128-280	2.34	10.66	0	20
42	7-DL	Xgwm37	15.00	168-412	2.4	12.6	2	20
		Total	**771**	**-**	**137.5**	**633.5**	**59**	**1177**
		Mean	**18.36**	**-**	**3.27**	**15.8**	**1.41**	**28.23**

*Note: Chr.: Chromosomal location, bp: base pairs.

Table 4.3 Summary of amplification profile of 42 SSR markers loci among 618 Indian wheat accessions.

		Expected amplicons								
		Total	**Not amplified / Missing amplicon**	**Amplified/Amplicons**						
				Total	**Mono morphic**	**Polymorphic**				
						Total	**Major alleles**	**Minor alleles (<5%)**		
								Total	**Common**	**Unique**
For all genotypes and loci	Total	25,956	1,177	24,779	24,008	771	139	632	574	59
	(%)	100	4.53^{E}	92.47^{E}	92.49^{E}, 96.89^{A}	2.9^{E}, 3.11^{A}	0.5^{E}, 0.55^{A}, 17.89^{P}	2.4^{E}, 2.56^{A}, 82.1^{P}	2.2^{E}, 2.32^{A}, 74.45^{P}, 90.68^{Mi}	0.24^{A}, 7.65^{P} 9.32^{Mi}
For each genotype	Mean	42	2	40.13	38.85	1.25	0.22	1.02	0.93	0.1
	(%)	100	4.76^{E}	95.55^{E}	92.5^{E}, 96.8^{A}	2.98^{E}, 3.12^{A}	0.5^{E}, 0.55^{A}, 17.6^{P}	2.43^{E}, 2.54^{A}, 81.6^{P}	2.2^{E}, 2.32^{A}, 74.4^{P}	0.24^{A}, 8.0^{P} 9.8^{Mi}
For each marker locus	Mean	618	28	590	571.62	18.36	3.29	15.1	13.7	1.45
	(%)	100	4.53^{E}	95.47^{E}	92.5^{E}, 96.88^{A}	2.97^{E}, 2.3^{A}	0.53^{E}, 0.56^{A}, 17.92^{P}	2.44^{E}, 2.56^{A}, 82.24^{P}	2.22^{E}, 3.32^{A}, 74.62^{P}, 90.73^{Mi}	0.25^{A}, 7.89^{P} 9.6^{Mi}

*Note: (E): Of total expected amplicons, (A): of total amplicons, P: of total polymorphic amplicons, Ma: of major alleles, Mi: of total minor alleles.

amplicons were expected, observed were 24779 (constituting 92.47% of expected) as 1177 amplicons (making 4.53 % of expected) did not show-up. Among 24779 amplified alleles, 24008 (comprising 92.49% and 96.89%, respectively, of expected and amplified) were monomorphic and 771 (2.9% of expected and 3.11% of amplified) were polymorphic. Based on frequency of their occurrence among a set of 618 Indian wheat accessions, the 771 polymorphic alleles were further partitioned into 139 major- and 632 minor alleles; denoting more- and less than 5% of their occurrence among 618 genotypes, respectively. Major alleles constituted 0.5%, 0.55% and 17.89% of expected, amplified and polymorphic alleles, respectively; whereas minor alleles were 2.4%, 2.56% and 82.1% of expected, amplified and polymorphic alleles, respectively. Out of 632 minor alleles, 574 (comprising 2.2%, 2.32%, 74.45% and 90.68% of expected, amplified and polymorphic and minor alleles, respectively) were common/ordinary alleles shared by two or more accessions, and 59 (7.65% of polymorphic and 9.32% of minor alleles) were unique to their respective single accessions.

On an average basis, each of 618 Indian wheat accessions had, out of 42 expected alleles per genotype, 2 null-, 40.13 amplified-, 38.85 monomorphic-, 1.25 polymorphic-, 0.22 major- and 1.02 minor alleles; comprising 4.76%, 95.555, 92.5%, 2.98%, 0.5% and 2.43% of expected number of alleles per accession, respectively (Table 4.3). Out of an average of 40.13 amplicons per genotype across 42 loci, 96.89% and 3.11% were monomorphic- and polymorphic alleles, respectively. An average of 1.25 polymorphic alleles per accession, which included, on average basis, 0.22 major- and 1.02 minor alleles per genotype, constituted 0.5% and 2.43% of expected number of alleles per genotype, respectively. An average of 0.93 minor alleles per genotype was common allele among two or more of 618 genotypes whereas an average of 0.1 minor allele per genotype was unique to single accession.

Among 42 SSR marker loci, each locus on an average had, out of 618 expected alleles, 28 null-, 590 amplified-, 571.62 monomorphic-, 18.36 polymorphic-, 3.29 major- and 15.1 minor alleles; comprising 4.53%, 95.47%, 92.5%, 2.97%, 0.53% and 2.44%, of average expected number of alleles per locus, respectively. Out of 590 amplified alleles per SSR marker locus across 618 genotypes, on an average basis, 571.62 and 18.36 per locus (constituting 92.5% and 2.97% of expected number of alleles per locus, respectively) were monomorphic- and polymorphic alleles,

respectively. An average of 18.36 polymorphic alleles per locus was further partitioned, on an average basis, into 3.29 major- and 15.1 minor alleles per locus, constituting 17.92% and 82.24%, respectively. An average of 13.7 minor alleles per locus was common allele among two or more of 618 accessions, whereas an average of 1.45 alleles per accession was unique to single genotype.

4.1.1.2.1 Analysis of null alleles

Out of 25956 expected number of alleles, 1177 alleles were not PCR-amplified (Table 4.3). Failure of amplification can be attributed either to presence of null allele among concerned genotypes or to faulty protocol, or reagents/polymerase, at any stage of the method from DNA isolation through PCR to gel/capillary electrophoresis. However, after the repetition of confirmatory experiments using different sets of DNAsamples, PCR-reagents/enzymes and involving positive and negative controls, it was proved that there was a consistency in the absence of amplificons in the same panel where the negative control failed and positive control amplified. Based on this evidence, the procedural and handling defects were ruled-out as soucrces of PCR-failure, indicating that the absence of amplicons corresponded to null alleles. On an average, each of the 618 genotypes had two null alleles, constituting 4.76% of the total 42 expected amplicons per genotype. There were, on an average, 28 null alleles per locus, constituting 4.53% of total 618 expected alleles per locus.

4.1.1.2.2 Analysis of unique alleles

Details of unique alleles, i.e., alleles of 42 SSR loci (Tables 3.3, Table 3.4 and Table 3.5) that are found only in any one of 618 Indian wheat accessions, are given in Table 4.4.

4.1.1.2.2.1 Unique alleles per locus

The list of marker loci, number of unique alleles and allele size in base pairs and the accession number in which such allele(s) is(are) present is given in Tables 4.4. Among 42 SSR markers loci surveyed in a panel of 618 Indian wheat landrace accessions, 35 loci showed unique alleles, and the number of alleles per locus ranged from 4 to 0. The maximum of four unique alleles per locus was exhibited by two loci: xgwm2 and xgwm443. Two marker loci, xgwm497 and xgwm169, revealed three unique alleles each. Two alleles per locus was detected in 14 out of 42 marker loci

Table 4.4 Number of unique alleles, their sizes and genotypes carrying them as revealed by 42 SSR markers loci amplified across 618 Indian wheat accessions.

	S. No.	SSR Locus	No of unique alleles	Allele size in base pairs (Accession number)
Subgenome-A	1	Xgwm136	1	275 (IC-532262)
	2	Xgwm497	3	133 (IC-118758), 160 (IC-82370) and 185 (IC-82440)
	3	Xgwm448	2	321 (IC-212145) and 357 (IC-30284)
	4	Xgwm445	2	192 (IC-73210) and 326 (IC-30276)
	5	Xgwm2	4	148 (IC-212185), 160 (IC-212145), 198 (IC-30276) and 206 (IC-31496)
	6	Xgwm494	2	276 (IC-118758) and 312 (IC-532149)
	7	Xgwm165	2	162 (IC-212185) and 232 (IC-31496)
	8	Xgwm637	2	182 (IC-57983) and 222 (IC-31496)
	9	Xgwm443	4	202 (IC-57983), 236 (IC-31496), 350 (IC-82425) and 370 (IC-82388)
	10	Xgwm666	2	164 (IC-57983) and 190 (IC-31496)
	11	Xgwm459	2	158 (IC-57983) and 196 (IC-31496)
	12	Xgwm169	3	164 (IC-212145), 172 (IC-212185), and 250 (IC-31496)
	13	Xgwm635	2	160 (IC-57983) and 194 (IC-31496)
	14	Xgwm282	1	208 (IC-57983)
	15	Xgwm268	1	238 (IC-57983)
Subgenome-B	16	Xgwm257	2	238 (IC-57983) and 268 (IC-31496)
	17	Xgwm47	1	186 (IC-57983)
	18	Xgwm533	1	186 (IC-57983)
	19	Xgwm181	1	194 (IC-118758)
	20	Xgwm368	1	194 (IC-57983)
	21	Xgwm513	2	178 (IC-57983) and 234 (IC-31496)
	22	Xgwm544	1	320 (IC-57983)
	23	Xgwm335	2	196 (IC-57983) and 252 (IC-31496)
	24	Xgwm361	1	249 (IC-31496)
	25	Xgwm219	1	248 (IC-118758)
	26	Xgwm400	2	222 (IC-118758) and 258 (IC-31496)
	27	Xgwm131	2	292 (IC-104551) and 304 (IC-82426)
Subgenome-D	28	Xgwm458	1	148 (IC-118758)
	29	Xgwm642	1	258 (IC-118758)
	30	Xgwm183	1	175 (IC-118758)
	31	Xgwm3	1	148 (IC-118758)
	32	Xgwm194	1	171 (IC-118758)
	33	Xgwm624	1	164 (IC-118758)
	34	Xgwm182	1	204 (IC-118758)
	35	Xgwm37	2	168 (IC-212185) and 231 (IC-118758)
		Total	**59**	-

investigated. While majority (17) loci possessed one allele per locus, seven did not show any unique allele.

4.1.1.2.2.2 Distribution of unique alleles at subgenome level

All the 14 SSR loci belonging to subgenome A had at least one unique allele per locus (Table 4.4). Further, A-subgenome had, in its fold, all the loci that showed 4-3 alleles per locus, 8 of 14 loci having 2 unique alleles each, and two loci with one per locus each. Among 14 loci with two alleles per locus, 4 belonged to subgenome-B. Only 8 of 14 SSR marker loci in subgenome-D showed unique alleles ranging from 1-2 allele(s) per locus.

4.1.1.2.2.3 Unique alleles among Indian wheat landraces

Among 618 Indian wheat accessions investigated for 42 SSR loci, 17 genotypes had unique alleles ranging from 14 to 1 allele(s) per genotype (Table 4.4). Indian wheat accession IC-57983 had the highest number of 14 unique alleles in the size-range of 158 bp to 320 bp. The maximum and minimum allele sizes were observed on at two marker loci, *xgwm459* and *xgwm544*, located on the short arms of chromosomes 6 and 5 of subgenome A, respectively. Thirteen unique alleles per genome was observed in the accession IC-31496, where the allele sizes were in the range of 190 bp to 258 bp. The maximum and minimum allele sizes were recorded at loci, *xgwm666* and *xgwm400*, located on long- and short arms, of chromosomes 5 and 7 of subgenomes A and B, respectively. Twelve unique alleles in the range of 113 bp to 276 bp, with the maximum and minimum corresponding to marker loci *xgwm497* and *xgwm494* located on long arms of chromosomes 1 and 3 of subgenome A, respectively, were recorded in the wheat accession IC-118758. Two of the remaining 14 accessions, IC-30276 and IC-212185, possessed two unique alleles each, corresponding to loci *xgwm445* and *xgwm2* in first accession; and, *xgwm497* and *xgwm182* in latter. Twelve of 17 genotypes showed only one unique allele per locus. Four unique alleles: 160 bp, 185 bp, 321 bp and 148 bp; were revealed by 4 loci: Xgwm497, Xgwm185, Xgwm321 and Xgwm497; located on long arm of chromosome 1 of subgenome A of accessions IC-82370, IC-82440, IC-212145 and IC-212185, respectively.

4.1.1.3 Genet diversity among homoeologous groups

The complex genome of wheat has 7 homoeologous groups, each having 3 chromosomes belonging to three distinct subgenomes: AA, BB and DD. The details of

gene diversity among homoeologous chromosomes as revealed by 42 SSR marker loci amplified among 618 Indian wheat accessions is furnished in Table 4.5. Each homoeologous group was investigated for gene diversity deploying 6 SSR markers.

4.1.1.3.1 Polymorphic alleles among homoeologous groups

Each homoeologous group had, on an average, 110 polymorphic alleles for 6 SSR loci, whereas 7 homoeologous group together had 128.35 per locus (Table 4.5). Group 1 had the highest number of 117 polymorphic alleles with a mean of 19.5 alleles per locus, followed by groups 3 and 7 which had 114 and 111 polymorphic alleles with means of 19.0 and 18.5 alleles per locus, respectively. The lowest polymorphic alleles of 105 with an average of 17.5 alleles per locus was observed in the homoeologous group 2. Each of the groups 4 to 6 had 108 polymorphic alleles with a mean of 18 per locus. Each of the 7 homoeologous group had the minimum of 100 alleles for 6 SSR loci. Marker loci in group 1 exhibited the highest number of alleles per locus in the range of 15-26, while the lowest number of 12 alleles per locus was detected in the homoeologous group 6.

4.1.1.3.2 Major alleles among homoeologous groups

On an average, each group had 19.7 major alleles (i.e., amplified in > 5% of 618 genotypes) corresponding to 6 SSR loci, whereas it was 23.7 per locus across the 7 homoeologous groups (Table 4.5). Group 2 had the highest number of 23 major alleles with a mean of 3.83 alleles per locus, followed by groups 4 and 5 which, equally, had 22 major alleles each with a mean of 3.67 alleles per locus. The lowest major alleles of 16 with a mean of 2.67 per locus was observed in the homoeologous group 6, followed by 17 and 19 major alleles with means of 2.83 and 3.17 per locus in homoeologous groups 3 and 1, respectively. When compared across the 7 homoeologous groups, group 2 had the highest number 23 out of 138 major alleles which was above the group average of 19.7 per homoeologous group. However, all the 7 homoeologous groups had 16 major alleles of 42 SSR markers. An average number of major alleles per homoeologous group was ranging from 16 (in group 6) to 23 (in group 2), with three groups: 16 (group 6),17 (group 3) and 19 (group 1) having a little below the group average of 19.7. This indicated that 42 SSR loci had polymorphic alleles that were distributed among all the 7 homoeologous groups, ranging from 16 to 23 per group, and they occurred among more than 5% of the 618 Indian wheat landrace accessions.

Table 4.5 Genetic diversity in 7 homoeologous groups of the Indian wheat accessions as revealed by 42 SSR markers.

Homoeo* groups	Polymorphic alleles			Number of								PIC	GDI
	Total	Mean	Range	Major alleles			Minor alleles			Unique alleles	Null alleles		
				Total	Mean	Range	Total	Mean	Range				
1	117	19.5	15-26	19	3.17	2.55-4.16	97.95	16.33	12.45-21.84	7	292	0.902	0.907
2	105	17.5	13-23	23	3.83	3.25-4.64	82.12	13.69	9.75-19.09	7	85	0.873	0.883
3	114	19.0	15-24	17	2.83	1.76-4.08	97.23	16.21	12.90-20.47	10	182	0.910	0.912
4	108	18.0	15-22	22	3.67	2.72-5.72	86.28	14.38	10.65-17.20	9	135	0.887	0.895
5	108	18.0	13-23	22	3.67	2.25-4.83	86.05	14.34	10.14-18.17	10	93	0.883	0.893
6	108	18.0	12-23	16	2.67	1.8-2.64	92.47	15.41	10.2-20.70	7	207	0.902	0.910
7	111	18.5	13-24	20	3.33	2.34-5.46	91.4	15.23	10.66-19.92	9	183	0.897	0.905
Total	**771**	**128.35**	**-**	**138**	**23.7**	**1.80-5.46**	**633**	**105.5**	**-**	**59**	**1177**	**-**	**-**
Mean	**110**	**18.5**	**-**	**19.7**	**3.31**	**-**	**90.43**	**15.08**	**-**	**8.42**	**168**	**0.893**	**0.901**

*Foot Note: Homoeo: Homoeologous, SSR: Simple sequence repeats, loci: SSR loci, PIC: Polymorphic information content, GDI: Gene diversity index.

4.1.1.3.3 Minor alleles among homoeologous groups

Each group had, on an average, 90.43 minor alleles for 6 SSR loci, whereas there was an average of 105.5 per locus across all the 7 homoeologous groups (Table 4.5). Group 1 had the highest number of 97.95 minor alleles with a mean of 16.33 alleles per locus, followed by groups 3 and 6, which had 97.23 and 92.47 polymorphic alleles with means of 16.21 and 15.41 alleles per locus. The lowest number (82.12) of minor alleles with a mean of 13.69 alleles per locus was observed in the homoeologous group 5 followed by the group 4 that had 86.05 polymorphic alleles with a mean of 14.34 polymorphic alleles per locus. The highest number of 21.84 minor alleles was recorded in the homoeologous group 1, while the lowest number of 9.75 is shown by group 2.

4.1.1.3.4 Unique alleles among homoeologous groups

A total of 59 unique alleles appearing among 17 genotypes were distributed among all the 7 homoeologous groups of Indian wheat accessions (Table 4.5). On an average basis, there were 8.42 unique alleles per homoeologous group, ranging from 7 to 10 each with the maximum recorded in groups 1, 2 and 6; and, the minimum in groups 3 and 5, respectively. Groups 4 and 7 had 9 unique alleles each.

4.1.1.3.5 Null alleles among homoeologous groups

Null alleles amounting to 1177 (4.53% of the alleles sampled; Table 4.3) were distributed across all the homoeologous groups, with an average of 168 per group. The highest number of 292 null alleles were found in group 1, followed by group 6 that had 207 null alleles. Homoeologous group 2 had the lowest number of 85 null alleles, followed by group 5 that showed 93.

4.1.1.3.6 Polymorphism information content (PIC) among homoeologous groups

Each of the seven homoeologous groups had marker loci with high degree of polymorphism information content (PIC) values with an average of 0.893 per group, ranging from 0.910 in groups 1 to 0.873 in group 2 (Table 4.5). The Homoeologous groups 3, 4, 5, 6 and 7 had PIC values of 0.910, 0.887, 0.883, 0.902 and 0.897, respectively. High PIC values distributed among 7 homoeologous groups indicate the informativeness of the 42 SSR markers, deployed in this study, for deciphering genetic diversity in studied set of 618 Indian wheat accessions.

4.1.1.3.7 Gene diversity among homoeologous groups

Shannon-wiener method-based gene diversity analyzed in this study was high among all the 7 homoeologous groups of 618 Indian wheat landrace accessions (Table 4.5). The comparative analysis of gene diversity among the seven homoeologous groups, each analyzed with 6 marker loci, revealed presence of high level of genetic diversity, with an average diversity index of 0.901, for all the 7 homoeologous groups (Table 4.5). Homoeologous group 3 showed the highest genetic diversity of 0.912, while group 2 had the lowest (0.883). Further, homoeologous groups 3 (0.912), 6 (0.91) and 1 (0.907) are more diverse than the groups 4 (0.895), 5 (0.893), and 2 ((0.883). In addition, the average genetic diversity indices for homoeologous groups 3 (0.912), 6 (0.91), 1 (0.907) and 7 (0.905) were higher than the group average of 0.901, whereas those for 4 (0.895), 5 (0.893), and 2 (0.883) were lower than the group average.

4.1.1.4 Gene diversity among homologous groups

In order to explain the details of allelic information and genetic diversity among 21 homologous groups as revealed by 42 SSR markers among 618 Indian wheat landraces, total, mean and range of polymorphic alleles; the number of major-, minor, unique and null allele; polymorphism information content (PIC); and, gene diversity indices (GDIs) are presented in Table 4.6. Each of the homologous group was investigated with 2 SSR marker loci, one each on their short- and long arms. Considering all the 21 homologous chromosomes, total polymorphic alleles ranged from 28 (7D and 5D) to 51 (1A) for two marker loci (Table 4.6); average allele-richness for 2 marker loci vared from 14.0 (5D and 7D) to 25.5 (1A); allele sizes were in the range of 104 bp (1A, 1D, 7B and 7D) to 422 bp (5B); there were 4 (6D) to 10 (7A) major alleles, with an average major-allele richness of 6.62 per homologous chromosome; minor alleles ranged from 21 (6D) to 43 (1A), with overall minor-allele richness of 30.10 per homologous chromosome; there were zero (2D and 6D) to six (3A and 5A) unique alleles, with an overall unique-allelic richness of 2.81 per homologous chromosome; null-alleles ranged from 20 (2B) to 170 (1B) with exception of chromosome 5B that had no null-allele; average polymorphism information content (PIC) for 2 alleles varied between 0.835 (2D) and 0.920 (2A, 3A and 6A); and, gene diversity indices ranged from 0.855 (2D) to 0.925 (6A).

Table 4.6 Details of allelic information and genetic diversity among 21 homologous groups as revealed by 42 SSR markers amplified across 618 Indian wheat landraces.

S. No.	Hom. Chr*.	No. of Mr. loci	Polymorphic alleles			Number of				PIC	GDI
			Total	Mean	Range of allele sizes (bp)	Major alleles	Minor alleles	Unique alleles	Null alleles	Mean	Mean
1	1A	2	51	25.5	104-321	8	43	4	82	0.910	0.915
2	1B	2	35	17.5	150-350	6	29	1	170	0.895	0.900
3	1D	2	31	15.5	104-378	5	26	2	40	0.900	0.905
4	2A	2	45	22.5	130-375	8	37	4	25	0.920	0.920
5	2B	2	33	16.5	128-310	8	25	3	20	0.865	0.875
6	2D	2	27	13.5	140-320	7	20	0	40	0.835	0.855
7	3A	2	47	23.5	124-332	7	40	6	117	0.920	0.920
8	3B	2	36	18.0	108-282	6	30	2	25	0.900	0.900
9	3D	2	31	15.5	110-336	4	27	2	40	0.910	0.915
10	4A	2	42	21.0	144-320	9	33	4	63	0.890	0.900
11	4B	2	35	17.5	124-292	6	29	3	32	0.885	0.895
12	4D	2	31	15.5	109-352	7	24	2	40	0.885	0.890
13	5A	2	45	22.5	114-370	9	36	6	53	0.895	0.905
14	5B	2	35	17.5	128-422	8	27	3	0	0.865	0.875
15	5D	2	28	14.0	154-382	5	23	1	40	0.890	0.900
16	6A	2	45	22.5	114-354	5	40	5	101	0.920	0.925
17	6B	2	38	19.0	118-368	6	32	2	66	0.900	0.910
18	6D	2	25	12.5	109-256	4	21	0	40	0.885	0.895
19	7A	2	45	22.5	108-368	10	35	3	91	0.895	0.900
20	7B	2	38	19.0	104-304	5	33	4	52	0.905	0.915
21	7D	2	28	14.0	104-412	5	23	2	40	0.890	0.900
	Total	**42**	**771**	**385.5**	**104-422**	**139**	**632**	**59**	**1177**	**-**	**-**
	Mean	**2**	**36.74**	**18.37**	**-**	**6.62**	**30.10**	**2.81**	**56.05**	**0.893**	**0.901**

*Note: Hom: Homologous, Chr.: chromosome, Mr.: Marker, MAF: Major allele frequency, PIC: Polymorphic information content.

In addition, when compared to the mean genetic diversity index of 0.893 for all the 21 chromosomes, most of the homologous groups had the genetic diversity indices above average, while it was below the average for eight chromosomes: 2B, 2D, 4A, 4B, 5B, 5D,6D and 7D. However, all the 21 chromosomes had higher level of genetic diversity, indicating that a panel of 618 Indian wheat landrace accessions had a high genetic diversity, and that the 42 SSR markers and computational algorithms used in this study were highly informative to decipher the extent and pattern of diversity.

4.1.1.5 Genetic diversity at subgenome level

As each of the three sub- or component-genomes exhibits homoeologous identity, it was pertinent to understand gene- and genetic diversity at subgenome levels too. Genetic diversities at 3 subgenome (AA, BB and DD) levels were analysed with respect to the following features: major allele frequency; total, mean and range of polymorphic alleles; number of major-, minor-, unique- and null alleles; polymorphism information content (PIC) values; and gene diversity indices (GDIs).

4.1.1.5.1 Subgenome-A

Major allele frequency; total, mean and range of polymorphic alleles; number of major-, minor-, unique- and null alleles; polymorphism information content (PIC) values; and gene diversity indices (GDIs) for subgenome-A are tabulated in Table 4.7.

Seven homologous chromosomes of subgenome-A, analyzed with 14 marker loci in a set of 618 genotypes, revealed that major allele frequencies (MAF) ranged from 0.12 (xgwm169 on 6-AL) to 0.16 (xgwm136 on 1-AS); there were 21 (xgwm635 on 7-AS) to 26 (xgwm136 on 1-AS) total polymorphic alleles with a mean of 22.86 per locus; allele sizes ranged from 104 bp (xgwm497 on 1-AL) to 375 bp (xgwm448 on 2-AS), with the widest spectrum of allele-sizes, 104-321 bp, with the maximum size detected in xgwm497 locus located on 1-AL; there were 2.3 (xgwm459 on 6-AS) to 5.46 (xgwm635 on 1-AS) major alleles, with an average of 3.9 per locus; minor alleles ranged from 15.54 (xgwm635 on 1-AS) to 21.84 (xgwm136 on 1-AS) with a mean of 18.96 per locus; there were 1 (xgwm136 on 1-AS and xgwm282 on 7-AL) to 4 (xgwm2 on 3-AS and xgwm443 on 5-AS) unique alleles with a mean of 2.29 per locus; null alleles ranged from 3 (xgwm448 on 2-AS) to 82 (xgwm497 on 1-AL) with an exception of the locus xgwm136 (on 1-AS) that had no null allele; polymorphism information content (PIC) values ranged between 0.87 (xgwm137 on 4-AL) and 0.93 (xgwm459 on

Table 4.7 Details of allelic information and gene diversity of subgenome-A as revealed by 14 SSR markers loci amplified among 618 wheat landrace genotypes.

S. No.	Hom. Chr*.	Mr. Name Used	MAF (<5%	Number of Polymorphic Alleles						PIC	GDI
				Total	Range of allele size (bp)	Major alleles	Minor alleles	Unique alleles	Null alleles		
1	1-AS	Xgwm136	0.16	26.00	236-321	4.16	21.84	1	0	0.91	0.91
2	1-AL	Xgwm497	0.15	25.00	104-321	3.75	21.25	3	82	0.91	0.92
3	2-AS	Xgwm448	0.17	23.00	201-375	3.91	19.09	2	3	0.92	0.92
4	2-AL	Xgwm445	0.19	22.00	130-326	4.18	17.82	2	22	0.92	0.92
5	3-AS	Xgwm2	0.17	24.00	124-304	4.08	19.92	4	59	0.91	0.91
6	3-AL	Xgwm494	0.11	23.00	138-332	2.53	20.47	2	58	0.93	0.93
7	4-AS	Xgwm165	0.14	20.00	150-320	2.8	17.2	2	26	0.91	0.92
8	4-AL	Xgwm637	0.26	22.00	144-316	5.72	16.28	2	37	0.87	0.88
9	5-AS	Xgwm443	0.21	23.00	150-370	4.83	18.17	4	23	0.89	0.90
10	5-AL	Xgwm666	0.19	22.00	114-300	4.18	17.82	2	30	0.90	0.91
11	6-AS	Xgwm459	0.10	23.00	114-326	2.3	20.7	2	55	0.93	0.93
12	6-AL	Xgwm169	0.12	22.00	164-354	2.64	19.36	3	46	0.91	0.92
13	7-AS	Xgwm635	0.26	21.00	108-302	5.46	15.54	2	54	0.88	0.88
14	7-AL	Xgwm282	0.17	24.00	160-368	4.08	19.92	1	37	0.91	0.92
		Total	**-**	**320.0**	**104-375**	**54.62**	**265.38**	**32.0**	**532.0**	**-**	**-**
		Mean	**0.17**	**22.86**	**-**	**3.90**	**18.96**	**2.29**	**38**	**0.907**	**0.912**
		(%)	**-**	**41.5**	**-**	**39.72**	**41.89**	**54.23**	**45.43**	**-**	**-**

*Note: Hom: Homologous, Chr.: chromosome, Mr.: Marker, MAF: Major allele frequency, PIC: Polymorphic information content, GDI: Genetic diversity index, %: percent of whole genome (ABD).

6-AS); and, gene diversity indices varied from 0.88 (xgwm637 on 4-AL and xgwm635 on 7-AS) to 0.93 (xgwm494 on 3-AL and xgwm459 on 6-AS).

In the subgenome A, there were a total of 320 polymorphic-, 54.62 major-, 265.38 minor-, 32 unique- and 532 null-alleles, that make 41.5 %, 39.72 %, 41.89 %, 54.23 % and 45.43 % of those recorded in the total genome (ABD), respectively). The average PIC value of 14 SSR marker loci located in subgenome-A was 0.907, while average genetic diversity index (GDI) was 0.912. With respect to two SSR marker loci on each chromosome, both the short- and long-arms of chromosomes 2 and 6 of subgenome-A were more diverse than the group average. Similarly, short arm of chromosome 4 and long arms of chromosomes 1, 3 and 7 were also more diverse than the average diversity of subgenome-A.

4.1.1.5.2 Subgenome-B

Major allele frequency; total, mean and range of polymorphic alleles; number of major-, minor-, unique- and null alleles; polymorphism information content (PIC) values; and gene diversity indices (GDIs) for subgenome-B are listed in Table 4.8.

For all the 14 SSR marker loci surveyed among 7 homologous chromosomes of subgenome-B, major allele frequencies (MAF) ranged from 0.14 (xgwm361 on 6-BS, xgwm400 on 7-BS and xgwm131 on 7-BL) to 0.29 (xgwm257 on 2-BS); total number of polymorphic alleles were in the range of 15 (xgwm550 on 1-BS) to 20 (xgwm268 on 1-BL, xgwm219 on 6-BL and xgwm131 on 7-BL), with a mean of 17.86 per locus; allele sizes ranged from 104 bp (xgwm131 on 7-BL) to 422 bp (xgwm544 on 5-BS), with the widest spectrum of allele-sizes, 148-368 bp, detected in xgwm219 locus located on 6-BL; there were 2.52 (xgwm361 on 6-BS and xgwm400 on 7-BS) to 4.64 (xgwm257 on 2-BS) major alleles with an average of 3.31 per locus; minor alleles ranged from 11.36 (xgwm257 on 2-BS) to 17.20 (xgwm131 on 7-BL) with a mean of 14.55 per locus; all the chromosomes had at least one unique allele, except at xgwm550 (1-BS) locus that had no unique allele; null alleles ranged from 17 (xgwm131 on 7-BL) to 106 (xgwm550 on 1-BS) with an exception of five loci (xgwm47 on 2-BL, xgwm181 on 3-BL, xgwm368 on 4-BS, xgwm544 on 5-BS, and xgwm335 on 5-BL) that did not have any null-allele; polymorphism information content (PIC) values ranged between 0.84 (xgwm544 on 5-BS) and 0.91 (xgwm533 on 3-BS and xgwm131 on 7-BL), while

Table 4.8 Details of allelic information and gene diversity of subgenome-B as revealed by 14 SSR markers loci amplified among 618 wheat landrace genotypes.

S. No.	Hom. Chr*.	Mr. Name Used	MAF (<5%)	Number of Polymorphic Alleles						PIC	GDI
				Total	Range of allele size (bp)	Major alleles	Minor alleles	Unique alleles	Null alleles		
1	1-BS	Xgwm550	0.17	15.00	150-300	2.55	12.45	0	106	0.90	0.90
2	1-BL	Xgwm268	0.19	20.00	184-350	3.8	16.2	1	64	0.89	0.90
3	2-BS	Xgwm257	0.29	16.00	174-310	4.64	11.36	2	20	0.85	0.86
4	2-BL	Xgwm47	0.20	17.00	128-278	3.4	13.6	1	0	0.88	0.89
5	3-BS	Xgwm533	0.18	18.00	108-282	3.24	14.76	1	25	0.91	0.91
6	3-BL	Xgwm181	0.17	18.00	120-276	3.06	14.94	1	0	0.89	0.89
7	4-BS	Xgwm368	0.17	17.00	128-280	2.89	14.11	1	0	0.88	0.89
8	4-BL	Xgwm513	0.18	18.00	124-292	3.24	14.76	2	32	0.89	0.90
9	5-BS	Xgwm544	0.27	17.00	236-422	4.59	12.41	1	0	0.84	0.85
10	5-BL	Xgwm335	0.18	18.00	128-328	3.24	14.76	2	0	0.89	0.90
11	6-BS	Xgwm361	0.14	18.00	118-308	2.52	15.48	1	34	0.90	0.91
12	6-BL	Xgwm219	0.19	20.00	148-368	3.8	16.2	1	32	0.90	0.91
13	7-BS	Xgwm400	0.14	18.00	134-312	2.52	15.48	2	35	0.90	0.91
14	7-BL	Xgwm131	0.14	20.00	104-304	2.8	17.2	2	17	0.91	0.92
		Total	**-**	**250.0**	**104-422**	**46.29**	**203.71**	**18.0**	**365.0**	**-**	**-**
		Mean	**0.19**	**17.86**	**-**	**3.31**	**14.56**	**1.29**	**26.07**	**0.887**	**0.895**
		(%)	**-**	**32.41**	**-**	**33.67**	**32.15**	**30.5**	**31.67**	**-**	**-**

*Note: Hom: Homologous, Chr.: chromosome, Mr.: Marker, MAF: Major allele frequency, PIC: Polymorphic information content, GDI: Genetic diversity index, %: percent of whole genome (ABD).

gene diversity indices varied from 0.85 (xgwm544 on 5-BS) to 0.92 (xgwm131 on 7-BL).

Subgenome-B had a total of 250 polymorphic-, 46.29 major-, 203.71 minor-, 18 unique- and 365 null-alleles, that constituted 32.4 %, 33.67 %, 32.15 %, 30.5 % and 31.67 % of those detected in the total genome (ABD), respectively. The average PIC value for 14 SSR marker loci analysed in subgenome-B was 0.887, while an average genetic diversity index (GDI) was 0.895. Therefore, 89.5% of genetic diversity of subgenome-B, with respect to the 14 loci, was revealed by 14 SSR markers amplified among 618 Indian wheat landrace accessions. With respect to two SSR marker loci on each chromosome, both the short- and long-arms of chromosomes 1, 6 and 7 of subgenome-B were more diverse than the group average. Similarly, short arm of chromosome 3 and long arms of chromosomes 4 and 5 were also more diverse than the average diversity of subgenome-B.

4.1.1.5.3 Subgenome-D

Major allele frequency; total, mean and range of polymorphic alleles; number of major-, minor-, unique- and null alleles; polymorphism information content (PIC) values; and gene diversity indices (GDIs) for subgenome-D are presented in Table 4.9.

Short- and long arms of all the seven homologous chromosomes in subgenome-D were assessed for the genetic diversity deploying one SSR locus each. Upon analyzing all the 14 loci among seven homologous chromosomes of subgenome-D, it was found that major allele frequencies (MAF) ranged from 0.11 (xgwm183 on 3-DS) to 0.29 (xgwm624 on 4-DL); there were 12 (Xgwm141 on 6-DS) to 16 (Xgwm642 on 1-DL, xgwm183 on 3-DS and xgwm194 on 4-DS) total polymorphic alleles with a mean of 14.35 per locus; allele sizes ranged from 104 bp (xgwm458 on 1-DS) to 412 bp (xgwm37 on 7-DL), with the widest spectrum of allele-sizes, 168-412 bp, also detected in xgwm37 locus located on 7-DL; there were 1.76 (xgwm183 on 3-DS) to 4.35 (xgwm624 on 4-DL) major alleles, with an average of 2.61 per locus; minor alleles ranged from 9.75 (xgwm349 on 2-DL) to 13.76 (xgwm642 on 1-DL) with a mean of 11.74 per locus; on chromosome 7-DL there were 2 unique alleles, while most of the chromosomes had one each, six chromosomes (2-DS, 2-DL, 5-DS, 6-DS, 6-DL and 7-DS) did not have any unique alleles, and a mean of 0.64 per locus was detected for all the 14 chromosomes; null alleles were uniformly 21 in each chromosome, contributed

Table 4.9 Details of allelic information and gene diversity of subgenome-D as revealed by 14 SSR marker loci amplified among 618 wheat landrace genotypes.

S. No.	Hom. Chr*.	Mr. Name Used	MAF (<5%)	Number of Polymorphic Alleles						PIC	GDI
				Total	Range of allele size (bp)	Major alleles	Minor alleles	Unique alleles	Null alleles		
1	1-DS	Xgwm458	0.17	15.00	104-248	2.55	12.45	1	21	0.89	0.90
2	1-DL	Xgwm642	0.14	16.00	214-378	2.24	13.76	1	21	0.91	0.91
3	2-DS	Xgwm484	0.25	14.00	140-278	3.5	10.5	0	21	0.82	0.84
4	2-DL	Xgwm349	0.25	13.00	198-320	3.25	9.75	0	21	0.85	0.87
5	3-DS	Xgwm183	0.11	16.00	114-336	1.76	14.24	1	21	0.92	0.92
6	3-DL	Xgwm3	0.14	15.00	110-252	2.1	12.9	1	21	0.90	0.91
7	4-DS	Xgwm194	0.17	16.00	109-352	2.72	13.28	1	21	0.91	0.91
8	4-DL	Xgwm624	0.29	15.00	120-252	4.35	10.65	1	21	0.86	0.87
9	5-DS	Xgwm190	0.15	15.00	218-382	2.25	12.75	0	21	0.90	0.91
10	5-DL	Xgwm182	0.22	13.00	154-292	2.86	10.14	1	21	0.88	0.89
11	6-DS	Xgdm141	0.15	12.00	132-242	1.8	10.2	0	21	0.88	0.89
12	6-DL	Barc1121	0.19	13.00	109-256	2.47	10.53	0	21	0.89	0.90
13	7-DS	Xgwm44	0.18	13.00	128-280	2.34	10.66	0	21	0.88	0.89
14	7-DL	Xgwm37	0.16	15.00	168-412	2.4	12.6	2	21	0.90	0.91
		Total	**-**	**201.0**	**104-412**	**36.59**	**164.41**	**9.0**	**280.0**	**-**	**-**
		Mean	**0.18**	**14.35**	-	**2.61**	**11.74**	**0.64**	**20.0**	**0.89**	**0.89**
		(%)	**-**	**26.07**	-	**26.61**	**25.95**	**15.25**	**22.9**	**-**	**-**

*Note: Hom: Homologous, Chr.: chromosome, Mr.: Marker, MAF: Major allele frequency, PIC: Polymorphic information content, GDI: Genetic diversity index, %: percent of whole genome (ABD).

by of 21 tetrapoids that lacked D-genome, suggesting that subgenome-D had no null alleles among 597 hexaploids; polymorphism information content (PIC) values ranged between 0.82 (xgwm484 on 2-DS) and 0.92 (xgwm183 on 3-DS); and, gene diversity indices (GDI) varied from 0.84 (xgwm183 on 3-DS) 0.91 (xgwm37 on 7-DL).

In total, subgenome-D had 201 polymorphic-, 36.59 major-, 164.41 minor-, 9 unique- and 294 null-alleles, that comprised 26.07 %, 26.61%, 25.95%, 15.25% and 22.9% of those detected in the whole genome (ABD), respectively. Both average PIC value as well as average GDI for 14 SSR marker loci across subgenome-D was 0.89. Thus, 14 SSR markers explained 89% of the genetic diversity of subgenome-D, with respect to 14 loci, in a panel of 618 Indian wheats.

With respect to two SSR marker loci on each chromosome, both the short- and long-arms of chromosomes 2 and 3 of subgenome-D were more diverse than the group average. Similarly, short arms of chromosomes 4 and 5, and long arms of chromosomes 6 and 7 were also more diverse than the average diversity of subgenome-D.

4.1.1.6 Comparison of molecular genetic diversity at subgenome and genome levels

Since three subgenomes had distinct allelic profiles, and extent- and pattern of gene- and genetic diversities (Table 7, Table 8 and Table 9), it was prudent to compare the diversity of subgenomes A, B and D. The subgenome-wise details: *viz.*, total, mean and range of polymorphic alleles; mean and range of polymorphism information content (PIC); mean and range of genetic diversity; and, mean and range of major allele frequencies are summarized in Table 4.10.

Even though 42 SSR marker loci were evenly distributed among three subgenomes, their polymorphic alleles were not. Among a total of 771 polymorphic alleles belonging to 42 SSR loci amplified among 618 hexaploid genomes, the highest (320) polymorphic alleles belonged to 14 loci in subgenome-A; followed by 250 and 201 belonging to 14 loci each in subgenomes B and D, respectively (Table 4.10). Similarly, subgenome-A had the highest genetic richness as indicated by 22.857 polymorphic alleles per chromosome, across 14 loci; followed by subgenomes B and D having17.857 and 14.357 alleles per chromosome, respectively. In addition, subgenome-A exhibited the widest range of 20-26 alleles per locus; followed by subgenomes B and D, where the number of polymorphic alleles were in the range, of 15-20 and 12-26 per locus, respectively. The major allele frequency (MAF) was highest

Table 4.10 Genetic diversity at subgenome and genome levels of a set of 618 ndian wheat landraces as revealed by 42 SSR markers.

Subgenome/ Genome	Number of polymorphic alleles			MAF* (>5%)		Polymorphism (PIC)		Genetic Diversity	
	Total	Mean	Range	Mean	Range	Mean	Range	Mean	Range
A	320	22.857	20-26	0.176	0.88-0.93	0.907	0.87-0.93	0.912	0.88-0.93
B	250	17.857	15-20	0.187	0.85-0.92	0.887	0.84-0.91	0.896	0.85-0.92
D	201	14.357	12-16	0.182	0.11-0.92	0.885	0.82-0.92	0.890	0.84-0.92
Whole genome (ABD)	**771**	**257**	**12-26**	**0.180**	**0.11-0.93**	**0.893**	**0.82-0.93**	**0.901**	**0.84-0.93**

*Note: MAF= Major allele frequency, PIC: polymorphic information content, (>5%): Occurring in more than 5 % of the genotypes.

in the subgenome-B (0.187), followed by subgenome-D (0.182), and subgenome-A (0.176) in the ranges, of 0.85-0.92, 0.11-0.92 and 0.88-0.93, respectively; and, the overall MAF for the hexaploid genome ranged from 0.11-0.93 with a mean of 0.180.

The highest polymorphism was found among the SSR loci of subgenome-A; as revealed by mean PIC values varying from 0.87-0.93 per locus with a mean of 0.907 per subgenome; followed by PIC values ranging from 0.84-0.91 and 0.82-0.92, with means of 0.887 and 0.885 per locus, respectively. The overall polymorphism (PIC) of the 618 hexaploid wheat genome varied from 0.82-0.93, with a mean of 0.893 per SSR locus across 618 genotypes.

The subgenome-A had the highest genetic diversity (0.912) ranging from 0.88-0.93, followed by subgenomes B (0.895) and D (0.890) having 0.85-0.92 and 0.84-0.92, respectively.

4.1.2 Genetic diversity analyses

4.1.2.1 Cluster analysis

A simple agglomerative (bottom-up) hierarchical clustering method was used to generate unweighted pair group method with arithmatic mean (UPGMA) based tree. The resultant UPGMA tree-based dendrogram is depicted in Figure 4.3 and Figure 4.4 and clustering of the wheat accessions into different groups is detailed in Table 4.11.

The clustering of 618 Indian wheat accessions is illustrated in Figure. 4.3 and the details of the clustering and membership in each major group, group and subgroup along with alphanumeric codes are detailed in Table 4.11, and depicted pictorially in the dendrogram displayed in Figure 4.4. Owing to the larger number of genotypes involved in the study and the difficulty associated with depicting the entire cluster at one place, the clusters were assigned alpha-numeric symbols in dendrogram (Figure. 4.4) and corresponding group membership list (Table 4.11). Most (510) of the studied set of accessions belonged to the category “unknown” and a few (108) had the information about their geographical origin (Table 3.2). These 108 accessions belonged to eight geopolitical states of India, *viz*., Gujarat (2 accessions), Himachal Pradesh (18), Jammu and Kashmir (3), Karnataka (3), Madya Pradesh (2), Maharashtra (9), Rajasthan (27) and Uttar Pradesh (44). For easy identification, analysis and discussion, unique Indigenous Collection (IC) numbers of all the accessions were colour-coded (Table 4.11) as follows: dark red: Gujarat, red: Rajasthan, yellow: Himachal Pradesh, light

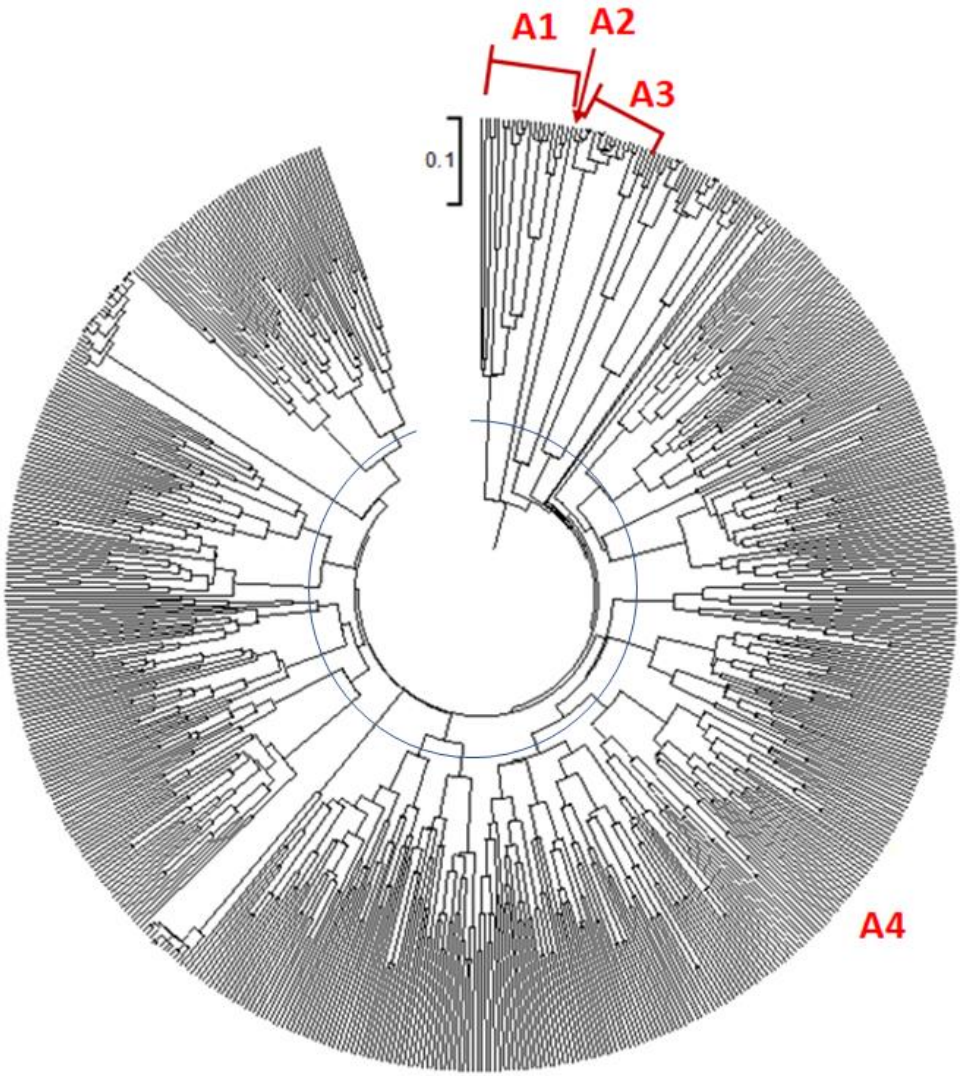


Figure 4.3 UPGMA-tree-based dendrogram displaying grouping of 618 Indian wheat landrace accessions into 4 clads (or major groups) A1-A4. Blue colored line indicates an incision line drawn at genetic dissimilarity coefficient 0.7 resulting in 14 clusters (or groups); Clusters in this dendrogram correspond to pictorially resolved and alpha-numerically coded clusters (or groups) shown in Figure 4.4 and detailed in Table 4.11.

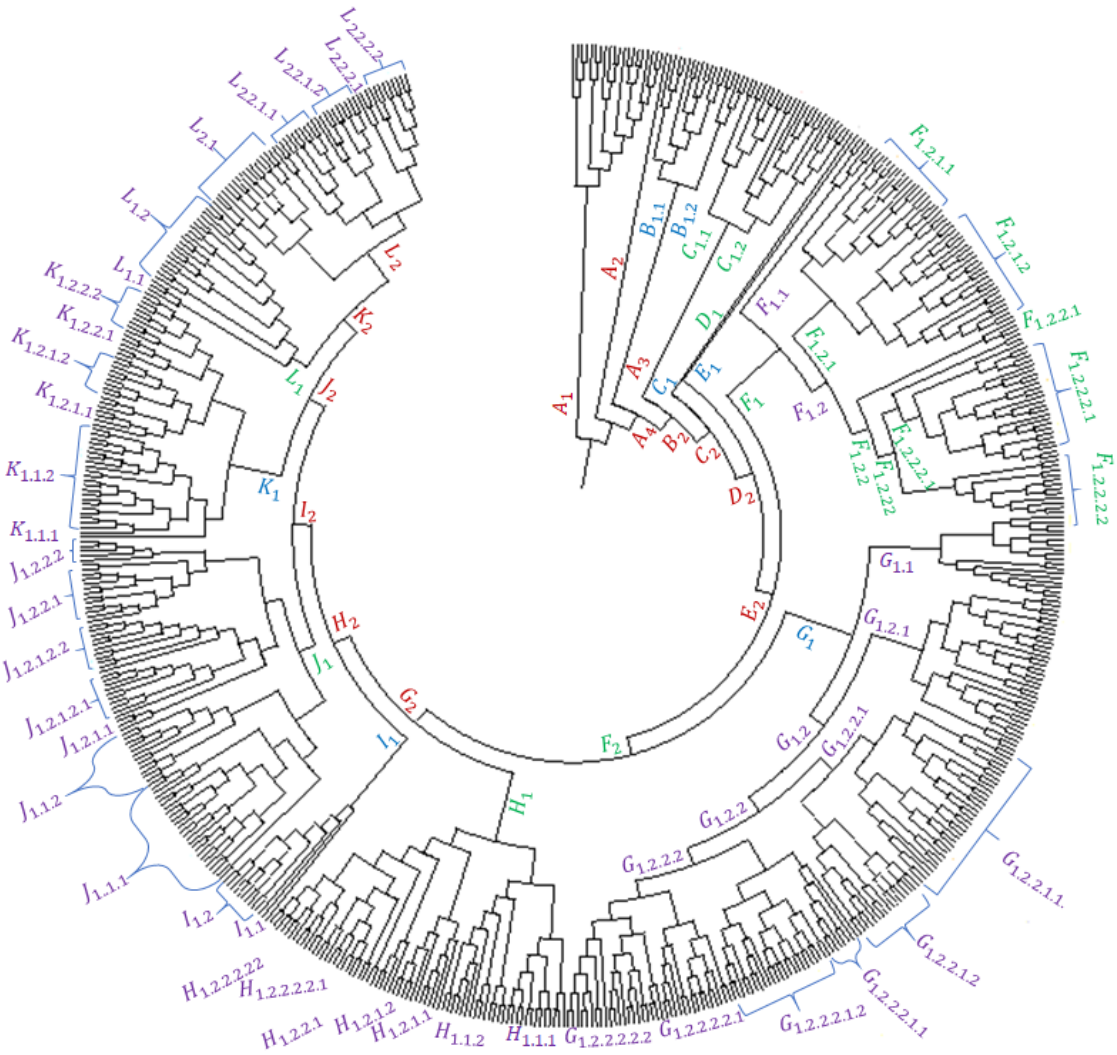


Figure 4.4 The resolved UPGMA tree-based dendrogram derived from genetic dissimilarity index of 0.7 as revealed by 42 SSR marker loci amplified across 618 Indian wheat landrace accessions. Alpha-numeric codes correspond to clusters (or groups) detailed in Table 4.11.

Table 4.11 Accessions with membership in different clads, clusters, major groups, or groups according to dendrogram (Figure 4.4).

Codes*			Accessions with membership in respective clads/clusters/subclusters or Major groups/groups/subgroups
Clad/ Major group	**Cluster/ Group**	**Subcluster/ Sub-group**	
A_1			IC-532887, IC-78937, IC-32520, IC-35070, IC-212145, IC-28974, IC-73232, IC-64238, IC-73205, IC-73198, IC-73201, IC-31496, IC-73493, IC-75327, IC-75328, IC-75339, IC-78933, IC-78877, IC-78096, and IC-78860
A_2			IC-118758
A_3	B_1	$B_{1.1}$	IC-78987, IC-78982, IC-78968, IC-78981, IC-78990, IC-79056, IC-79091, IC-79090, IC-79077, IC-79085, IC-78996, IC-79065, and IC-79068
		$B_{1.2}$	IC-138898, IC-78991, IC-82338, IC-82286, and IC-82303
A_4	C_1	$C_{1.1}$	IC-28958, IC-28969, IC-82136, IC-82145, IC-82125, IC-82126, IC-82133, IC-82134, and IC-82135
		$C_{1.2}$	IC-82132, IC-82146, IC-82179, IC-82157, IC-82147, IC-82159, IC-82160, IC-82156, IC-82158, IC-28664, IC-212179, IC-210548, IC-212142, IC-104540, and IC-104543,
	D_1		IC-30276
	E_1		IC-212185
	$F_{1.1}$		IC-73207, IC-73210, IC-118727, IC-118723, and IC-118730
	$F_{1.2.1}$	$F_{1.2.1.1}$	IC-138893, IC-532473, IC-532475, IC-532229, IC-532245, IC-82165, IC-79100, IC-79110, IC-79940, IC-532141, IC-532276, IC-532281, IC-532282, IC-532185, IC-532209, IC-532216, IC-532204, and IC-532218
		$F_{1.2.1.2.}$	IC-532110, IC-532074, IC-532116, IC-28926, IC-47478, IC-532055, IC-104568, IC-532133, IC-532085, IC-104564, IC-532058, and IC-532072
			IC-82118, IC-79053, IC-79055, IC-79026, IC-79067, IC-82394, IC-82386, and IC-82398
			IC-534566, IC-532268, IC-78940, IC-532483, IC-532490, IC-532238, IC-532243, and IC-532244
$F_{1.2.2}$	$F_{1.2.2.1}$		IC-57983, IC-29035, and IC-30284
	$F_{1.2.2.2.1.1}$		IC-82193, and IC-82236,
$F_{1.2.2.2}$	$F_{1.2.2.2.1.2.1}$		IC-82116, IC-534750, IC-534752, IC-534741, and IC-534744,
	$F_{1.2.2.2.1.2.2.1}$		IC-532480, IC-532093, IC-532088, IC-532095, IC-532274, IC-532737, IC-534739, IC-534605, IC-534737, IC-534697, and IC-534736
	$F_{1.2.2.2.1.2.2.2}$		IC-532083, IC-532699, IC-138849, IC-532474, IC-532723, IC-55657, IC-532219, IC-532700, IC-532077, and IC-79047
	$F_{1.2..2.2}$		IC-532406, IC-78931, IC-532775, IC-532794, IC-532807, IC-532815, IC-532832, IC-532833, IC-532851, IC-532497, IC-532476, and IC-532482

Contd…

Table 4.11 Contd…

$G_{1.1}$		IC-532290, IC-532481, IC-532502, IC-532697, IC-82426, IC-82421, IC-82432, IC-534871, IC-534858, IC-534864, IC-118740, IC-532318, IC-104573, IC-534766, IC-534751, IC-534745, IC-534742, and IC-534748
$G_{1.2.1}$		IC-534363, IC-534419, IC-534435, IC-534509, IC-534564, IC-534855, IC-532813, IC-534820, IC-532768, IC-532784, IC-532839, IC-534721, IC-79046, IC-534857, IC-79095, IC-82177, IC-82180, IC-82256, IC-82367, IC-532059, IC-138896, IC-138897, IC-532061, IC-532842, IC-532853, IC-532841, IC-82190, IC-82195, IC-82206, IC-532790, IC-82257, IC-104537, IC-118729, and IC-118741
$G_{1.2.2.1}$	$G_{1.2.2.1.1}$	IC-82263, IC-82285, IC-79043, IC-78999, IC-79041, IC-532698, IC-532704, IC-534786, IC-4792, IC-534562, IC-532725, IC-532872, IC-534773, IC-534776, IC-534767, IC-534768, IC-534770, IC-532285, IC-532292, IC-532297, IC-532309, IC-532241, IC-532250, IC-532251, IC-532256, IC-532486, IC-82247, IC-532257, IC-534720, IC-532489, IC-532258, IC-532273, IC-532271, and IC-532284
	$G_{1.2.2.1.2}$	IC-82163, IC-82164, IC-82181, IC-82189, IC-78728, IC-78832, IC-79052, IC-82127, IC-82161, IC-82185, IC-82187, IC-82372, IC-82375, IC-82200, IC-82210, IC-82342, and IC-82381
$G_{1.2.2.2}$	$G_{1.2.2.2.1.1}$	IC-532068, IC-532240, IC-118736, IC-118737, IC-532210, IC-532231, and IC-532237
	$G_{1.2.2.2.1.2}$	IC-534197, IC-534323, IC-534572, IC-534688, IC-534723, IC-534743, IC-532261, IC-532263, IC-532267, IC-532289, IC-534747, IC-534749, IC-534754, IC-534757, IC-534884, IC-78707, IC-73215, IC-79063, IC-533965, IC-534193, IC-533966, IC-532835, and IC-532837
$G_{1.2.2.2.2}$	$G_{1.2.2.2.2.1}$	IC-534760, IC-534761, IC-534777, IC-534765, IC-534775, IC-55659, IC-59612, IC-57844, IC-59511, IC-55656, IC-78872, IC-59544, IC-533970, IC-534190, and IC-534886
	$G_{1.2.2.2.2.2}$	IC-532217, IC-532183, IC-532214, IC-532187, IC-532188, IC-532147, IC-532155, IC-532175, IC-35069, IC-532092, IC-532073, IC-532144, IC-82197, IC-532935, IC-533962, IC-534553, IC-534560, IC-534565, IC-78928, IC-534543, and IC-534568
$H_{1.1}$	$H_{1.1.1}$	IC-534451, IC-78923, IC-78924, IC-532863, IC-534322, IC-532930, IC-534609, IC-79039, IC-79042, IC-82387, IC-82393, IC-79062, and IC-82167
	$H_{1.1.2}$	IC-532279, IC-532736, IC-532078, IC-532228, IC-82369, IC-534549, IC-534763, IC-532062, IC-532106, IC-532067, IC-532066, IC-532180, IC-532143, and IC-532201
$H_{1.2.1}$	$H_{1.2.1.1}$	IC-118732, IC-532225, IC-532779, IC-118731, and IC-532094
$H_{1.2.1.2}$	$H_{1.2.1.2.1}$	IC-532084
	$H_{1.2.1.2.2}$	IC-532102, IC-532119, IC-532120, IC-532064, and IC-532098

Contd…

Table 4.11 Contd…

$H_{1.2.2.}$	$H_{1.2.2.1}$	IC-532091, IC-532089, IC-118762, and IC-532090
	$H_{1.2.2.2.1}$	IC-79038, and IC-79040
	$H_{1.2.2.2.2.1}$	IC-534778, IC-534798, IC-534822, IC-78925, IC-79106, IC-82555, IC-534805, IC-534432, IC-534814, IC-534455, IC-534430, and IC-534480
$H_{1.2.2.2}$	$H_{1.2.2.2.2.2}$	IC-534481, IC-534554, IC-534556, IC-532773, IC-532787, IC-534758, IC-534235, IC-534883, IC-534334, IC-534561, IC-104576, and IC-118739
I_1	$I_{1.1}$	IC-75333
	$I_{1.2}$	IC-532097, IC-78839, IC-75351, IC-78715, IC-78824, IC-78859, IC-78865, IC-78828, IC-78837, IC-78897, IC-78836, and IC-78862
$J_{1.1}$	$J_{1.1.1}$	IC-532137, IC-532138, IC-532184, IC-534764, IC-534787, IC-534772, IC-534823, IC-534759, IC-532868, IC-534769, IC-532096, IC-55593, IC-532239, IC-532503, IC-532855, IC-532910, IC-532923, IC-532708, IC-532727, IC-78927, IC-532701, IC-534756, IC-532103, IC-532134, IC-532131, and IC-532478
	$J_{1.1.2}$	IC-534072, IC-534123, IC-533964, IC-533971, IC-532242, IC-532492, IC-533953, IC-532937, IC-532891, IC-532936, IC-532728, and IC-532880
$J_{1.2}$	$J_{1.2.1.1}$	IC-532057, IC-138845, and IC-138895
	$J_{1.2.1.2.1}$	IC-104551, IC-532707, IC-532298, IC-532310, IC-532246, IC-532247, IC-532272, IC-532264, IC-532248, and IC-532262
	$J_{1.2.1.2.2}$	IC-532071, IC-532186, IC-532205, IC-118722, IC-532181, IC-532227, IC-532156, IC-532182, IC-35071, IC-532100, and IC-532101
	$J_{1.2.2.1}$	IC-59179, IC-534859, IC-55664, IC-55685, IC-55617, IC-82370, IC-79102, IC-82440, IC-82400, IC-82410, IC-82425, IC-104566, and IC-534774
	$J_{1.2.2.2}$	IC-82385, IC-82259, IC-82344, IC-82221, Iand C-82371
$K_{1.1}$	$K_{1.1.1}$	IC-532213
	$K_{1.1.2}$	IC-532934, IC-534223, IC-534192, IC-534524, IC-79028, IC-79050, IC-79107, IC-79108, IC-534534, IC-534555, IC-534582, IC-534690, IC-532142, IC-532146, IC-532140, IC-532905, IC-532729, IC-532821, IC-532857, IC-532847, IC-532849, IC-532223, and IC-532886
$K_{1.2.}$	$K_{1.2.1.1}$	IC-534802, IC-82130, IC-78889, IC-78895, IC-82119, IC-534189, and IC-534557
	$K_{1.2.1.2}$	IC-82199, IC-82204, IC-82192, IC-82198, IC-55636, IC-55652, IC-534870, IC-534873, and IC-534885
	$K_{1.2.2.1}$	IC-532726, IC-533954, IC-532504, IC-532705, and IC-532719
	$K_{1.2.2.2}$	IC-532118, IC-532130, IC-532150, IC-532259, IC-532253, IC-532265, IC-532139, IC-532145, IC-532277, and IC-533963
L_1	$L_{1.1}$	IC-78888, IC-78843, and IC-78854
	$L_{1.2}$	IC-79083, IC-79066, IC-79079, IC-78935, IC-78945, IC-78936, IC-78914, IC-78919, IC-79080, IC-78912, IC-78915, IC-78901, IC-78905, IC-78948, IC-78908, IC-78911, IC-78899, IC-78960, IC-78918, and IC-78941,

Contd…

Table 4.11 Contd…

L_2	$L_{2.1}$	IC-55578, IC-534887, IC-55507, IC-78930, IC-57998, IC-59191, IC-78890, IC-78891, IC-532132, IC-532124, IC-532126, IC-532206, IC-532221, IC-532129, IC-532224, IC-532230, and IC-532232
	$L_{2.2.1.1}$	IC-534771, IC-534755, IC-534762, IC-534806, IC-534808, IC-78869, IC-78920, IC-79022, IC-79023, IC-534794, and IC-534811
	$L_{2.2.1.2}$	IC-532122, IC-532249, IC-532121, IC-532136, IC-532081, IC-532087, IC-532151, IC-532153, IC-532148, and IC-532149
	$L_{2.2.2.1}$	IC-82169, IC-534819, and IC-534854
	$L_{2.2.2.2}$	IC-532286, IC-532300, IC-532485, IC-82377, IC-82388, IC-532805, IC-532811, IC-532487, and IC-532738

*Note: Cluster code: Alpha numeric code corresponding to that of Figure 4.4, Color dark Red: Gujarat, Red: Rajasthan, Yellow: Himachal Pradesh, Light Green: Uttar Pradesh, Light blue: Maharashtra, Green: Jammu & Kashmir, Purple: Madhya Pradesh, Blue: Karnataka, Black: Unknown.

green: Uttar Pradesh, light blue: Maharashtra, green: Jammu & Kashmir, purple: Madhya Pradesh, blue: Karnataka, and black: Unknown.

Analysis of UPGMA tree revealed that the studied set of 618 wheat accessions were broadly clustered into 4 main major groups: A_1, A_2, A_3 and A_4, each diverging from an unknown common ancestor. Major group A_2 had a single accession, IC-118758, of unknown origin. On the other hand, major group A_4 had the largest number of 579 accessions that were highly diverse among themselves. At genetic dissimilarity coefficient of 0.7, major group A_4 was further clustered into 11 groups, each of which had different number of sub-clusters. Major group A1 had 20 wheat accessions comprising of: 6 accessions, IC-78937, IC-75327, IC-75328, IC-75339, IC-78096 and IC-78860, from Uttar Pradesh; 3 accessions, IC-212145, IC-28974 and IC-73493, belonging to Rajastan; 2 accessions, IC-32520 and IC-31496, of Madhya Pradesh origin; 3 accessions, IC-73205, IC-73198 and IC-73201, from Karnataka; and, 6 accessions, IC-532887, IC-35070, IC-73232, IC-64238, IC-78933 and IC-78877 of unknown origin. Majour group A3 had 18 wheat accessions of which majority (10) of genotypes, *viz*., IC-78987, IC-78982, IC-78981, IC-78990, IC-79091, IC-79090, IC-79077, IC-78996, IC-78991 and IC-82303, belonging to Himachal Pradesh grouped together with two accessions, IC-78968 and IC-79085, of Uttar Pradesh origin, along with 6 accessions, *viz*., IC-79056, IC-79065, IC-79068, IC-138898, IC-82338, and IC-82286, of unknown origin.

In total, there were 14 groups, at 0.7 coefficient of genetic dissimilarity, as described in Table 4.11, *viz*., A_1, A_2, A_3, C_1, D_1, E_1, F_1, G_1, H_1, I_1, J_1, K_1, L_1 and L_2 with membership strengths of 20, 1, 13, 24, 1, 1, 94, 169, 68, 13, 80, 55, 23 and 50, respectively. Among them, three groups, *viz*., A_2, D_1 and E_1, had only one member each, namely, IC-118758, IC-30276 and IC-212185, respectively. Out of remaining 11 subgroups, the largest group was the G_1 comprising of 169 accessions that included Indian wheat landraces, the geographical origins of most of which are unknown; among those with known origin, three accessions (IC-79095, IC-78999 and IC-82342) belonging to Himachal Pradesh, four accessions (IC82180, IC-104537, IC-82127 and IC-82161) from Rajasthan, two accessions (IC-118729 and IC-73215) of Maharashtra were grouped into single cluster along with one accession (IC-35069) belonging to Karnataka, and two accessions (IC-78832 and IC-78707) from Uttar Pradesh. The second largest and the diverse subgroup was F_1 comprising of 94 accessions, followed

by subgroups J_1 and H_1 with membership sizes of 80 and 68 accessions, respectively. The smallest multi-membered group was I_1 comprising of 13 accessions, followed by groups A_1 and C_1 having 20 and 24 accessions, respectively. most (10) of which belonged to the political state of Himachal Pradesh. Out of 108 accessions of known origin, most of the landraces (20) belonging to Uttar Pradesh were clustered in group L_1, which had 24 accessions altogether. Group C_1 included most, 18 of 27, of the Indian wheat landraces from Rajasthan.

Except the groups A_2, D_1, E_1 and L_1, none of the other groups showed the clustering of accessions according to their origin. While groups A_2, D_1 and E_1 had a single member each, IC-118758 of unknown origin, IC-30276 belonging to state of Jammu and Kashmir, and IC-212185 from state of Rajasthan, respectively; group L_1 had a total of 23 accessions most (20) of which belonged to state of Uttar Pradesh, two, IC-79066 and IC-79079, were of Himachal Pradesh, and one, IC-78941, was of unknown origin.

Overall, there was no group, except three single-membered groups (A_2, D_1 and E_1), that had all the members of single origin, suggesting that origin of Indian wheat landraces mostly did not determine the genetic relatedness. However, IC-118758 of unknown origin, IC-30276 belonging to state of Jammu and Kashmir, and IC-212185 from state of Rajasthan were clustered into separate distinct groups, A_2, D_1 and E_1, respectively.

4.1.2.2 Analysis of molecular variance

The Software tool determined the variance components and estimated the total variation within and among the 9 populations, grouped according to the origin of the accessions. The results are depicted in Table 4.12 and illustrated in Figure 4.5. The total molecular genetic variance at subgenome (A, B and D) and whole genome levels were partitioned into molecular variance within the populations and among the populations. According to genotyping data for 42 SSR marker loci amplified across 618 wheat genotypes, the molecular genetic variance within the populations was 90%, whereas it was 10% among the populations (Table 4.12).

Further, it was observed that the similar partitioning into 90% within- and 10% among the populations of total variability assessed among 618 landrace accessions was existed in the molecular variances of the whole genome, ABD (Ψ PT ($p<0.001$) 0.099)

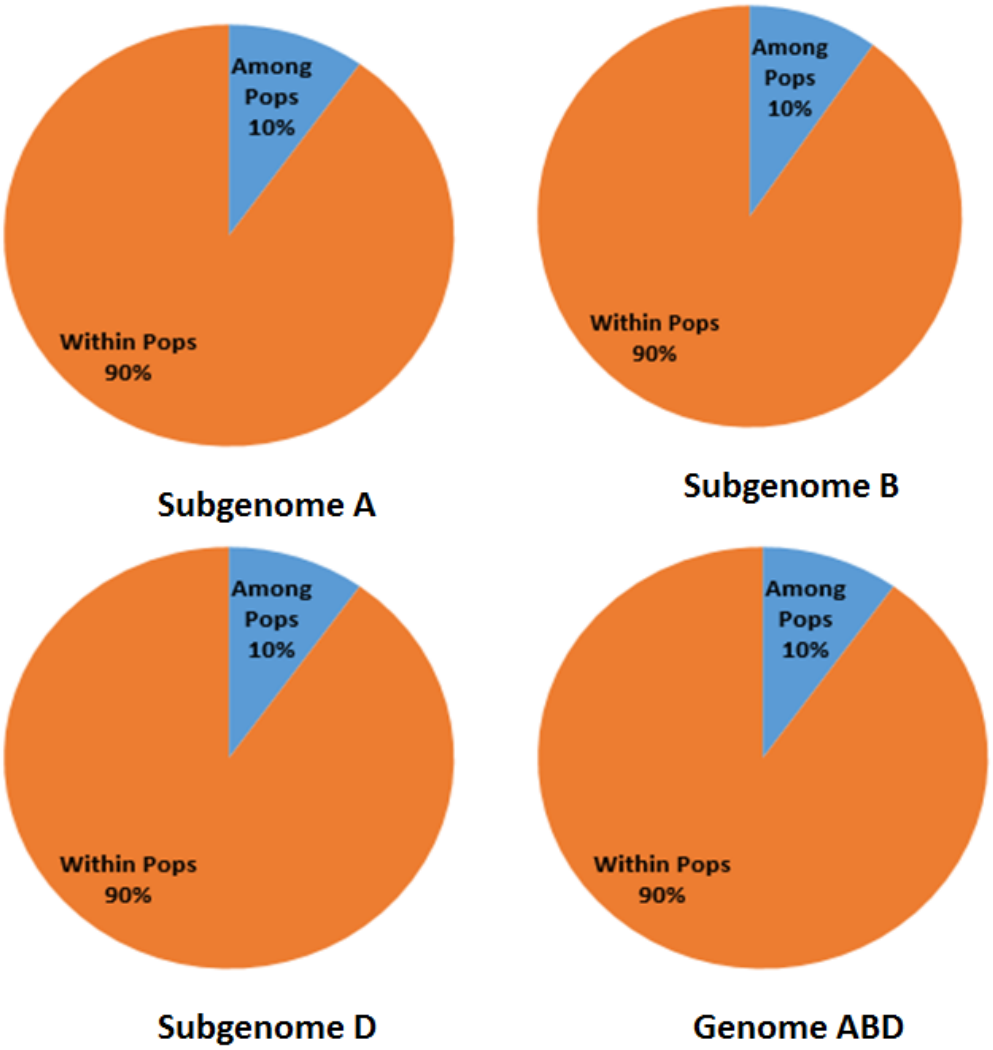


Figure 4.5 Percent molecular variance partitioned among and within the 9 populations as revealed by analysis of molecular variance using the genotyping data of 42 SSR markers loci amplified in a set of 618 Indian wheat accessions.

Table 4.12 Partitioning of the total molecular variance at subgenome and genome levels as revealed by analysis of molecular variance (AMOVA) based on genotyping data of 42 SSR markers loci amplified among 618 Indian wheat (*Triticum* spp.) landraces.

Parameter	Subgenome A			Subgenome B			Subgenome D			Genome ABD		
	Among Pops	**Within Pops**	**Total**	**Among Pops**	**Within Pops**	**Total**	**Among Pops**	**Within Pops**	**Total**	**Among Pops**	**Within Pops**	**Total**
Ψ PT ($p<0.001$)	-	-	0.099	-	-	0.099	-	-	0.097	-	-	0.099
Variance (Est.)	2.708	24.728	27.436	2.699	24.265	26.964	2.620	24.273	26.893	8.027	73.267	81.293
Variance (%)	10	90	100	10	90	100	10	90	100	10	90	100

Note: (-): Not applicable, Est.: Estimated, Pops: Populations.

and those of subgenomes: subgenome A (Ψ PT ($p<0.001$) 0.099), subgenome B (Ψ PT ($p<0.001$) 0.099) and subgenome D (Ψ PT ($p<0.001$) 0.097).

4.1.3 Study of population structure and phylogeny among the Indian wheat landraces

4.1.3.1 Model based population structure

Model-based population structure of the 618 Indian wheat accessions was estimated using STRUCTURE (version 2.3.3) software tool using the genotyping data of 42 SSR markers amplified across 618 Indian wheat landrace accessions. Evanno plot (Figure 4.6) obtained through STRUCTURE analysis revealed the K-value four, suggesting the presence of four subpopulations in a panel of 618 Indian wheat accessions. The number of groups deciphered by using STRUCTURE, analysis is given in Figure 4.7.

According to a threshold value of membership-probability 0.80, 176 accessions were assigned to subpopulation A, 144 to B, 135 to C and 163 to D (Table 4.13, Figure 4.7 and appendix A.11). Accordingly, subpopulation-A was the largest, with 155 accessions having pure ancestry and 21 accessions belonging to admixture, out of total 176 accessions in the group. On the other hand, the smallest was the subpopulation-C having a total of 135 accessions segregated into 130 pure and 5 admixtures. The highest number of 23 accessions with admixture ancestry was observed in subpopulation-B that had 121 pure accessions out of a total of 144 accessions. The lowest number of 2 admixture accessions was recorded, in subpopulation-D along with 161 accessions with pure ancestry out of total 163 accessions.

4.1.3.2 Principal coordinate analysis (PCoA)

In order to further confirm the results of STRUCTURE analysis, the genotyping data of a marker panel of 42 SSR markers was subjected to three-dimensional of Principal coordinate analysis (PCoA). Multidimensional scaling (MDS) plots were constructed and the genetic interrelationships among the accessions were visualized in PCoA plot given in Figure 4.8. A Genotype panel of 618 Indian wheat landrace accessions was separated into 4 major groups on three-dimensional coordinates. The first, second and third axes explained 4.14 %, 3.34 % and 2.77%, respectively, of the total genetic variation.

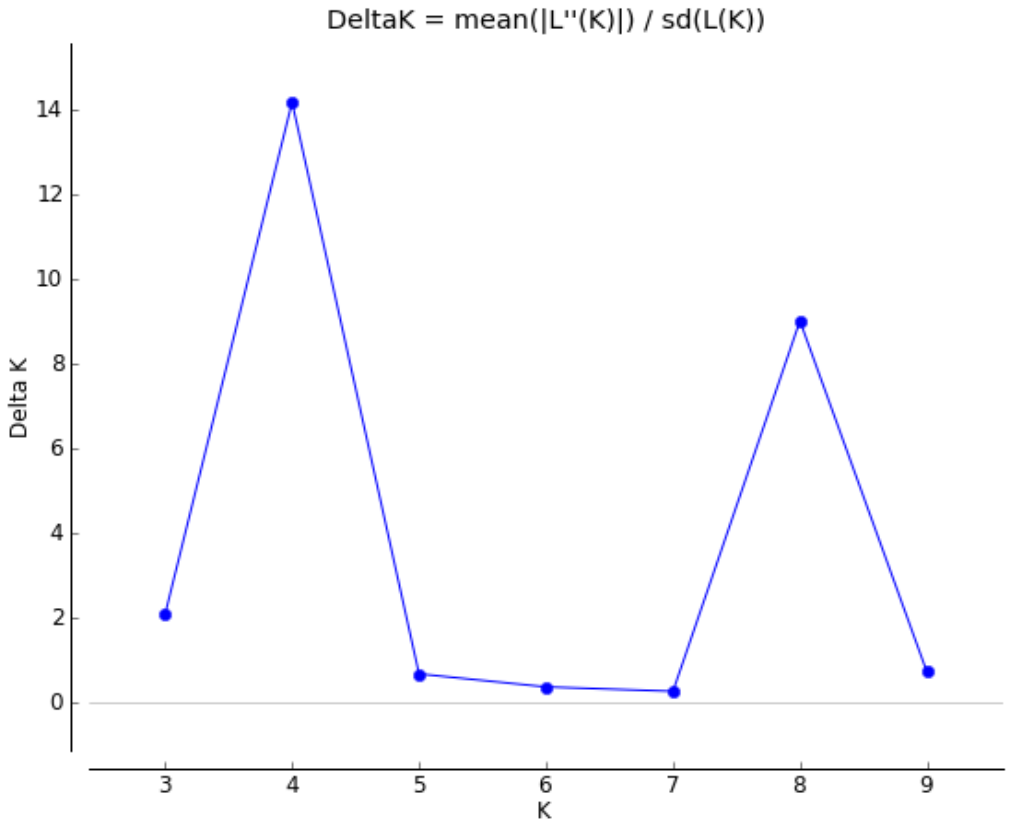


Figure 4.6 Evanno plot of delta *K* vs *K* showing the *K*=4. As indicated by the *K* value on x-axis corresponding to first peak of delta *K* value on y-axis, there are four sub-populations in the studied set of 618 Indian wheat accessions.

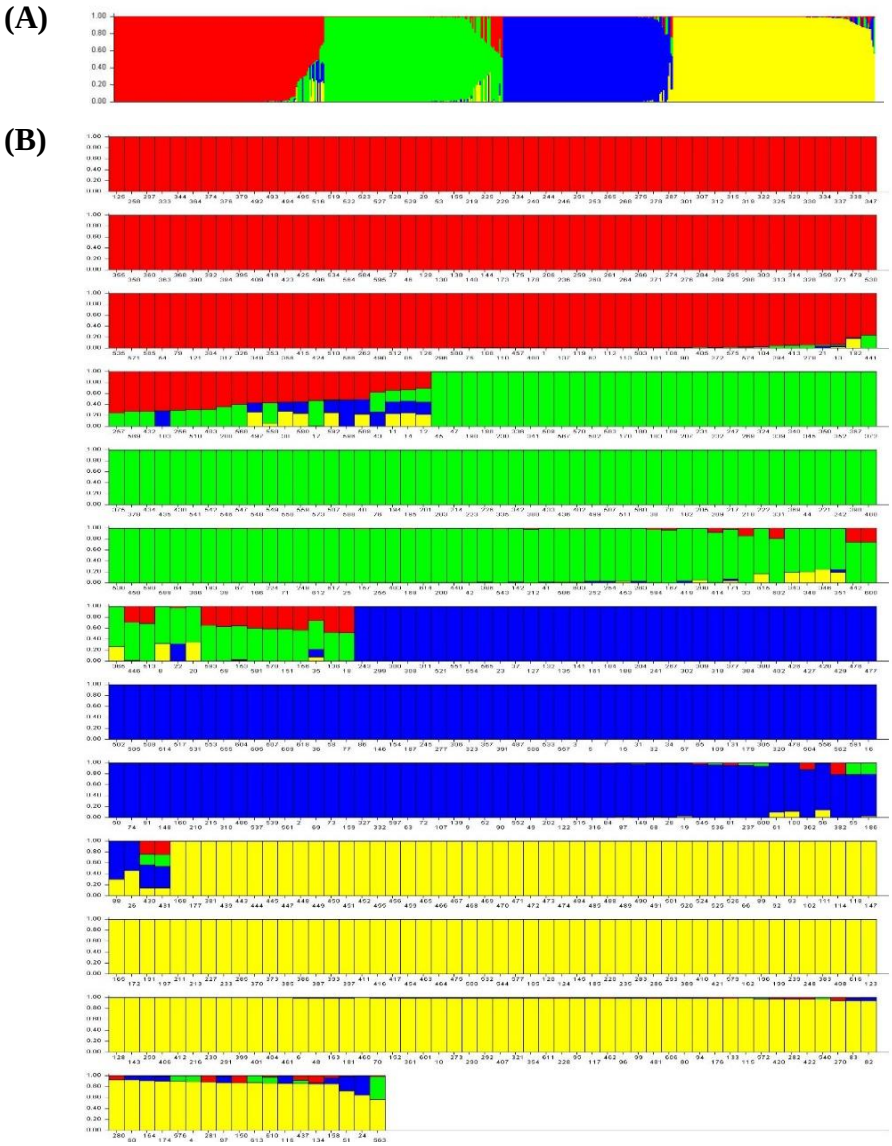


Figure 4.7 Barplots showing model-based subpopulations inferred from STRUCTURE analysis. (A): Overall population-wise barplots, and (B): Genotype-wise barplots. There are four populations (K=4), color coded are as follows: red: population A, green: population B, blue: population C, and yellow: population D. Population codes correspond to those described in Table 4.13. Vertical coordinates correspond to membership coefficients and each vertical bar separated by vertical line along with horizontal coordinates denotes individual genotypes. Distinct color within a vertical bar denotes an individual genome belonging to a distinct population denoted by the respective color code. Distinct colors within a vertical bar indicate that the corresponding genotype is an admixture, whereas single color indicates purity. In case of admixtures, proportions of distinct colors within the bar correspond to proportions of distinct populations contributing to genome of the individual represented by the bar.

Table 4.13 Ancestry details, number and list of individual members in each of the model-based populations obtained from STRUCTURE analysis.

Model based population	No. of accessions			Individual accessions constituting	
	Pure	Admixture	Total	Pure	Admixture
A	155	21	176	IC_104537, IC_118729, IC_118731, IC_118741, IC_138896, IC_138897, IC_35069, IC_532059, IC_532061, IC_532073, IC_532092, IC_532096, IC_532103, IC_532110, IC_532120, IC_532131, IC_532134, IC_532138, IC_532140, IC_532142, IC_532144, IC_532146, IC_532147, IC_532155, IC_532175, IC_532183, IC_532184, IC_532187, IC_532188, IC_532213, IC_532214, IC_532217, IC_532223, IC_532239, IC_532261, IC_532263, IC_532267, IC_532478, IC_532503, IC_532701, IC_532708, IC_532727, IC_532729, IC_532768, IC_532784, IC_532790, IC_532794, IC_532813, IC_532821, IC_532835, IC_532837, IC_532839, IC_532841, IC_532842, IC_532847, IC_532849, IC_532853, IC_532855, IC_532857, IC_532868, IC_532886, IC_532887, IC_532905, IC_532910, IC_532923, IC_532934, IC_532935, IC_533962, IC_533965, IC_533970, IC_534190, IC_534192, IC_534193, IC_534197, IC_534223, IC_534323, IC_534363, IC_534419, IC_534435, IC_534509, IC_534524, IC_534534, IC_534543, IC_534549, IC_534553, IC_534555, IC_534560, IC_534564, IC_534565, IC_534568, IC_534572, IC_534582, IC_534688, IC_534690, IC_534721, IC_534723, IC_534743, IC_534747, IC_534749, IC_534754, IC_534756, IC_534757, IC_534759, IC_534760, IC_534761, IC_534764, IC_534765, IC_534769, IC_534772, IC_534775, IC_534777, IC_534787, IC_534820, IC_534823, IC_534855, IC_534857, IC_534886, IC_55593, IC_55656, IC_55659, IC_57844, IC_59511, IC_59544, IC_59612, IC_78872, IC_78927, IC_78928, IC_78968, IC_78981, IC_78982, IC_78987, IC_78990, IC_78996, IC_79028, IC_79046, IC_79050, IC_79056, IC_79065, IC_79068, IC_79077, IC_79085, IC_79090, IC_79091, IC_79095, IC_79107, IC_79108, IC_82177, IC_82180, IC_82190, IC_82195, IC_82197, IC_82206, IC_82256, IC_82257, IC_82367	IC_118723, IC_118723, IC_118727, IC_118730, IC_118736, IC_138898, IC_30276, IC_532137, IC_532289, IC_533966, IC_534884, IC_73215, IC_78707, IC_78991, IC_79063, IC_82161, IC_82185, IC_82187, IC_82286, IC_82303, IC_82338
B	121	23	144	IC_138849, IC_212145, IC_28958, IC_28969, IC_28974, IC_29035, IC_30284, IC_31496, IC_32520, IC_35070, IC_532077, IC_532083, IC_532088, IC_532093, IC_532095, IC_532219, IC_532237, IC_532241, IC_532250, IC_532251, IC_532256, IC_532257, IC_532258, IC_532271, IC_532273, IC_532274, IC_532285, IC_532290, IC_532292, IC_532297, IC_532309, IC_532318, IC_532406, IC_532474, IC_532476, IC_532480, IC_532481, IC_532482, IC_532486, IC_532489, IC_532497, IC_532502, IC_532697, IC_532698, IC_532699, IC_532700, IC_532704, IC_532723, IC_532725, IC_532737, IC_532775, IC_532807, IC_532815, IC_532832, IC_532833, IC_532851, IC_532872, IC_534562, IC_534605, IC_534697, IC_534720, IC_534736, IC_534737, IC_534739, IC_534741, IC_534744, IC_534750, IC_534752, IC_534767, IC_534768, IC_534770, IC_534773, IC_534776, IC_534786, IC_534792, IC_534858, IC_534864, IC_534871, IC_55657, IC_57983, IC_73232, IC_73493, IC_75327, IC_75328, IC_75339, IC_78096, IC_78860, IC_78877, IC_78931, IC_78933, IC_78999, IC_79041, IC_79043, IC_79047, IC_82116, IC_82125, IC_82126, IC_82127, IC_82133, IC_82134, IC_82135, IC_82136, IC_82145, IC_82163, IC_82164, IC_82181, IC_82189, IC_82193, IC_82236, IC_82247, IC_82259, IC_82263, IC_82285, IC_82344, IC_82371, IC_82372, IC_82381, IC_82385, IC_82421, IC_82426, IC_82221	IC_104573, IC_118737, IC_118740, IC_118758, IC_212185, IC_532068, IC_532210, IC_532231, IC_532240, IC_534742, IC_534745, IC_534748, IC_534751, IC_534766, IC_78728, IC_78832, IC_79052, IC_82169, IC_82200, IC_82210, IC_82342, IC_82375, IC_82432

Contd…

Table 4.13 Contd…

C	130	5	135	IC_104540, IC_104543, IC_104564, IC_104568, IC_104576, IC_118732, IC_118739, IC_118762, IC_138893, IC_210548, IC_212142, IC_212179, IC_28664, IC_28926, IC_47478, IC_532055, IC_532058, IC_532062, IC_532064, IC_532066, IC_532067, IC_532072, IC_532074, IC_532078, IC_532084, IC_532085, IC_532089, IC_532090, IC_532091, IC_532094, IC_532098, IC_532102, IC_532106, IC_532119, IC_532133, IC_532141, IC_532143, IC_532180, IC_532185, IC_532201, IC_532204, IC_532209, IC_532216, IC_532218, IC_532225, IC_532228, IC_532229, IC_532238, IC_532243, IC_532244, IC_532245, IC_532268, IC_532276, IC_532281, IC_532282, IC_532284, IC_532473, IC_532475, IC_532483, IC_532490, IC_532736, IC_532773, IC_532779, IC_532787, IC_532863, IC_532930, IC_534235, IC_534322, IC_534334, IC_534430, IC_534432, IC_534451, IC_534455, IC_534480, IC_534481, IC_534554, IC_534556, IC_534561, IC_534566, IC_534609, IC_534758, IC_534763, IC_534778, IC_534798, IC_534805, IC_534814, IC_534822, IC_534883, IC_64238, IC_73198, IC_73201, IC_73205, IC_78923, IC_78924, IC_78925, IC_78937, IC_78940, IC_79026, IC_79038, IC_79039, IC_79040, IC_79042, IC_79053, IC_79055, IC_79062, IC_79067, IC_79100, IC_79106, IC_79110, IC_79940, IC_82118, IC_82132, IC_82146, IC_82147, IC_82156, IC_82157, IC_82158, IC_82159, IC_82160, IC_82165, IC_82167, IC_82179, IC_82385, IC_82369, IC_82386, IC_82387, IC_82393, IC_82394, IC_82398, IC_8255	IC_138895, IC_532116, IC_532279, IC_73207, IC_73210
D	161	2	163	IC_104551, IC_104566, IC_118722, IC_35071, IC_532071, IC_532081, IC_532087, IC_532097, IC_532100, IC_532101, IC_532118, IC_532121, IC_532122, IC_532124, IC_532126, IC_532129, IC_532130, IC_532132, IC_532136, IC_532139, IC_532145, IC_532148, IC_532149, IC_532150, IC_532151, IC_532153, IC_532156, IC_532181, IC_532182, IC_532186, IC_532205, IC_532206, IC_532221, IC_532224, IC_532227, IC_532230, IC_532232, IC_532242, IC_532246, IC_532247, IC_532248, IC_532249, IC_532253, IC_532259, IC_532262, IC_532264, IC_532265, IC_532272, IC_532277, IC_532286, IC_532298, IC_532300, IC_532310, IC_532485, IC_532487, IC_532492, IC_532504, IC_532705, IC_532707, IC_532719, IC_532726, IC_532728, IC_532738, IC_532805, IC_532811, IC_532880, IC_532891, IC_532936, IC_532937, IC_533953, IC_533954, IC_533963, IC_533964, IC_533971, IC_534072, IC_534123, IC_534189, IC_534557, IC_534755, IC_534762, IC_534771, IC_534774, IC_534794, IC_534802, IC_534806, IC_534808, IC_534811, IC_534819, IC_534854, IC_534859, IC_534870, IC_534873, IC_534885, IC_534887, IC_55507, IC_55578, IC_55617, IC_55636, IC_55652, IC_55664, IC_55685, IC_57998, IC_59179, IC_59191, IC_75333, IC_75351, IC_78715, IC_78824, IC_78828, IC_78836, IC_78837, IC_78839, IC_78843, IC_78854, IC_78859, IC_78862, IC_78865, IC_78869, IC_78888, IC_78889, IC_78890, IC_78891, IC_78895, IC_78897, IC_78899, IC_78901, IC_78905, IC_78908, IC_78911, IC_78912, IC_78914, IC_78915, IC_78918, IC_78919, IC_78920, IC_78930, IC_78935, IC_78936, IC_78941, IC_78945, IC_78948, IC_78960, IC_79022, IC_79023, IC_79066, IC_79079, IC_79080, IC_79083, IC_79102, IC_82119, IC_82130, IC_82192, IC_82198, IC_82199, IC_82204, IC_82377, IC_82388, IC_82400, IC_82410, IC_82425, IC_82440	IC_138845, IC_532057
Total	**567**	**101**	**618**	**-**	

Note: Dark red: Gujarat, Red: Rajasthan, Yellow: Himachal Pradesh, Light green: Uttar Pradesh, Light blue: Maharashtra, Green: Jammu & Kashmir, Purple: Madhya Pradesh, Blue: Karnataka, Black: Unknown.

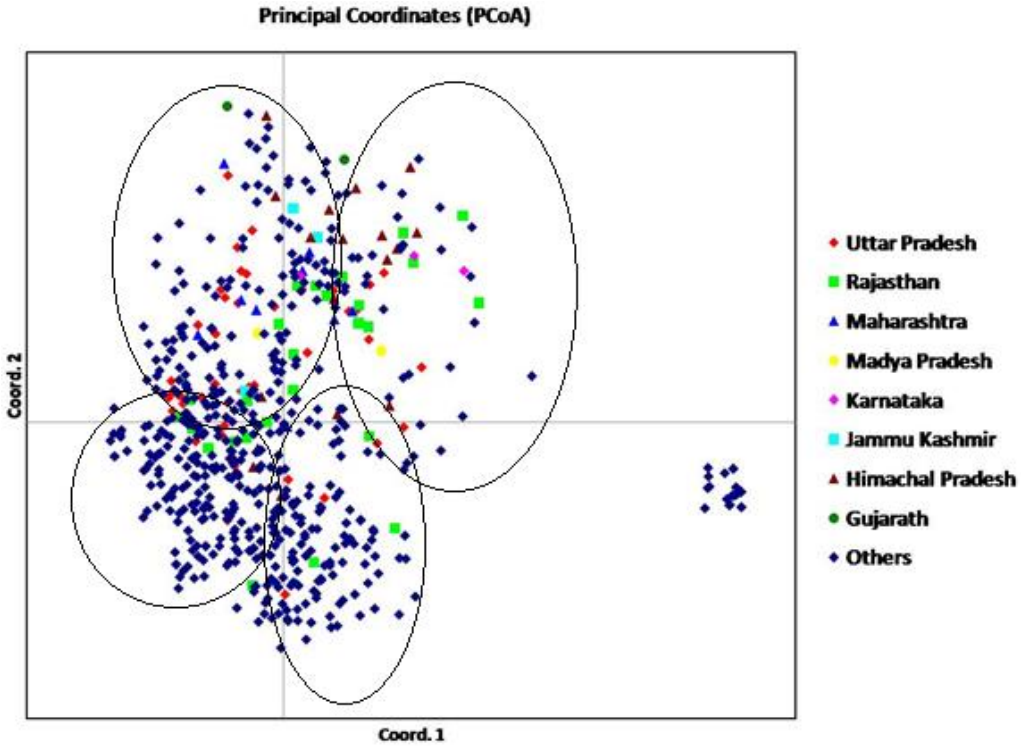


Figure 4.8 Principal coordinate analysis of 618 wheat landraces based on geographic origin of accessions and variations for 42 SSR markers loci. There are 4 subgroups inferred by STRUCTURE based on maximum membership probability. Different colored spots correspond to geographic origins (i.e., Indian geopolitical states).

4.1.4 Phylogenetic analysis

The resultant UPGMA tree-based dendrogram is depicted in Figure 4.3 and Figure 4.4, and grouping of the wheat accessions into different clusters is detailed in Table 4.11. Analysis of UPGMA-based phylogenetic tree revealed that the studied set of 618 wheat accessions were broadly grouped into 4 major phylogenetic clads: Clad-A1, Clad -A2, Clad -A3 and Clad-A4, each diverging from an unknown common ancestor. Clad A2 had a single genotype having accession number IC-118758 of unkown origin. On the other hand, Clad A4 had the largest number of 579 accessions that were highly diverse among them. Clad A1 had 20 wheat accessions comprising of: 6 accessions, IC-78937, IC-75327, IC-75328, IC-75339, IC-78096 and IC-78860, from Uttaer Pradesh; 3 accessions, IC-212145, IC-28974 and IC-73493, belonging to Rajastan; 2 accssessions, IC-32520 and IC-31496, of Madhya Pradesh origin; 3 accessions, IC-73205, IC-73198 and IC-73201, from Karnataka; and, 6 accessions, IC-532887, IC-35070, IC-73232, IC-64238, IC-78933 and IC-78877 of unkown origin. Clad A3 had 18 wheat accessions of which majority (10) of genotypes, *viz*., IC-78987, IC-78982, IC-78981, IC-78990, IC-79091, IC-79090, IC-79077, IC-78996, IC-78991 and IC-82303, belonging to Himachal Pradesh grouped together with two accessions, IC-78968 and IC-79085, of Uttar Pradesh origin, along with 6 accessions, *viz*., IC-79056, , IC-79065, IC-79068, IC-138898, IC-82338, and IC-82286, of unkown origin.

4.2 Study of population structure and phylogeny based on agro-morphological and phenological traits

4.2.1 Study of agro-morphological diversity

4.2.1.1 Extent and pattern of genetic diversity

Fifteen quantitative morphological traits investigated in this study showed a wide range of agro-morphological variations in the evaluated wheat accessions as revealed by the percent coefficients of variation (CV) given in Table 4.14. Majority of the accessions were of spreading type with medium plan height (90 to 105 cm), intermediate tillering capacity (5 to 10 cm), medium spike length (5 to 10 cm), having 15 to 20 spikelets per spike, single kernel per floret, 25 to 50 seeds per spike, awned, glume color purple to black, pubescent glume, intermediate type of spike density, red seed color, intermediate seed size, intermediate seed shriveling, early emergence (12 to 15 days), days to tillering medium (20 to 25 days), days to booting late (75 to 90 days),

Table 4.14 Statistical description of 15 quantitative traits evaluated in 618 Indian wheat accessions.

Traits	Mean	SEM	Min	Max	Var	SD	CV	Skew	Kurtosis
SPL	9.21	0.07	2.50	15.00	2.93	1.71	18.58	-0.04	0.21
NSLS	14.83	0.09	11.00	23.00	5.16	2.27	15.31	0.42	-0.05
NFSL	3.33	0.06	1.00	7.00	2.03	1.42	42.78	0.43	0.28
NSSL	2.90	0.06	1.00	5.00	0.98	0.99	34.06	11.51	221.58
MNSF	1.50	0.03	1.00	3.00	0.42	0.65	43.01	0.92	-0.26
NSS	41.37	0.87	11.00	101.00	263.91	16.25	39.27	7.52	119.49
TKW	25.01	0.12	6.68	69.99	50.35	7.10	28.37	-1.21	-0.46
PLH	95.76	0.03	61.95	120.55	74.64	8.64	9.02	-0.53	-1.44
TILC	4.30	0.03	1.00	20.00	5.74	2.40	55.69	-0.41	-0.95
DEM	15.48	0.03	3.70	21.80	6.89	2.62	16.96	-0.16	-0.17
DTIL	27.08	0.29	15.70	38.70	33.98	5.83	21.53	0.56	2.77
DBOO	64.73	0.03	47.70	98.10	458.46	21.41	33.08	-0.07	-1.28
DSEM	78.61	0.04	59.70	113.10	459.75	21.44	27.28	0.58	-0.39
DFLO	86.65	0.05	69.70	120.10	457.10	21.38	24.67	0.01	-0.08
DMAT	111.45	0.35	90.80	150.60	465.44	21.57	19.36	-0.35	0.53

SPL, spike length in cm; NSLS, number of spikelets per spike; NFSL, number of florets per spikelet; NSSL, number of seeds per spikelet; MNSF, maximum number of seeds per floret; NSS, number of seeds per spike; AWN, awnness; GLCL, glume color; GLHR, glume hairiness; SPD, spike density; TKW, thousand kernels weight in gram; SDCL, seed color; SDZ, seed size (relative); DSHL, degree of seed shriveling; PLH, plant height in cm; GRHB, growth habit; TILC, tillering capacity; DEM, days to emergence; DTIL, days to tillering; DBOO, days to booting; DSEM, days to spike emergence; DFLO, days to flowering; DMAT, days to physiological maturity.

days to spike emergence early (60 to 75 days), very early flowering (≤ 70 days), early physiological maturity (90 to 105 days), and low thousand kernel weight (20 to 30 g). Two accessions, IC_534777 and IC_534787, recorded the highest 1000-kernel weights of more than 50 gram.

Estimated phenotypic diversities (H') for individual traits is presented in Table 4.15. Phenotypic diversity (H') value ranged from 0.93, for glume hairiness, to 0.49, for the maximum number of seeds per spike. Analyzed accessions showed high phenotypic diversity (H'>0.60), except for NSS (0.59) and NSSL (0.49).

Values of most of the agro-morphological characters analysed in the present study were negatively skewed and biased towards higher values of phenotypes.

4.2.1.2 Correlation among the 23 analyzed traits

The correlation coefficients among 23 morphological traits are given in Table 4.16 and the overview is graphically represented in form of a heat-map in Figure 4.9. There existed highly statistically significant ($p<0.01$) positive relationships of days to tillering with all other phenological traits: days to booting ($r=0.92$, $p<.01$), days to flowering ($r=0.92$, <0.01), days to spike emergence ($r=0.91$, $p<.01$) and days to maturity ($r=0.91$, $p<0.01$). Days to booting had the highest degree of positive association with days to spike emergence ($r=0.99$, $p<0.01$), days to flowering ($r=0.99$, $p<0.01$), and days to maturity ($r=0.98$, $p<0.01$), The relationship of days to spike emergence was highly positive with days to flowering ($r=0.99$, $p<0.01$), and days to physiological maturity ($r=0.97$, $p<0.01$). The high degree of positive relationship was also observed between days to flowering and days to physiological maturity ($r=0.99$, $p<0.01$). Statistically highly significant positive correlation existed between days to emergence and days to tillering ($r=0.87$, $p<0.01$). Days to emergence showed moderate degree of positive relationship with days to booting ($r=0.62$, $p< 0.01$), days to spike emergence ($r=0.61$, $p<0.01$), days to flowering ($r=0.62$, $p<0.01$) and days to physiological maturity ($r=0.62$, $p< 0.01$).

Degree of seed shriveling showed statistically highly significant ($p<0.01$) low degree of positive relationship with days to spike emergency ($r=0.28$), days to flowering ($r=0.28$), days to physiological maturity ($r=0.28$), days to tillering ($r=0.27$), days to booting ($r=0.27$), and days to emergency ($r=0.2$). However, it showed highly significant ($p<.01$) low degree of negative relationship with plant height ($r=-0.12$).

Table 4.15 Agro-morphological traits class, frequency of accessions and estimated trait phenotypic diversity index (H') in the total 618 indigenous wheat accessions evaluated.

Traits	State (class) of the trait	NA	Freq (%)	H'	Trait	State (class) of the trait	NA	Freq (%)	H'	Trait	State (class) of the trait	NA	Freq (%)	PDI
GRHB	3 = erect	112	18.12	0.94	NSS	1 = very low (≤25)	90	14.56	0.59		7 = late (>18 to 21 days)	114	18.45	
	5 = semi spreading	213	34.47			3 = low (>25 to 50)	428	69.26			9 = very late (>21 days)	11	1.78	
	7 = spreading	293	47.41			5 = medium (>50 to 75)	69	11.17		DTIL	3 = early (15 to 20 days)	64	10.36	0.88
PLH	1 = very short (≤75 cm)	11	1.78	0.60		7 = high(>75 to 100)	29	4.69			5 = medium (>20 to 25 days)	264	42.72	
	3 = short (>75 to 90 cm)	138	22.33			9 = very high(>100)	2	0.32			7 = late (>25 to 30 days)	78	12.62	
	5 = medium (>90 to 105 cm)	391	63.27		AWN	1 = awnless	134	21.68	0.62		9 = very late (>30 days)	212	34.3	
	7 = long (>105 to 120 cm)	76	12.3			5 = awnletted	25	4.05		DBOO	1 = very early (≤50 days)	357	57.77	0.71
	9 = very long (>120cm)	2	0.32			9 = awned	459	74.27			3 = early (>50 to 60 days)	19	3.07	
TILC	3 = low (≤5)	277	44.82	0.71	GLCL	1 = white	168	27.18	0.92		5 = medium (>60 to 75 days)	22	3.56	
	5 = intermediate (>5 to 10)	328	53.07			5 = red to brown	120	19.42			7 = late (>75 to 90 days)	89	14.4	
	7 = high (>10)	13	2.1			9 = purple to black	330	53.4			9 = very late (>90 days)	131	21.2	
SPL	1 = very short (≤6.0 cm)	24	3.88	0.79	GLHR	1 = absent	96	15.53	0.93	DSEM	1 = very early (≤60 days)	28	4.53	0.75
	3 = short (>6 to 8 cm)	135	21.84			5 = low	267	43.2			3 = early (>60 to 75 days)	348	56.31	
	5 = medium (>8 to 10 cm)	282	45.63			9 = high	255	41.26			5 = medium (>75 to 90 days)	27	4.37	
	7 = long (>10 to 12 cm)	156	25.24		SPD	1 = very lax	12	1.94	0.77		7 = late (>90 to 105 days)	84	13.59	
	9 = very long (>12 cm)	21	3.4			3 = lax	105	16.99			9 = very late (>105 days	131	21.2	
NSLS	3 = few (≤15)	244	39.48	0.67		5 = intermediate	276	44.66		DFLO	1 = very early (≤70 days)	256	41.42	0.88
	5 = medium (>15 to 20)	366	59.22			7 = dense	200	32.36			3 = early (>70 to 85 days)	121	19.58	
	7 = high (>20)	8	1.29			9 = very dense	25	4.05			5 = medium (>85 to 105 days)	53	8.58	
NFSL	1 = low (1)	89	14.4	0.76	SDCL	1 = white	167	27.02	0.99		7 = late (>105 to 120 days)	150	24.27	
	3 = medium (3)	364	58.9			5 = red	256	41.42			9 = very late (>120 days)	38	6.15	
	5 = high (5)	140	22.65			9 = purple	195	31.55		DMAT	1 = early (90 to 105 days)	375	60.68	0.73
	7= very high (7)	25	4.05		SDZ	3 = small	12	1.94	0.62		3 = medium (>105 to 120 days)	24	3.88	
NSSL	1 = very low (1)	90	14.56	0.49		5 = intermediate	437	70.71			5 = late (>120 to 135 days)	69	11.17	
	3 = low (2)	5	0.81			7 = large	169	27.35			7 = very late (>135 days)	150	24.27	
	5 = medium (3)	460	74.43		DSHL	3 = plump	62	10.03	0.77	TKW	1 = very low (≤20 g)	148	23.95	0.68
	7 = high (4)	3	0.49			5 = intermediate	407	65.86			3 = low (>20 to 30 g)	326	52.75	
	9 = very high (5)	60	9.71			7 = shriveled	149	24.11			5 = medium (>30 to 40 g)	132	21.36	
MNSF	3 = low (1)	358	57.93	0.81	DEM	1 = very early (≤12 days)	59	9.55	0.82		7 = high (>40 to 50 g)	10	1.62	
	5 = medium (2)	208	33.66			3 = early (>12 to 15 days)	265	42.88			9 = very high (>50 g)	2	0.32	
	7 = high (3)	52	8.41			5 = medium (>15 to 18 days)	169	27.35						

Rows: SPL, spike length in cm; NSLS, number of spikelets per spike; NFSL, number of florets per spikelet; NSSL, number of seeds per spikelet; MNSF, maximum number of seeds per floret; NSS, number of seeds per spike; AWN, awnness; GLCL, glume color; GLHR, glume hairiness; SPD, spike density; TKW, thousand kernels weight in gram; SDCL, seed color; SDZ, seed size (relative); DSHL, degree of seed shriveling; PLH, plant height in cm; GRHB, growth habit; TILC, tillering capacity; DEM, days to emergence; DTIL, days to tillering; DBOO, days to booting; DSEM, days to spike emergence; DFLO, days to flowering; DMAT, days to physiological maturity. Columns: NA, number of accessions; Freq, frequency as percent of total number of accessions; PDI, phenotypic diversity index (H') estimated from Shannon-Weaver diversity index (H`).

Table 4.16 Pearson correlation coefficients for twenty agro-morphological and phenological traits of 618 Indian wheat accessions; Probability values, Pearson coefficients, N=618, Prob<[r] under HO: Rho=0.

	SPL	NSLS	NFSL	NSSL	MNSF	NSS	AWN	GLCL	GLHR	SPD	TKW	SDCL	SDZ	DSHL	PLH	GRHB	TILC	DEM	DTIL	DBOO	DSEM	DFLO
NSLS	**0.51****																					
NFSL	**0.21****	0.01																				
NSSL	**0.18****	0.01	**0.49****																			
MNSF	0.029	0.028	**0.109****	-0.048																		
NSS	**0.3****	**0.33****	**0.51****	**0.93****	-0.04																	
AWN	**-0.18****	**-0.18****	0.05	0.06	-0.02	0.01																
GLCL	**0.18****	**0.14****	**0.10***	0.03	0.08	0.07	-0.02															
GLHR	**-0.12***	-0.05	0.08	0.03	**0.10***	0.01	**0.16****	0.04														
SPD	**-0.23****	0.08	**-0.10***	-0.06	0.02	-0.03	0.07	-0.03	0.07													
TKW	-0.05	-0.03	0.00	0.02	0.00	0.01	0.01	-0.14	-0.02	0.03												
SDCL	-0.04	-0.01	**-0.11****	**-0.08***	-0.04	**-0.08***	-0.04	0.01	-0.02	-0.06	0.04											
SDZ	0.01	0.07	0.03	0.02	-0.01	0.04	-0.02	-0.05	0.05	-0.03	**0.35****	**0.10***										
DSHL	0.04	0.01	0.01	-0.01	-0.03	-0.01	-0.04	**0.14****	0.00	-0.01	**-0.84****	-0.05	**-0.32****									
PLH	0.06	0.01	-0.06	0.00	0.05	0.00	0.03	0.02	0.01	-0.04	**0.15****	-0.01	0.00	**-0.12****								
GRHB	-0.07	-0.04	0.02	-0.02	0.03	-0.04	0.06	-0.04	0.05	0.02	**-0.10***	-0.06	-0.03	0.07	-0.05							
TLCP	**-0.08***	-0.07	0.04	-0.07	-0.08	-0.08	-0.04	0.01	0.03	**0.13****	**-0.10***	**-0.12****	**-0.12****	0.06	0.03	0.05						
DEM	0.06	-0.01	-0.01	0.02	-0.04	0.02	0.04	0.05	-0.03	**-0.13****	**-0.19****	0.08	-0.07	**0.20****	-0.05	0.03	-0.03					
DTIL	0.04	-0.01	-0.01	0.00	-0.02	0.00	0.02	0.06	-0.02	**-0.12****	**-0.27****	0.04	**-0.08***	**0.27****	-0.08	0.05	0.03	**0.87****				
DBOO	0.02	-0.01	-0.01	-0.01	0.01	-0.02	-0.01	0.06	0.00	-0.07	**-0.28****	-0.01	**-0.08***	**0.27****	-0.07	0.05	0.05	**0.62****	**0.92****			
DSEM	0.02	-0.01	0.00	0.00	0.00	-0.02	-0.01	0.05	0.00	-0.08	**-0.29****	-0.01	**-0.08***	**0.28****	**-0.08***	0.05	0.06	**0.61****	**0.91****	**0.99****		
DFLO	0.02	-0.01	0.00	-0.01	0.01	-0.02	0.00	0.05	-0.01	-0.08	**-0.29****	-0.01	**-0.09***	**0.28****	**-0.08***	0.05	0.05	**0.62****	**0.92****	**0.99****	**0.99****	
DMAT	0.02	-0.01	0.00	0.00	0.02	-0.02	0.01	**0.04***	0.00	-0.07	**-0.29****	-0.02	**-0.08***	**0.28****	**-0.08***	0.05	0.05	**0.62****	**0.91****	**0.98****	**0.97****	**0.99****

*p<0.05; **p<0.01; SPL, spike(ear) length; NSLS, number of spikelets per spike; NFSL, number of florets per spikelet; NSSL, number of seeds per spikelet; MNSF, maximum Number of seeds per floret; NSS, number of seeds per spike; AWN, awnedness; GLCL, glume color; GLHR, glume hairiness; SPD, spike density (SPD); TKW, thousand kernels weight; SDCL, seed color; SDZ, seed size; DSHL, degree of seed shrivelling; PLH, plant height (PLH); GRHB, growth habit; TILC, tillering capacity; DEM, number of days to emergence; DTIL, number of days to tillering; DBOO, number of days to booting; DSEM, number of days to spike emergence; DFLO, number of days to flowering; DMAT, number of days to physiological maturity.

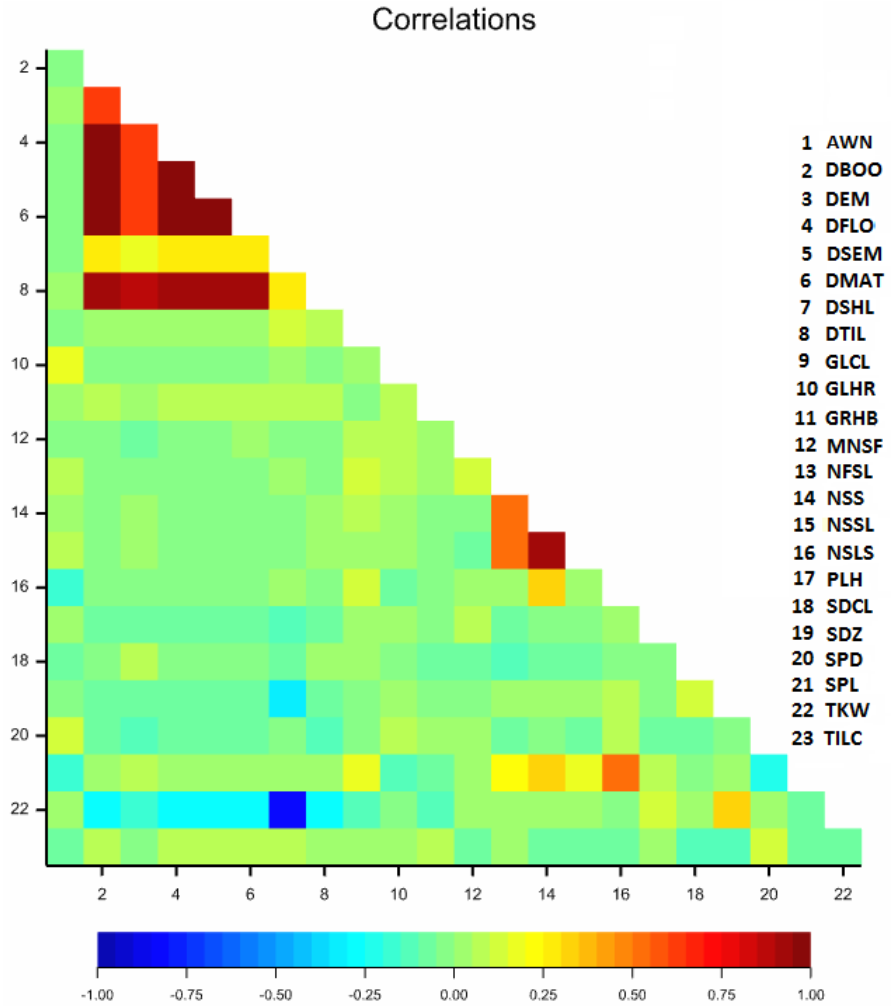


Figure 4.9 Heat map for Carl Pearson's coefficients of correlation among 23 agro-morphological traits in a set of 218 Indian wheat accessions. SPL: spike length in cm, NSLS: number of spikelets per spike, NFSL: number of florets per spikelet, NSSL: number of seeds per spikelet; MNSF: maximum number of seeds per floret, NSS: number of seeds per spike, AWN: awnness, GLCL: glume color, GLHR: glume hairiness, SPD: spike density, TKW: thousand kernel, SDCL: seed color, SDZ: seed size (relative), DSHL: degree of seed shriveling, PLH: plant height, GRHB: growth habit, TILC: tillering capacity, DEM: days to emergence, DTIL: days to tillering, DBOO: days to booting, DSEM: days to spike emergence; DFLO: days to flowering, and DMAT: days to physiological maturity.

Degree of seed shriveling showed highly statistically significant low degree of positive corelatedness with glume color (r=0.14, p<0.01) but high degree of negative relationship with thousand kernels weight (r=-0.84, p<0.01). Thousand kernels weight showed low degree of positive relationship with seed size (r=0.35, p<0.01). Plant height showed statistically highly significant low degree of positive association with thousand kernels weight (r=0.15, p<0.01).

Spike length showed moderate positive association with number of spikelets per spike (r=0.51, p<0.01), number of florets per spikelet (r=0.21, p<0.01), number seeds per spikelet (r=0.18, p<0.01), number seeds per spike (r=0.3, p<0.01) and glume color (r=0.18, p<0.01); whereas it showed negative association with glum hairiness (r=-0.12, p<0.01) and tillering capacity (r=-0.08, p<0.01). Number of spikelets per spike varied positively with number of seeds per spike (r=0.33, p<0.01), glume color (r=0.14, p<0.01); and, negatively with awnedness (r=-0.18, p<0.01). Number of florets per spikelet showed positive association with number of seeds per spike (r=0.51, p<0.01), number of seeds per spikelet (r=0.49, p<0.01), the maximum number of seeds per floret (r=0.109, p<0.01), and glume color (r=0.1, p<0.05); and negative association with seed color (r=-0.11, p<0.01). Number of seeds per spikelet showed high positive relationship with seed number per spike (r=0.93, p<0.01), and negative correlation with seed color (r=-0.08, p<0.05). The Maximum number of seeds per floret exhibited low degree of positive correlation with glume hariness (r=0.1, p<0.05). A low degree of negative relationship between number of seeds per spike and seed color (r=-0.08, p<0.05) was observed. Awnedness showed positive association with glum hairiness (r=0.16, p<0.05).

Seed color and seed size were positively related (r=0.1, p<0.05). Low degree of negative association was recorded between seed size and degree of seed shriveling (r=-0.32, p<0.01) and tillering capacity (r=-0.12, p<0.01).

Seed size which is very important economic trait showed negative relationship with days to tillering (r=-0.08, p<0.05), days to booting (r=-0.08, p<0.05), days to spike emergence (r=-0.08, p<0.05), days to flowering (r=-0.09, p<0.05), and days to physiological maturity (r=-0.08, p<0.05).

Plant height had negative association with days to spike emergence (r=-0.08, p<0.05), days to physiological maturity (r=-0.08, p<0.05), and days to anthesis (r=-

0.08, p<0.05). Glume color showed a low level of association with days to physiological maturity (0.04, p<0.05). Correlation between spike density and tillering capacity was positive and highly significant (r=0.13, p<0.01); whereas it was negative with days to emergence (r=-0.13, p<0.01) and days to tillering (r=-0.12, p<0.01).

4.2.1.3 Principal component analysis (PCA) of the traits

The PCA variable loadings, percentage of variance, and cumulative variance for the first four principal components are given in Table 4.17 and procedural paragraph in Figure 4.10. In this study, the criteria used by Mishra *et al.,* (2015) to select the number of components to be retained that the Eigen values should be more than one, was used. Therefore, in this analysis, the first, second, third and fourth components, PC-1, PC-2, PC-3 and PC-4, had Eigen values of 5.577, 2.821, 1.646 and 1.394, respectively (Table 4.17 and Figure 4.10). Of total variance, 76.35% was accounted for by the first four principal components, PC-1 to PC-4. The first principal component (PC-1) accounted for 37.25% of total variance and had high contributing factor loading from all the phenological traits, namely, days to emergence, days to tillering, days to booting, days to emergence, days to flowering and days to physiological maturity, and low variance for degree of seed shriveling. The second principal component (PC-2), accounting for further 18.8%, indicated primarily the patterns of variation in number of florets per spikelet, number of seeds per spikelet, and number of seeds per spike. The third principal component (PC-3) accounting for additional 11% of the total variation, had lower loading values except for thousand kernels weight. PC4, showing another 9.3% of the total variability, had factor loading from spike length, number of spikelets per spike, and plant height.

Except the maximum number of seeds per floret, all the quantitative traits, analyzed in the present study, contributed high proportion of variance amounting to 76.35% of the total variance observed in 618 Indian wheat accessions.

4.2.2 Clustering of the 618 wheat accessions based on morphological traits

Dendrogram projected on a scale of semi-partial R-Squared values, assembled 618 accessions into five clusters, as illustrated in Figure 4.11. Looking at the dendrogram graphically, it appears that two major clusters were formed. However, the following criteria were applied to to determine optimum number of clusters in the dendrogram.

Table 4.17 Principal components analysis showing the contributions of 15 quantitative traits to the total genetic variations among 618 Indian wheat accessions.

Traits	**PC* 1**	**PC 2**	**PC 3**	**PC 4**
Spike length in cm*	0.017	0.308	-0.073	0.543
Number of spkikelets per spike	-0.003	0.251	-0.091	0.652
Number of florets per spikelet	-0.005	0.484	0.041	-0.337
Number of seeds per spikelet	-0.01	0.531	0.061	-0.309
Maximum number of seeds per floret	-0.003	0.016	0.046	0.057
Number of seeds per spike	-0.013	0.57	0.016	0.007
Thousand kernel weight in gram	-0.175	-0.006	0.666	0.048
Degree of seed shriveling	0.17	0.015	-0.663	-0.071
Plant height in cm	-0.049	-0.03	0.183	0.244
Days to emergence	0.316	0.023	0.108	0.024
Days to tillering	0.405	0.009	0.112	0.017
Days to booting	0.411	-0.006	0.103	0.008
Days to spike emergence	0.41	-0.002	0.097	0.003
Days to flowering	0.411	-0.003	0.098	0.003
Days to physiological maturity	0.409	-0.002	0.095	0.002
Eigenvalue	**5.577**	**2.821**	**1.646**	**1.394**
Difference	**2.756**	**1.175**	**0.253**	**0.351**
Proportion of variance (%)	**37.25**	**18.8**	**11**	**9.3**
Cumulative variance (%)	**37.25**	**56.05**	**67.05**	**76.35**

*Note: PC = Principal component; cm = centimeter.

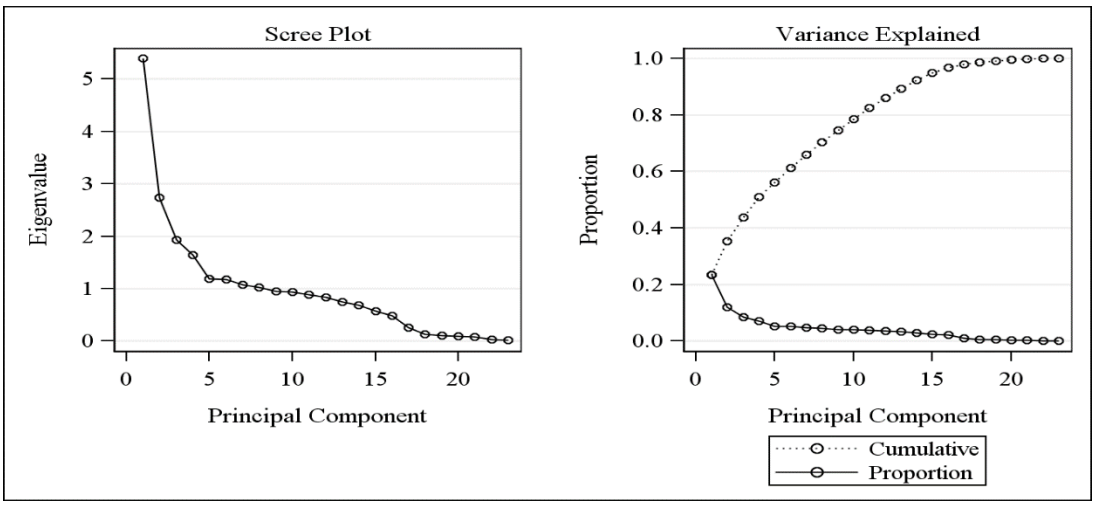


Figure 4.10 Principal component procedural graph showing the number of components to be retained to explain the diversity among 618 Indian wheat accessions. It is based on the concurrence of scree plot (Eigen value) and proportion of variance explained. The plots were constructed using statistical software SAS 9.3.

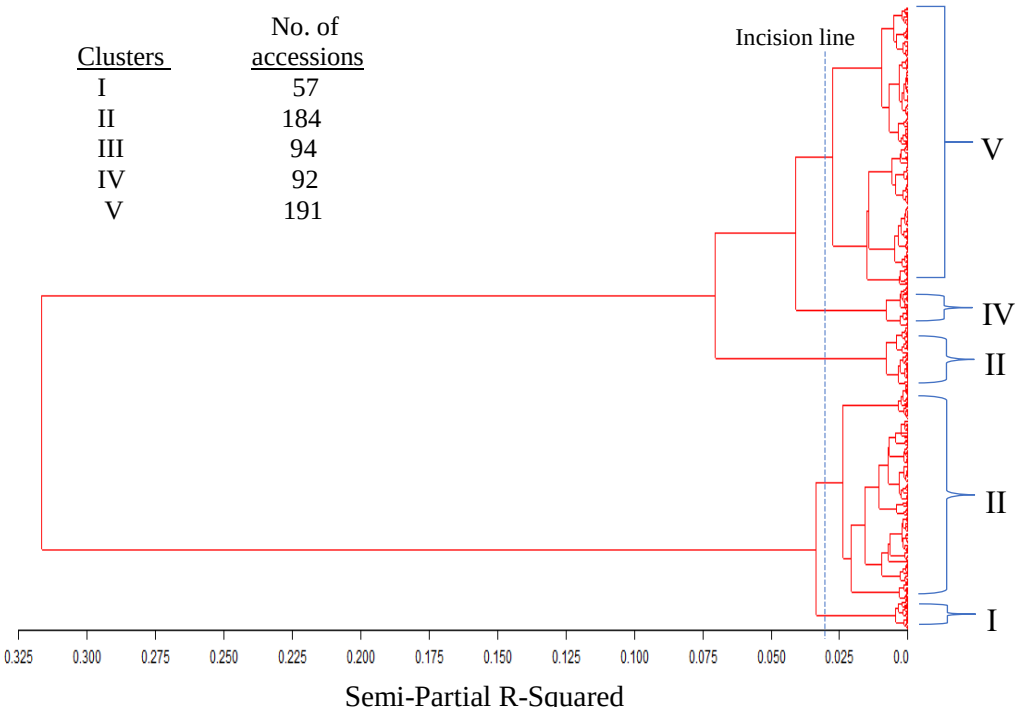


Figure 4.11: Ward's minimum variance dendrogram showing five clusters for 618 Indian wheat accessions based on genetic variations for agro-morphological and phenological characters.

4.2.2.1 Criteria for determining the number of clusters

In the present study, statistical criteria described by Millign and Cooper (1985) for determining number of clusters was deployed and the results are depicted in Figure 4.12. Statistically optimal number of clusters was determined based on the concurrence of these three criteria: (1) Cubic clustering criteria (CCC) curve should have the first peak with sudden fall; (2) Pseudo F curve should be declining slowly; and, (3) Pseudo T-squared curve should be showing a steady phase. An incision line was drawn as shown in Figure 4.12 such that all the three criteria are satisfied, and the number of clusters was determined to be five; based on the interception of this incision line on the x-axis.

4.2.2.2 GT-biplot and Minimum spanning tree

Pictorial display of clustering of 618 accessions using GT-biplot (Figure 4.13) and minimum spanning tree (Figure 4.14) further confirmed the presence of five clusters. In addition, GT-biplot provided for visualization of genetic correlation among genotypes and understanding of genotype-by-trait relationships.

4.2.2.3 Intra-cluster variability

Five clusters were further characterized for their intra-cluster variability (Table 4.18), as revealed by values of coefficient of variation (CV) with respect to each of the investigated traits. Cluster-I did not vary much for any of the traits; cluster-II, varied highly for days to booting (CV, 51.51), and slightly for days to tillering (CV, 24.63) and days to emergence (CV, 41.47); cluster-III showed high CVS for glume color (87.25), days to flowering (64.54) and days to physiological maturity (60.46); and, was slightly variable for degree of seed shriveling (CV, 23.87); cluster-IV, varied highly for awnedness (CV, 89.38) and moderately for growth habit (CV, 31.74); and, cluster V exhibited high variability for number of florets per spikelet (CV, 53.64) and number of productive florets per spikelet (CV, 51.30). All the five clusters exhibited comparable variability for spike length, number of spikelets per spike, number of seeds per spike, glume hairiness, spike density, thousand kernels weight, seed size, plant height, tillering capacity, and days to emergence.

It was interesting to note that cluster-IV was not at all variable for days to physiological maturity. This information is highly useful in making core collections

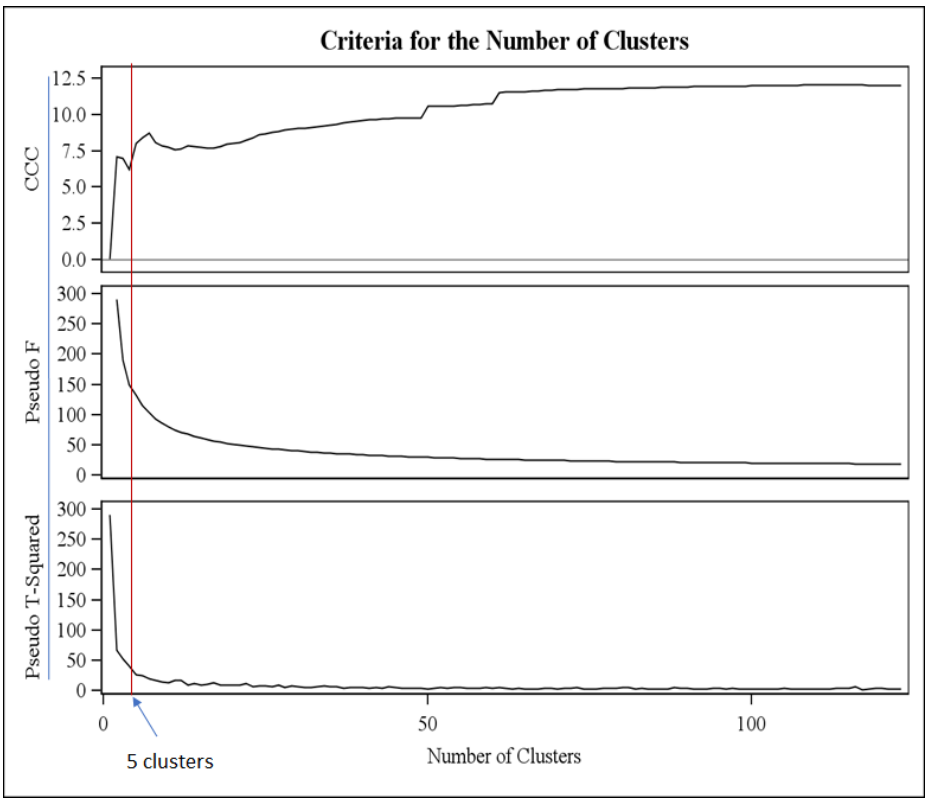


Figure 4.12: Criteria for determining statistically optimal number of clusters in the dendrogram. Five clusters on x-axis is determined based on the concurrence of three graph criteria: sudden fall in cubic clustering criterion (CCC) plot, slow decline of Pseudo F, and steady phase of Pseudo T-Squared. The graphs were constructed using statistical software SAS 9.3.

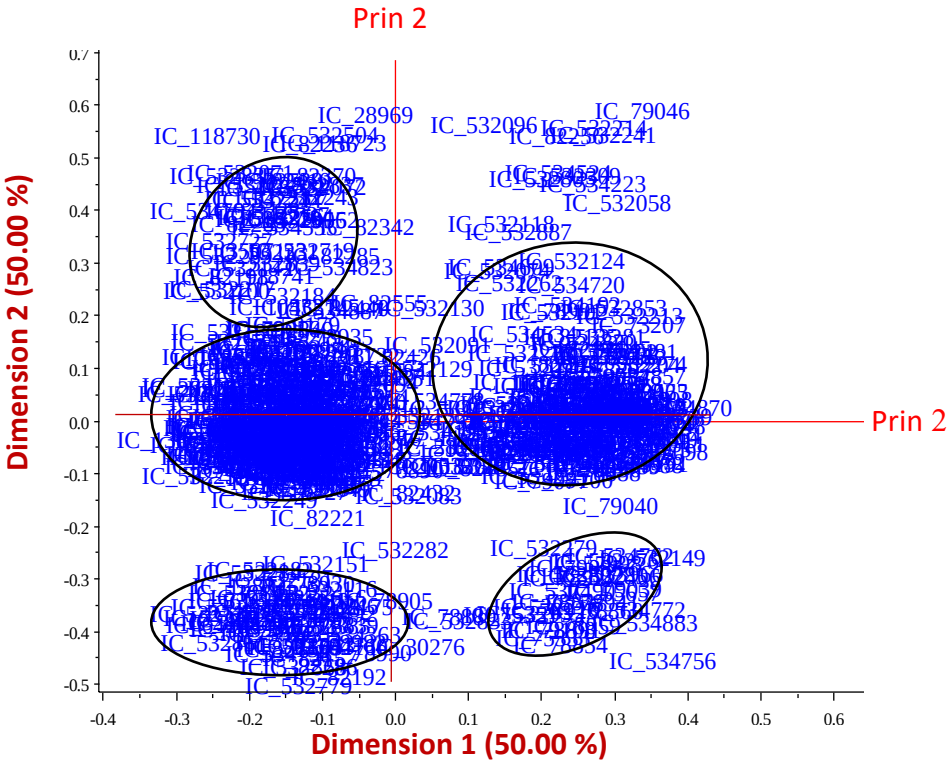


Figure 4.13: Genotype-by-trait (GT)-biplot graphically displaying inter-relationships among 618 Indian wheat accessions based on their variability for 23 agro-morphological traits.

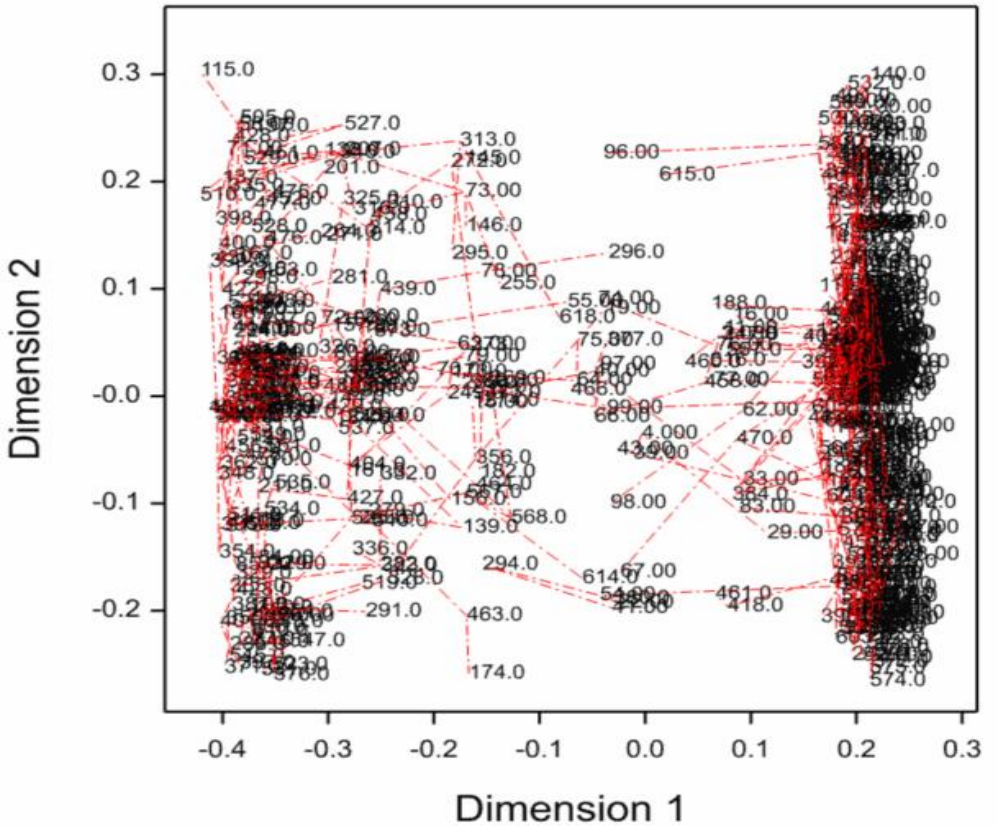


Figure 4.14: Minimum spanning tree showing clustering of 618 Indian wheat accessions into five distinct clusters. Dimension 1 corresponds to clustering and 2 to duplications. Numbers in the graph area indicate the serial number of accessions (Table 3.1).

Table 4.18 Coefficients of variations of 23 agro-morphological traits within each of the five clusters.

Traits	Cluster I	Cluster II	Cluster III	Cluster IV	Cluster V
Spike(ear) length in cm*	35.73	35.49	34.44	29.03	33.25
Number of spkikelets per spike	25.40	25.36	24.33	19.35	20.40
Number of florets per spikelet	37.67	47.74	29.69	46.92	53.64
Number of productive florets per spikelet	34.85	46.92	27.26	47.40	51.30
Maximum number of seeds per floret	31.69	33.81	28.17	32.41	33.56
Number of seeds per spike	40.51	47.03	33.68	48.48	50.57
Awnness	6.56	3.22	8.05	89.38	80.49
Glume colour	54.20	26.96	87.25	54.87	46.17
Glume hairiness	48.10	48.68	35.19	49.68	46.83
Spike density	29.89	29.92	25.26	35.86	36.49
Thousand kernel weight in gram	55.45	43.92	41.99	41.83	56.62
Seed colour	60.39	52.43	70.43	57.69	62.17
Seed size (relative)	16.77	18.51	17.68	16.32	15.52
Degree of seed shriveling	19.18	21.92	23.87	20.73	17.05
Plant height cm	26.86	26.43	29.87	25.52	28.84
Growth habit	24.87	28.17	25.75	31.74	24.36
Tillering capacity	25.47	26.22	27.38	25.31	24.49
Days to emergence	30.75	41.47	40.65	36.84	35.88
Days to tillering	7.87	24.63	22.97	20.74	9.52
Days to booting	17.37	51.51	28.10	27.83	13.06
Days to spike emergence	18.20	17.69	20.05	29.32	16.41
Days to flowering	18.30	58.13	64.54	57.21	19.41
Days to physiological maturity	22.77	19.99	60.46	0.00	23.45

* Note: cm = centimeter.

from the evaluated set of accessions listed according to their cluster-membership (Table 4.19).

4.3 Comparison of molecular- and morphological data-based grouping

Fo the purpose of comparing the molecular and morphological data-based grouping patterns, 108 accessions with passport data (Table3.2) were considered. The details of molecular marker-based grouping and morphological trait-based clustering are given in Table 4.20 and Table 4.21, respectively.

Comparison in terms of overall number of major groups revealed that molecular marker-based method clustered the 618 Indian accessions into 4 major groups (Table 4.20), while that derived based on morphological data revealed 5 major groups (Table 4.21), indicating that there was no correlation between molecular genotyping- and morphological data-based grouping.

Table 4.19 List of 618 Indian wheat accessions identified by their indigenous collection (IC) numbers and their memberships in five major clusters based on their agro-morphological and phenological variations.

Cluster (Size)	Individual accessions, identified by unique indigenous collection (IC) number, constituting clusters
Cluster I (57)	IC_532064, IC_532088, IC_532090, IC_532129, IC_532149, IC_532213, IC_532214, IC_532218, IC_532224, IC_532225, IC_532251, IC_532285, IC_532300, IC_532406, IC_532480, IC_532489, IC_532833, IC_532853, IC_532886, IC_532887, IC_532905, IC_534192, IC_534480, IC_534524, IC_534534, IC_534549, IC_534564, IC_534688, IC_534697, IC_534745, IC_534808, IC_534871, IC_534883, IC_534886, IC_55657, IC_59191, IC 73201, IC 73207, IC 75339, IC 78837, IC 78854, IC 78859, IC 78860, IC 78920, IC_78923, IC_78924, IC_79039, IC_79046, IC_79068, IC 79079, IC 79083, IC 79085, IC 79090, IC 79091, IC_82398, IC_82432, and IC_82555.
Cluster II (184)	IC 104551, IC 118731, IC_118739, IC_118758, IC 28969, IC 28974, IC 29035, IC 30276, IC_532058, IC_532061, IC_532062, IC_532074, IC_532077, IC_532081, IC_532083, IC_532087, IC_532089, IC_532091, IC_532092, IC_532095, IC_532096, IC_532097, IC 532098, IC_532102, IC_532103, IC_532106, IC_532118, IC_532124, IC_532130, IC_532132, IC_532141, IC_532150, IC_532153, IC_532204, IC_532216, IC_532223, IC_532240, IC_532241, IC_532245, IC_532250, IC_532261, IC_532262, IC_532273, IC_532274, IC_532276, IC_532277, IC_532279, IC_532281, IC_532290, IC_532292, IC_532309, IC_532473, IC_532476, IC_532481, IC_532486, IC_532490, IC_532700, IC_532701, IC_532705, IC_532707, IC_532723, IC_532805, IC_532807, IC_532857, IC_532872, IC_532891, IC_532935, IC_532936, IC_532937, IC_533953, IC_533963, IC_533966, IC_533970, IC_533971, IC_534072, IC_534190, IC_534193, IC_534223, IC_534235, IC_534322, IC_534455, IC_534543, IC_534560, IC_534561, IC_534565, IC_534568, IC_534572, IC_534605, IC_534609, IC_534690, IC_534720, IC_534736, IC_534737, IC_534739, IC_534747, IC_534748, IC_534749, IC_534750, IC_534752, IC_534755, IC_534756, IC_534757, IC_534759, IC_534761, IC_534762, IC_534763, IC_534772, IC_534774, IC_534777, IC_534778, IC_534787, IC_534792, IC_534806, IC_534811, IC_534820, IC_534857, IC_534859, IC_534864, IC_534870, IC_534873, IC_534885, IC 57998, IC_59511, IC 73198, IC 73205, IC 75333, IC 75351, IC 78707, IC 78824, IC 78828, IC_78836, IC 78862, IC 78865, IC 78869, IC_78872, IC 78888, IC_78891, IC_78895, IC_78897, IC 78899, IC 78901, IC_78925, IC_78927, IC_78941, IC 78991, IC_79023, IC_79026, IC_79028, IC_79040, IC_79041, IC_79047, IC_79050, IC_79052, IC_79063, IC_79065, IC 79077, IC 79080, IC_79100, IC_79107, IC_79108, IC_79110, IC_79940, IC_82116, IC_82118, IC_82119, IC 82127, IC_82130, IC_82132, IC 82133, IC 82134, IC 82135, IC 82136, IC 82145, IC 82146, IC_82163, IC_82181, IC_82185, IC_82189, IC_82193, IC_82256, IC_82393, IC_82394, IC_82400, and IC_82425.
Cluster III (94)	IC 104537, IC_138898, IC 35069, IC_47478, IC_532059, IC_532071, IC_532072, IC_532084, IC_532110, IC_532121, IC_532144, IC_532151, IC_532184, IC_532186, IC_532209, IC_532210, IC_532217, IC_532227, IC_532232, IC_532237, IC_532242, IC_532247, IC_532248, IC_532265, IC_532268, IC_532272, IC_532284, IC_532310, IC_532474, IC_532478, IC_532483, IC_532485, IC_532487, IC_532708, IC_532728, IC_532729, IC_532837, IC_532842, IC_532847, IC_532849, IC_532863, IC_532868, IC_532880, IC_532930, IC_534323, IC_534334, IC_534419, IC_534430, IC_534435, IC_534481, IC_534509, IC_534553, IC_534555, IC_534557, IC_534743, IC_534744, IC_534751, IC_534769, IC_534776, IC_534798, IC_534802, IC_534805, IC_534819, IC_55578, IC_55656, IC_59179, IC_59612, IC_64238, IC 73493, IC 75327, IC 78096, IC_78715, IC_78928, IC_78930, IC_78933, IC 78935, IC 78948, IC 78960, IC 78981, IC_79042, IC_79043, IC_79053, IC 79095, IC_79102, IC_79106, IC_82164, IC_82167, IC_82169, IC_82236, IC_82285, IC_82286, IC_82375, IC_82387, and IC_82440.
Cluster IV (92)	IC_104564, IC_104573, IC_118722, IC 118727, IC_118736, IC_118737, IC_118740, IC_138893, IC_138896, IC_138897, IC 212185, IC 30284, IC 32520, IC_532085, IC_532120, IC_532136, IC_532138, IC_532139, IC_532140, IC_532142, IC_532146, IC_532188, IC_532229, IC_532249, IC_532256, IC_532258, IC_532259, IC_532263, IC_532264, IC_532267, IC_532286, IC_532297, IC_532475, IC_532482, IC_532497, IC_532698, IC_532699, IC_532726, IC_532736, IC_532738, IC_532775, IC_532790, IC_532821, IC_532832, IC_532851, IC_532910, IC_533962, IC_534189, IC_534432, IC_534566, IC_534582, IC_534721, IC_534723, IC_534741, IC_534742, IC_534754, IC_534768, IC_534770, IC_534771, IC_534773, IC_534786, IC_534794, IC_534822, IC_534854, IC_534855, IC_534858, IC_55636, IC_55652, IC_57844, IC 75328, IC 78937, IC 78987, IC_79038, IC 79066, IC_82125, IC 82126, IC 82161, IC_82179, IC_82187, IC_82195, IC_82197, IC_82198, IC_82199, IC_82200, IC_82204, IC_82206, IC_82221, IC 82385, IC_82372, IC_82381, IC 82385, and IC_82426.
Cluster V (191)	IC_104540, IC 104543, IC_104566, IC_104568, IC_104576, IC_118723, IC 118729, IC_118730, IC_118732, IC_118741, IC 118762, IC_138845, IC_138849, IC_138895, IC 210548, IC 212142, IC 212145, IC 212179, IC 28664, IC 28926, IC_28958, IC 31496, IC_35070, IC 35071, IC_532055, IC_532057, IC_532066, IC_532067, IC_532068, IC_532073, IC_532078, IC_532093, IC_532094, IC_532100, IC_532101, IC_532116, IC_532119, IC_532122, IC_532126, IC_532131, IC_532133, IC_532134, IC_532137, IC_532143, IC_532145, IC_532147, IC_532148, IC_532155, IC_532156, IC_532175, IC_532180, IC_532181, IC_532182, IC_532183, IC_532185, IC_532187, IC_532201, IC_532205, IC_532206, IC_532219, IC_532221, IC_532228, IC_532230, IC_532231, IC_532238, IC_532239, IC_532243, IC_532244, IC_532246, IC_532253, IC_532257, IC_532271, IC_532282, IC_532289, IC_532298, IC_532318, IC_532492, IC_532502, IC_532503, IC_532504, IC_532697, IC_532704, IC_532719, IC_532725, IC_532727, IC_532737, IC_532768, IC_532773, IC_532779, IC_532784, IC_532787, IC_532794, IC_532811, IC_532813, IC_532815, IC_532835, IC_532839, IC_532841, IC_532855, IC_532923, IC_532934, IC_533954, IC_533964, IC_533965, IC_534123, IC_534197, IC_534363, IC_534451, IC_534554, IC_534556, IC_534562, IC_534758, IC_534760, IC_534764, IC_534765, IC_534766, IC_534767, IC_534775, IC_534814, IC_534823, IC_534884, IC_534887, IC_55507, IC_55593, IC_55617, IC_55659, IC_55664, IC_55685, IC_57983, IC_59544, IC 73210, IC 73215, IC_73232, IC_78728, IC 78832, IC 78839, IC 78843, IC_78877, IC 78889, IC_78890, IC 78905, IC 78908, IC 78911, IC 78912, IC 78914, IC 78915, IC 78918, IC 78919, IC_78931, IC 78936, IC_78940, IC 78945, IC 78968, IC 78982, IC 78990, IC 78996, IC 78999, IC_79022, IC_79055, IC_79056, IC 79062, IC 79067, IC 82147, IC 82156, IC 82157, IC 82158, IC 82159, IC 82160, IC_82165, IC_82177, IC 82180, IC_82190, IC_82192, IC_82210, IC_82247, IC_82257, IC_82259, IC_82263, IC 82303, IC_82338, IC 82342, IC_82344, IC_82367, IC_82369, IC_82370, IC_82371, IC_82377, IC_82386, IC_82388, IC_82410, and IC_82421.

*Note: Dark Red: Gujarat, Red: Rajasthan, Yellow: Himachal Pradesh, Light Green: Uttar Pradesh, Light blue: Maharashtra, Green: Jammu & Kashmir, Purple: Madhya Pradesh, Blue: Karnataka, Black: Unknown.

Table 4.20 Grouping of 108 Indian wheat accessions with passport information into four major groups according to molecular data-based STRUCTURE analysis.

Major groups	GJ		RJ		HP		UP		MH		J&K		MP		KA		Total
	P	A	P	A	P	A	P	A	P	A	P	A	P	A	P	A	
A	0	0	2	1	9	2	2	1	2	2	0	1	0	0	1	0	23
B	2	0	10	1	2	1	5	1	0	0	1	0	2	0	0	0	25
C	0	0	12	0	2	0	3	0	3	2	0	0	0	0	1	0	23
D	0	0	1	0	2	0	32	0	0	0	1	0	0	0	1	0	37
Total	**2**	**0**	**25**	**2**	**15**	**3**	**42**	**2**	**5**	**4**	**2**	**1**	**2**	**0**	**3**	**0**	**108**

*Note: GJ: Gujarat, RJ: Rajasthan, HP: Himachal Pradesh, UO: Uttar Pradesh, MH: Maharashtra, J&K: Jammu and Kashmir, MP: Madhya Pradesh, KA: Karnataka, P: Pure, A: Admixture.

Table. 4.21 Grouping of 108 Indian wheat accessions with passport information into five major groups according to morphology based dendrogram.

Major Groups	GJ*	RJ	HP	UP	MH	J&K	MP	KA	Total
I	0	0	3	8	2	0	0	0	13
II	0	10	2	13	3	2	0	0	30
III	0	2	2	5	0	0	0	1	10
IV	0	3	4	2	1	1	1	0	12
V	2	11	8	16	3	0	1	2	43
Total	**2**	**27**	**19**	**44**	**9**	**3**	**2**	**3**	**108**

*Note: GJ: Gujarat, RJ: Rajasthan, HP: Himachal Pradesh, UO: Uttar Pradesh, MH: Maharashtra, J&K: Jammu and Kashmir, MP: Madhya Pradesh, KA: Karnataka.

CHAPTER 5

DISCUSSION

5. DISCUSSION

Wheat (*Triticum* spp. L.*)* is the oldest and one of the most important cereal crops. Wheat cultivation has been practiced since ancient times, 10000-8000 BC, (Heun *et al*., 1997; Nesbitt and Samuel, 1998; Luo *et al*., 2007). It is the staple food for majority of the world's population and, therefore, it is the most widely grown crop. Wheat cultivation in India started 5000 years ago (Feldman, 2001). Three types of wheats are grown in India: 1) *Triticum aestivum* (Bread wheat), 2) *Triticum durum* (Durum/Macaroni wheat), and 3) *Triticum dicoccum* (Dicoccum wheat).

Even though huge size, complex nature and allohexaploidy of the wheat genome has been a bottle-neck for the genetic method of increasing its productivity, the recently decoded genome sequence (International Wheat Genome Sequencing Consortium, 2014) and deployment of a variety of diversity analyses, have contributed to furthering our understanding of wheat genome (Khan *et al*., 2015). Based on the review of the literature up-to-date, it is evident that there is no prior work on deployment of fluorescence-labelled SSR marker mediated high-precision genotyping of Indian wheat landrace accessions to study genetic diversity, population structure and phylogeny. Hence, the present study entitled 'Study of population structure and phylogeny in Indian wheat (*Triticum* spp.) landraces using DNA markers', is the first attempt to assess and understand genetic diversity, population structure and phylogeny among a large panel of wheat genotypes, comprised of 618 Indian wheat accessions collected from across India (Figure 3.1), using 42 fluorescence-labelled SSR markers, evenly distributed throughout the wheat genome, covering both short- and long-arms of all the homologous chromosomes spread across three homoeologous groups.

5.1 Genetic diversity

The term "diversity" refers to the range of variations, variety or differences found among a set of parameters (morphological or molecular) of different individuals and/or populations. Such variations contribute to biodiversity which specifically refers to the variety and variations within assemblages including the genetic differences among them; whether naturally occurring or induced artificially. Biodiversity has been defined as the variety of life and its processes that encompasses various living organisms, the communities, and ecosystem in which they occur (Keystone-Center, 1991).

The basic building blocks of diversity are the genes contained in plants and animals that, by their variations, can enable the whole organisms to adapt to the changing environment. Plant genetic diversity is a useful parameter that can be transmitted genetically from parents to their offsprings and is the source of tremendous variations in plants which support all other forms of life on the earth, and cover a wide range at the evolutionary and ecological levels. The diversity in plants has been the basis for food and other human needs for millennia; and, it continues to be so for the development of plant characters useful to unending human needs and adapting to ever changing environment. (Tanto and Demissie, 1996). It is important for broadening the genetic base and may be exploited *via* breeding (Melchinger, 1999).

Gene diversity assessed in the present study that ranged from 0.88-0.93 with a mean of 0.9 (Table 4.1), was higher than that of the previous reports on wheat diversity studies using SSRs. Wang *et al*. (2007) reported an average genetic diversity index of 0.60 with three subgenomes having values in the range of 0.05 to 0.86, in their analysis of genetic diversity among 52 accessions from three unique Chinese wheat varieties using 206 SSR marker loci. Nei's genetic diversity (H) of 0.2827 and null Shannon's information index (I) were reported by Abouzied *et al*. (2013) in their study involving 45 wheat cultivars comprised of two populations, 15 hexaploids (*Triticum aestvium*) and 30 tetraploids (*Triticum durum*), using 11 SSR markers. Arora *et al*. (2014), while studying genetic diversity among 319 Indian wheat varieties, deploying 16 SSR markers, reported genetic diversity ranging from 0.63 to 0.56. High genetic diversity found in the present study is attributed to involvement of a large number of genotypes (618) and deployment of fluorescent-labelled SSRs for high resolution genotyping.

5.2 Indian wheat landraces

Genetic diversity in populations of wild precursors of wheat has been eroded by domestication, in which a limited range of diversity is present in landraces that were initially selected and adopted for cultivation (Figure 2.2). Subsequent breeding has drawn on a limited range of the variations, present in landraces, to produce the elite cultivars that are used in the modern agriculture (Bevan *et al*., 2017).

Access to a range of genetic diversity is critical to the success of breeding programs. The global effort to assemble, document, and utilize these resources is enormous; and, the genetic diversity in the collections is critical to the world's fight

gainst hunger. The introgression of genes that reduced plant height and increased disease and viral resistance in wheat provided the foundation for the ‘‘Green Revolution’’ and demonstrated the tremendous impact that genetic resources can have on production (Hoisington *et al*., 1999).

Thus, agro-morphological characterization of crop germplasm collections is very essential preliminary step for quantifying the extent of genetic diversity within accessions; and, detecting duplications within germplasm collections, and acts as a starting point for the establishment of a “core collection” where the whole variability spectrum of a species in the collection is represented in a small subset of accessions (Zarkti *et al.* 2012). Little attempts had been made in past to identify genetic potential of Indian wheat accessions.

The plant material of the present study comprised of 618 Indian wheat landrace accessions (Table 3.1). Among these accessions, 597 were bread wheats (hexaploid) and 21 were macaroni or durum (tetraploid) wheats. The selected accessions represented almost all the wheat growing regions of India (Figure 3.1). Geopolitical State from which wheat landraces were collected is given Table 3.2 for 108 accessions (89 hexaploids and 19 diploids).

In the present study, the number of genotypes (618) utilized for SSR-based diversity assessment is relatively larger than those reported of previous studies. For instance, a study by Wang *et al*. (2007) involved 52 accessions from three unique Chinese wheat varieties; forty-five wheat cultivars comprising of two populations, 15 hexaploids (*Triticum aestvium*) and 30 tetraploids (*Triticum durum*), were studied by Abouzied *et al*. (2013); 319 Indian wheat varieties were studied by Arora *et al*. (2014); and, Mwale *et al*. (2016) investigated genetic diversity of 60 wheat cultivars.

5.3 DNA markers

While Morphological, biochemical and DNA markers can be deployed for diversity studies (Mohammadi and Prasanna, 2003; Sudre *et al*., 2007; Goncalves *et al*., 2009; Hailu *et al.,* 2010; Hafida *et al*., 2012; Mengistu *et al.,* 2015), the DNA markers are advantageous over the other types, for they are not affected by environmental factors and gene interactions. In addition, for they involve advanced methods of multiplexing and high-resolution genotyping (Binneck *et al*., 2002; Antonio *et al*., 2004; Saker *et al*.,

2005; Goncalves *et al*., 2008; Souza *et al*, 2008), they are useful for detailed study of genetic diversity involving the large number of genotypes or accessions.

In the present study, for achieving high resolution genotyping, 42 fluorescent-labelled SSR markers and capillary gel electrophoresis were deployed. The number of fluorescent-labelled SSR markers used in the present study are more than what is reported by other researchers. In a genetic discrimination study reported by Zhu *et al*. (2011) while studying 8 closely related wheat cultivars, 5 ISSR markers and 2 fluorescently labelled SSR markers were deployed. In a report by Spanic *et al*. (2012), 24 fluorescent-labelled microsatellite (SSR) markers are utilized to analyze molecular genetic diversity of 30 wheat genotypes.

Therefore, the present study is unique in that it not only deployed 42 fluorescent-labelled SSR markers for the high-precision genotyping but also involved a large panel of 618 Indian wheat landrace accessions.

5.4 Allelic variations among marker loci

Analysis of variance (ANOVA) helps to infer the reliability of the statistical analysis of experimental results. Thus, two crucial factors are needed to be sonsidered in the analysis of wheat amplicons (alleles). Firstly, due to the complex and hexaploid nature of bread wheat genome, enormous duplication and translocation events may result in multi-locus nature of SSR amplicons. Therefore, ruling-out the possibility of multi-locus amplicons considered for analysis is a very crucial prerequisite. In the present study, multi-locus amplicons were not expected in any of the studied SSR loci, as all the 42 SSR markers were selected for single-locus amplification, after repeated confirmation, through 4% agarose gel as well as capillary gel electrophoresis, before their deployment for high-resolution genotyping of the accessions. Secondly, the failure of amplification can be attributed either to presence of the null allele among concerned genotypes or to a faulty protocol, or reagents/polymerase, at any stage of the method from DNA isolation through PCR to gel/capillary electrophoresis. However, after the repetition of confirmatory experiments using different sets of DNAsamples; PCR-reagents/enzymes and involving positive and negative controls; it was proved that there was a consistency in the absence of amplificons in the same panel where the negative control failed, and the positive control amplified. This is very important in view of the presence of null alleles, as the missing information is considered in error reduction

procedure. In the present study, one-way analysis showed that F-value was uniformly one for every locus (Table 4.1), suggesting that the genotyping data obtained were statistically reliable to study the extent and pattern of genetic diversity (i.e., gene and genomic diversity) in a set of 618 Indian wheat accessions by genotyping with 42 marker loci. In addition, heterozygosity estimated was zero, indicating that all the 42 SSR marker loci amplified in a panel of 618 genotypes were fixed. Therefore, a panel of 618 Indian wheat landraces involved in this study are homozygous and heterogeneous while source populations were presumably homogeneous.

In the present analyses, a total of 24779 amplicons, constituting 92.47% of total 25956 expected, were detected (Table 4.3). The remaining 1177 amplicons (4.53% of expected) that did not show-up (Table 4.3) correspond to null-alleles. Among 24779 detected amplicons, 771 (3.11%) were polymorphic. The total number of amplicons detected in the present study is higher than those of other reports. Chen *et al*. (2012) reported a total of 1358 amplicons in a set of 90 Chinese winter wheat accessions genotyped by 269 SSR markers loci; and, Abouzied *et al*. (2013) reported a total of 3840 amplicons observed uponon 6% denaturing polyacrylamide gel electrophoresis (PAGE).

In the present study, forty-two SSR marker loci revealed a total of 771 size-based polymorphic alleles, with an average of 18.36 per locus, with amplicon-sizes ranging from 104 bp to 422 bp (Table 4.1 and Table 4.2). The highest number of 26 alleles per locus was observed at SSR locus *xgwm136* (located on 1-AS) followed by 25 per locus at *xgwm497* (1-AL). The lowest number of 12 alleles per locus was recorded for *xgdm141* located on 6-DS. The number of major alleles per locus ranged from 5.72 at *xgwm637* (4-AL) to 1.76 at *xgwm183* (3-DS), with an average of 3.27 alleles per locus. The number of minor alleles per locus was highest (21.84) at *xgwm136* (1-AS), followed by 21.25 and 20.7 at *xgwm497* (1-AL) and *xgwm459* (6-AS), respectively. The lowest average number of 9.75 minor alleles per locus was present at *xgwm349* (2-DL), followed by 10.2 and 10.5 at *xgdm141* (6-DS) and *xgwm484* (2-DS), respectively. The number of polymorphic loci (771) detected in the present study is more than those of previous reports. Wang *et al*. (2007), in their analysis of genetic diversity in 52 accessions using 206 SSR marker loci, reported a total of 488, 472 and 308 allelic variants in 31 Yunnan, 15 Tibetan and 6 Xinjiang wheat accessions, respectively. In a recent report, Mwale *et al*. (2016) reported a total of 276 polymorphic

alleles across 60 SSR markers amplified in 60 wheat cultivars. Therefore, the findings of the present study corroborated the report of Chao (2009) that asimple sequence repeats are very informative due to their co-dominant nature, and, additionally, they contain more allelic diversity per marker; and, they have higher mutation rates than the other marker types have.

Therefore, a large number of total amplicons as well as polymorphic alleles, observed in the present study, entail the utility of SSR markers and fluorescent-labelling, deployed in the present study, for assessment of diversity and phylogenetic relationship among Indian wheat landraces.

5.4.1 Different categories of polymorphic alleles

Based on frequency of their occurrence among 618 Indian wheat accessions, the 771 polymorphic alleles were further partitioned into 139 major- and 632 minor alleles, constituting 17.89% and 82.1%, respectively (Table 4.3). Minor alleles (< 5%) found in this study represent rare alleles. Out of 632 rare alleles, 59 were unique for they occurred in only one of the 618 genotypes.

Out of 618 alleles expected for each locus, an average of 28 null-alleles (4.53%) and 590 amplified alleles (95.47%) are detected. The highest polymorphism information content (PIC) of 0.93 was observed in each of the two markers, xgwm494 (3-AL) and xgwm459 (6-AS) (Table 4.1); while the lowest PIC of 0.82 was exhibited by the marker xgwm484 (2-DS). An overall gene diversity, computed based on Shannon-wiener diversity index, revealed by 42 SSR loci among 618 genotypes, was 0.9. The highest gene diversity of 0.93 was recorded for each of the two 2 loci, *xgwm494* (3-AL) and *xgwm459* (6-AS), while most of the loci (12) had 0.91. The lowest gene diversity index of 0.84 was recorded for *xgwm484* locus (2-DS), followed by 0.85 in case of *xgwm544* (5-BS).

5.4.1.1 Null alleles

Null alleles are impacted by three factors: one, unstable microsatellites-flanking regions; two, microsatellite-unit-repeat length; and, lastly, SSR-motif complexity. The last two factors influence mutation rate of the microsatellite repeat region (Jin *et al*., 1996; Chakraborty *et al*., 1997). Several studies demonstrated and reported that the microsatellites-flanking regions may be less stable than other genomic regions (Angers and Bernatchez, 1997; Grimaldi and Crouau-Roy, 1997; Meglecz *et al*., 2004), whereas

microsatellite-unit-repeat length and SSR-motif complexity have no correlation with null-allele frequency (Li *et al*., 2003a). In the present study, on an average, each of the 618 genotypes had two null alleles, constituting 4.76% of the total 42 expected amplicons per genotype. On the other hand, each locus had an average of 28 (4.53%) null alleles among total 618 expected alleles per locus.

5.4.1.2 Unique alleles

Information on unique allele entails applications of SSR markers in DNA-finger printing of wheat landraces. Among 42 SSR markers loci surveyed in a set of 618 Indian wheat landrace accessions, 35 loci showed a total of 59 unique alleles, and the number of alleles per locus ranged from 4 to 0. While majority (17) loci possessed one allele per locus, seven did not show any unique locus. All the 14 SSR loci belonging to subgenome-A showed at least one unique allele per locus (Table 4.4). Indian wheat accession IC-57983 had the highest number of 14 unique alleles, followed by 13 unique alleles per genome observed in the accession IC-31496, while IC-118758 recorded 12. Four unique alleles: of 160 bp, 185 bp, 321 bp and 148 bp in size; were revealed by 4 loci: *xgwm497*, *xgwm185*, *xgwm321* and *xgwm497*; located on long arm of chromosome 1 of subgenome A in accessions IC-82370, IC-82440, IC-212145 and IC-212185, respectively.

5.4.2 Genetic diversity among homoeologous groups

The complex genome of wheat has 7 homoeologous groups, each having 3 chromosomes belonging to three distinct subgenomes: AA, BB and DD. The details of gene diversity among homoeologous chromosomes, as revealed by 42 SSR marker loci amplified among 618 Indian wheat accessions, are furnished in Table 4.5. Each homoeologous group was investigated for gene diversity deploying 6 SSR marker loci.

5.4.2.1 Polymorphic alleles

Each homoeologous group had, on an average, 110 polymorphic alleles for 6 SSR loci, whereas 7 homoeologous group together had 128.35 per locus (Table 4.5). Group 1 had the highest number of 117 polymorphic alleles with a mean of 19.5 alleles per locus, followed by groups 3 and 7, which had 114 and 111 polymorphic alleles with means of 19.0 and 18.5 alleles per locus, respectively. The lowest polymorphic alleles of 105 with an average of 17.5 alleles per locus was observed in the homoeologous group 2. Each of the groups 4 to 6 had 108 polymorphic alleles with a mean of 18 per

locus. Each of the 7 homoeologous group had a minimum of 100 alleles for 6 SSR loci. Marker loci in group 1 exhibited the highest number of alleles per locus, in the range of 15-26, while the lowest number of 12 alleles per locus was detected in the homoeologous group 6.

5.4.2.1.1 Major alleles

On an average, each group had 19.7 major alleles (i.e., amplified in > 5% of 618 genotypes) corresponding to 6 SSR loci, whereas it was 23.7 per locus for all the 7 homoeologous groups (Table 4.5). Group 2 had the highest number of 23 major alleles with a mean of 3.83 alleles per locus, followed by groups 4 and 5, which equally had 22 major alleles each with a mean of 3.67 alleles per locus. The lowest major alleles of 16 with a mean of 2.67 per locus was observed in the homoeologous group 6, followed by 17 and 19 major alleles with means of 2.83 and 3.17 per locus in homoeologous groups 3 and 1, respectively. When compared across the 7 homoeologous groups, group 2 had the highest number of 23 out of 138 major alleles which was above the group average (19.7). Group 4 had the highest average number of 5.72 major alleles per locus, while the lowest (1.76) was detected in group 3.

Major alleles occur in more than 5% of individual genotypes, in the entire panel of 618 accessions, and, therefore, contribute majorly to gene diversity. (i.e., variability among allelic-loci).

5.4.2.1.2 Minor alleles (Rare alleles)

Minor (rare) alleles are present in 5% or less; and, therefore distinguishe the subgroups of genotypes harboring them from the rest, as waell as the genotypes harbouring them from one another. However, their contribution to gene diversity is minor.

Each group had, on an average, 90.43 minor alleles for 6 SSR loci, whereas there was an average of 105.5 per locus across the across all the 7 homoeologous groups (Table 4.5). The homoeologous group 1 had the highest number of 97.95 minor alleles with a mean of 16.33 alleles per locus; followed by groups 3 and 6, which had 97.23 and 92.47 polymorphic alleles with means of 16.21 and 15.41 alleles per locus, respectively. The lowest number (82.12) of minor alleles with a mean of 13.69 alleles per locus was observed in the homoeologous group 5 followed by the group 4 that had 86.05 polymorphic alleles with a mean of 14.34 polymorphic alleles per locus. The

highest number of 21.84 minor alleles was recorded in the homoeologous group 1, while the lowest number of 9.75 is shown by group 2.

All the groups revealed mostly the same degree of diversity, despite the fact that a large number of minor alleles were present. A total of 633 minor alleles suggested the presence of many rare alleles in the studied panel of landraces.

5.4.2.1.3 Unique alleles

A total of 59 unique alleles were found in all the homoeologous group, with an average of 8.42 per group for all the 6 marker loci (Table 4.5). The highest number of 10 unique alleles per group was recorded in two groups 3 and 5. Homoeologous groups 1, 2 and 6 had the lowest number of 7 unique alleles each. Each of the groups 4 and 7 recorded 9 unique alleles. Information on unique allele entails the applications of SSR markers, used in the present study, in DNA-finger printing of wheat landraces.

5.4.2.1.4 Null alleles

Null alleles amounting to 1177 (4.53% of the alleles sampled; Table 4.3) were distributed across all the homoeologous groups, with an average of 168 per group. The highest number of 292 null alleles were found in group 1, followed by group 6 that had 207 null alleles. Homoeologous group 2 had the lowest number of 85 null alleles, followed by group 5 that showed 93 null alleles. As the null alleles were confirmed after ruling-out of the other reasons for failure of PCR-amplification through repeated empirical experiments, they also played a key role in the revelation of the extent and pattern of the diversity that exist among 618 Indian wheat accessions.

5.5 Gene diversity of Indian wheats at chromosome level

5.5.1 Gene diversity of Indian wheats at homoeologous groups

5.5.1.1 Distribution of Informative markers among homoeologous groups

Each of the seven homoeologous groups had high degree of polymorphism information content (PIC) ranging from 0.910 in groups 1 to 0.873 in group 2, with an average of 0.893 per group (Table 4.5). The Homoeologous groups 3, 4, 5, 6 and 7 had PIC values 0.910, 0.887, 0.883, 0.902 and 0.897, respectively. High PIC values distributed among 7 homoeologous groups indicate the usefulness of the 42 SSR markers, deployed in this study, for deciphering genetic diversity in studied set of 618 Indian wheat accessions.

Microsatellite (SSR) markers used in the present study demonstrated a high degree of polymorphism compared to those of previous reports on barley and wheat. Varshney *et al*. (2007) combining 4 AFLP markers with 16 EST-SSRS and 15 EST-SNPs to analyze diversity among 43 wild, 35 cultivated and 12 elite barley lines, reported an average PIC 0.593. Kumar *et al.* (2016a), while studying 54 wheat genotypes (41 indigenous and 13 exotic) using 39 SSR markers, reported an average PIC value of 0.29 ranging from 0.03 to 0.49, indicating the lower level of genetic diversity than that discovered in the present study.

The high PIC values detected in this study are due to the high-resolution genotyping involving fluorescent-labelled SSR markers as well as the involvement of many genotypes. Fahima *et al.* (1998) showed that diversity in the wild species can be detected using a relatively small number of microsatellite markers. But in the present study, though relatively small numcer of SSR markers were used to study diversity in wheat landraces, their polymorphism information content values were enhanced by the high-resolution genotyping. Therefore, though a smaller number of markers are used in the present study, high-resolution genotyping by using fluorescent-labelled microsatellite markers and capillary gel electrophoresis rendered them a high capability, and, therefore, they are highly useful in revealing genetic diversity of Indian wheat landraces that can be utilized for wheat genetic improvement.

5.5.1.2 Gene diversity among homoeologous groups

Gene diversity index is a measure of gene diversity (allelic richness) of each marker locus analyzed and it reveals genetic diversity when combined for several loci. The comparative analysis of gene diversity among the seven homoeologous groups, each analyzed with 6 marker loci, revealed an existence of high level of genetic diversity, with an average gene diversity index of 0.901, for all the 7 homoeologous groups (Table 4.5). Homoeologous group 3 showed the highest genetic diversity of 0.912, while group 2 had the lowest (0.883). Further, homoeologous groups 3 (0.912), 6 (0.91) and 1 (0.907) are more diverse than groups 4 (0.895), 5 (0.893) and 2 ((0.883). Further, the average genetic diversity index for homoeologous groups 3 (0.912), 6 (0.91), 1 (0.907) and 7 (0.905) was higher than the group average of 0.901, whereas those for 4 (0.895), 5 (0.893), and 2 ((0.883) were lower.

5.5.2 Gene diversity of Indian wheats at homologous chromosomes

Each of the 21 homologous group was investigated with 2 SSR marker loci, one each on their short- and long arms. Considering all the 21 homologous chromosomes, 1A showed the highest number of 28 polymorphic alleles for two marker loci while the lowest (28) was in 7D (Table 4.6). Chromosome 1A is the richest in number of polymorphic alleles (25.5), while each of the 5D and 7D had the least (14.0). There was no correlation between the number of polymorphic alleles and gene diversity, suggesting that in addition to polymorphic alleles, the polymorphism information content (PIC) of each marker played role in determining the gene diversity. Chromosomes 3A and 5A had the highest number of 6 unique alleles each, while 2D and 6D had none. While the highest number of 43 rare alleles were detected in chromosome 1A, chromosome 6D had 21. While 5B had no null-allele, the remaining 20 chromosomes harboured at least one null allele each, with 1B having 170.

Although each of the homologous chromosomes 2A, 3A and 6A had the highest average polymorphism information content (PIC, 0.920) for two markers, chromosome 6A alone exhibited the highest gene diversity (GDI, 0.925), followed by 2A and 3A that had the same GDI of 0.920 (Table 4.6). On the other hand, 6A had a greater number of null-alleles (101) than 2A and 3A, each of which had 25, suggesting that null-alleles also contributed to the genetic diversity. In addition, with respect to the mean genetic diversity index, for all the 21 chromosomes and 42 marker loci, most of the homologous groups had the genetic diversity above the group average (GDI, 0.9), whereas it was below the group average for six chromosomes: 2B (GDI, 0.875), 2D (GDI, 0.855), 4B (GDI, 0.895), 4D (GDI, 0.89), 5B (GDI, 0.875) and 6D (GDI, 0.895). However, all the 21 chromosomes had higher level of genetic diversity, indicating that investigated panel of 618 Indian wheat landrace accessions had high genetic diversity, and that the 42 SSR markers and the computational algorithms used in this study were highly informative to decipher the extent and pattern of the diversity. These findings can be utilized in broadening the genetic base of Indian wheat landrace collections, for their exploitation in wheat breeding programs in the face of ever-changing climate and never receding human populations.

5.6 Genetic diversity of Indian wheat landraces at subgenome level

As each of the three sub- or component-genomes exhibits homoeologous relationship, it is pertinent to understand gene- and genetic diversity at subgenome levels too.

5.6.1 Subgenome-A

In the subgenome A, there were a total of 320 polymorphic-, 54.62 major-, 265.38 minor-, 32 unique- and 532 null-alleles, that constitute 41.5 %, 39.72 %, 41.89 %, 54.23 % and 45.43 % of those recorded in the total genome (ABD), respectively (Table 4.7). The average genetic diversity index (GDI) of a set of marker loci is a measure of percent diversity, with respect to investigated loci, explained by them in a set of genotypes used in the analysis (Tamura *et al*., 2007). The average PIC value of 14 SSR marker loci located in subgenome-A was 0.907, while its average genetic diversity index (GDI) was 0.912. This suggests that 14 markers explained the 91.2% of the genetic diversity of subgenome-A corresponding to 14 loci among 618 Indian wheat landraces. With respect to two SSR marker loci on each chromosome, one each on the short- and long-arms, chromosomes 2 and 6 of subgenome-A were more diverse than the group average. Similarly, short arm of chromosome 4 and long arms of chromosomes 1, 3 and 7 were also more diverse than the average diversity of subgenome-A.

5.6.2 Subgenome-B

The subgenome-B had a total of 250 polymorphic-, 46.29 major-, 203.71 minor-, 18 unique- and 365 null-alleles, that comprised 32.4 %, 33.67 %, 32.15 %, 30.5 % and 31.67 % of those detected in total genome (ABD), respectively (Table 4.8). The average PIC value for 14 SSR marker loci located in subgenome-B was 0.887, and an average genetic diversity index (GDI) was 0.895. Therefore, 89.5% of genetic diversity of subgenome-B, with respect to the 14 loci, was revealed by 14 SSR markers amplified among 618 Indian wheat landrace accessions. With respect to two SSR marker loci on each chromosome, one each on the short- and long-arms, chromosomes 1, 6 and 7 of subgenome-B were more diverse than the group average. Similarly, short arm of chromosome 3 and long arms of chromosomes 4 and 5 were also more diverse than the average diversity of subgenome-B.

5.6.3 Subgenome-D

In total, subgenome-D had 201 polymorphic-, 36.59 major-, 164.41 minor-, 9 unique- and 294 null-alleles, that constituted 26.07 %, 26.61%, 25.95%, 15.25% and 22.9% of those detected in the whole genome (ABD), respectively (Table 4.9). Both average PIC value as well as average GDI for 14 SSR marker loci distributed across subgenome-D was 0.89. Thus, 14 SSR markers explained 89% of the genetic diversity of subgenome-D, with respect to 14 loci, in a panel of 618 Indian wheats. With respect to two SSR marker loci on each chromosome, one each on the short- and long-arms, chromosomes 2 and 3 of subgenome-D were more diverse than the group average. Similarly, short arms of chromosomes 4 and 5, and long arms of chromosomes 6 and 7 were also more diverse than the average diversity of subgenome-D.

5.7 Comparison of molecular genetic diversity at subgenome and genome levels

Since three subgenomes had different allelic profiles, and extent- and pattern of gene- and genetic diversities (Table 4.7, Table 4.8 and Table 4.9), it is important to compare the subgenomes *vis-a-vis* whole genome with respect to their diversity.

The subgenome-A had the highest genetic diversity index (GDI) of 0.912 ranging from 0.88-0.93, followed by subgenomes B (0.896) and D (0.890) where GDI varied between 0.85-0.92 and 0.84-0.92, respectively. The results of the present study that subgenome-D is the least diverse is inconsistent with what is reported by Arora *et al*. (2014) in their study of genetic diversity among 319 Indian wheat varieties using 16 SSR markers. However, regarding the subgenomes-A and B, Arora *et al*. (2014) reported that subgenome-B is more diverse than subgebome-A, which is in contradiction with the findings of the present study. In addition, genetic diversity detected by the present study is higher than that reported by Arora *et al*. (2014). The highest genetic diversity reported by them in subgenome-B was 0.63, whereas it was 0.89. Similarly, in subgenome-D, the diversity found in the present study is 0.89 whereas the one reported by them is 0.56.

The present study demonstrated the presence of an high level of genetic diversity, with an average GDI of 0.901, ranging from 0.84 to 0.93 among three genomes, compared to those reported in previous studies. Recently, Mwale *et al*. (2016), while investgating genetic diversity of 60 wheat cultivars using 60 SSR marker

loci, reported that genetic diversity index varied from 0.56 to 0.87. In another report by Chen *et al*. (2012) on genetic diversity and population structure in a set of 90 Chinese winter wheat accessions, surveyed with 269 SSR markers loci, the average genetic diversity index was 0.60, with three genomes having GDI values in the range of 0.05 to 0.86. The high level of genetic diversity demonstrated by the results of the present study can be attributed to both the rich diversity present in the studied set of 618 Indian wheat accessions and advanced method deployed to reveal it. In conventional agarose gel electrophoretic method, two or more amplicons (or alleles), differing by a fewer bases in size may get scored as single allele, negatively impacting the number of polymorphic alleles, and, therefore, limiting the power of revealing diversity. Precise scoring of allele size coupled with high-resolution of polymorphism by fluorescent-labelled markers yielded high values of PIC for the amplified loci. Further, the resultant high values of GDI suggested the presence of high level of genetic diversity in a studied set of 618 Indian wheat accessions.

5.8 Genetic diversity among Indian wheat landraces

The diverse genetic resources with wider adaptation and broad utility have rendered cultivated wheat (*Triticum aestivum* L.) a status of world-wide food crop. Domestication and modern plant breeding have presumably narrowed the genetic base of bread wheat, which could jeopardize the future crop improvement (Reif *et al.*, 2005). Further, it has been argued that modern plant breeding has led to genetic homogeneity and possible vulnerability to biotic and abiotic stresses (Vellve, 1993; Russel, 2000; Roussel *et al*., 2004; Fu, 2005). While there is a multitude of high yielding modern wheat varieties in use, a variety that is successful today can overnight be rendered ineffective by changes in the agro-ecosystem and/or outbreaks of new races of pests and diseases. Landrace cultivars (LC) undoubtedly represent an important source of genetic variation in wheat. One of the prime examples is the use of *Rht* dwarfing genes that became available through the Japanese wheat 'Norin10', derived from the LC Shiro Daruma (Kihara, 1982). Approximately, 266589 hexaploid wheat collections are available worldwide of which 24% constitute landraces (FAO, 2002).

Further, due to an increasing pressure on arable land caused by increasing paces of human population and urbanization, and since there are environmental bottle-necks due to challenges of climate change, crop production needs to be increased to secure future food supplies, while minimizing its impact on ecosystems. The detailed

understanding of plant genomes and genetic diversity is crucial in the face of these challenges (Bevan *et al.*, 2017).

The genetic distance matrix based on genetic dissimilarity indices was derived from allelic information of evenly distributed 42 SSR loci amplified across the genomes of 618 Indian wheat accessions, deploying genetic distance method of Cavalli-Sforza (CS) chord. A simple agglomerative (bottom-up) hierarchical clustering method was used to generate unweighted pair group method with arithmetic mean (UPGMA)-based tree.

Most (510) of the studied set of accessions belonged to the category "unknown" and some (108) had the information about their geographical origin (Table 3.2). These 108 accessions belonged to eight political states of India, *viz*., Gujarat (2 accessions), Himachal Pradesh (18), Jammu and Kashmir (3), Karnataka (3), Madya Pradesh (2), Maharashtra (9), Rajasthan (27), and Uttar Pradesh (44). For easy identification, analysis and discussion, unique indigenous collection (IC) numbers of all the accessions were colour-coded (Table 4.13) as follows: dark red: Gujarat, red: Rajasthan, yellow: Himachal Pradesh, light green: Uttar Pradesh, light blue: Maharashtra, green: Jammu & Kashmir, purple: Madhya Pradesh, blue: Karnataka, and black: Unknown.

Analysis of UPGMA tree revealed that the studied set of the wheat accessions were broadly clustered into 4 main groups: A_1, A_2, A_3 and Gr-A_4, each diverging from an unknown common ancestor (Figure 4.3 and Figure 4.4). Group A2 had a single accession, IC-118758, of unknown origin. On the other hand, Gr-A_4 had the largest number of 579 accessions that were highly diverse and were further grouped into 11 groups at genetic dissimilarity coefficient of 0.7. Each of the 11 groups had different number of sub-groups. Major group A1 had 20 wheat accessions comprising of: 6 accessions, IC-78937, IC-75327, IC-75328, IC-75339, IC-78096 and IC-78860, from Uttar Pradesh; 3 accessions, IC-212145, IC-28974 and IC-73493, belonging to Rajastan; 2 accssessions, IC-32520 and IC-31496, of Madhya Pradesh origin; 3 accessions, IC-73205, IC-73198 and IC-73201, from Karnataka; and, 6 accessions, IC-532887, IC-35070, IC-73232, IC-64238, IC-78933 and IC-78877 of unkown origin. Clad A3 had 18 wheat accessions of which majority (10) of genotypes, *viz*., IC-78987, IC-78982, IC-78981, IC-78990, IC-79091, IC-79090, IC-79077, IC-78996, IC-78991 and IC-82303, belonging to Himachal Pradesh grouped together with two accessions,

IC-78968 and IC-79085, of Uttar Pradesh origin, along with 6 accessions, *viz*., IC-79056, , IC-79065, IC-79068, IC-138898, IC-82338, and IC-82286, of unkown origin.

Thus, there was a total of 14 grooups, at 0.7 coefficient of genetic distance, as described in Table 4.11, *viz*., A_1, A_2, A_3, C_1, D_1, E_1, F_1, G_1, H_1, I_1, J_1, K_1, L_1 and L_2 with membership strengths of 20, 1, 13, 24, 1, 1, 94, 169, 68, 13, 80, 55, 23 and 50, respectively. Among them, three groups, *viz*., A_2, D_1 and E_1, had only one member each, namely, IC-118758, IC-30276 and IC-212185, respectively. Out of remaining 11 groups, the largest group was the G_1 comprising of 169 accessions that included Indian wheat landraces, the geographical origins of most of which are unknown; among those with known origin, three accessions (IC-79095, IC-78999 and IC-82342) belonging to Himachal Pradesh, four accessions (IC82180, IC-104537, IC-82127 and IC-82161) from Rajasthan, two accessions (IC-118729 and IC-73215) of Maharashtra were grouped together in the group G_1 along with one accession (IC-35069) belonging to Karnataka, and two accessions (IC-78832 and IC-78707) from Uttar Pradesh. The second largest and the most diverse group was F_1 comprising of 94 accessions, followed by groups J_1 and H_1 with membership sizes of 80 and 68 accessions, respectively. The smallest multi-member group was I_1 comprising of 13 accessions, followed by groups A_1 and C_1 having 20 and 24 accessions, respectively; the most (10) of which belonged to the political state of Himachal Pradesh. Among the 108 accessions of known origin, most of the landraces (20) belonging to Uttar Pradesh were grouped in group L_1, which had 24 accessions altogether. The goup C_1 inluded most, 18 of 27, of the Indian wheat landraces from Rajasthan.

Except the groups A_2, D_1, E_1 and L_1, none other showed the grouping of accessions according to their origin. While groups A_2, D_1 and E_1 had a single member each, *viz*., IC-118758 of unknown origin, IC-30276 belonging to state of Jammu and Kashmir, and IC-212185 from state of Rajasthan, respectively; group L_1 had a total of 23 accessions the most (20) of which belonged to state of Uttar Pradesh, two, IC-79066 and IC-79079, were from Himachal Pradesh, and one, IC-78941, was of unknown origin.

Overall, there was no group, except three single-membered groups (A_2, D_1 and E_1), that had all the members belonging to a single origin, suggesting that origin of Indian landraces mostly did not determine the genetic relatedness. This reveals the enormous exchange of genetic material among the wheat growing regions of India.

However, IC-118758 of unknown origin, IC-30276 belonging to state of Jammu and Kashmir, and IC-212185 from state of Rajasthan were clustered into distinct groups, A_2, D_1 and E_1, respectively. These accessions can be utilized in broadening the genetic base of Indian wheat landrace collections, for their exploitation in wheat breeding programmes.

5.9 Molecular variations among Indian wheat landraces

Analysis of molecular variance (AMOVA) reveals the partitioning of total diversity into variability within populations and among populations. Selective adaptation towards the local growth conditions specific to the geopolitical state of origin of accessions largely contributes to the within group diversity (Khan *et al*., 2015).

The result of the present study revealed that according to genotypic data for 42 SSR marker loci amplified across 618 wheat genotypes, the molecular genetic variance within populations was 90%, whereas it was 10% among populations (Table 4.12 and Figure 4.5) and partitioning pattern was uniform across all the three subgenomes (A, B and D) as well as in the whole genome (ABD).

Our finding that there existed a high genetic variation within groups than among groups is in concurrence with those of previous reports. Further, variations within populations (90%) recorded in our study is higher than that (83%) recently reported by Soriano *et al*. (2016). Further, still lower level is reported Khan *et al*. (2015) that among ninety-five tetraploid and hexaploid wheat genotypes belonging to India and Turkey, 77% genetic variation was within populations. In a report on Indian wheat varieties, Arora *et al*. (2014) reported 78.5% variation within populations. However, variation among the populations (10%) found in the present study is lesser than what was reported (17%) by Soriano *et al*. (2016).

The results indicated that there exists a significant amount of molecular genetic variance (90% of the total) within groups, suggesting that individual genotypes within each group were genetically diverse among themselves. However, the remaining 10 % genetic variation was among populations, indicating that 9 groups, based on state-level origin of wheat accessions, did not vary much among themselves. This suggests that while 90% of the genetic variability was attributed to genotypc variance, only 10% was attributed to geographic variance. Therefore, total diversity studied in a set of 618 Indian wheat landraces indicated existence of more genetic variability than

geographical variations. Further, it indirectly suggests that aspects of agro-climatic conditions required for better agronomic performance of concerned landrace was not uniform throughout the state of origin. The other probable reason is that there existed neither cultural nor geographical barriers to historical exchange of most of the studied landraces among most places represented by sampling sites. The the other probable reason for small fraction of geographical variability (10 %) in total variability is due to within-group inconsistencies of landrace accessions. Therefore, future collection trips should focus on collecting more acccessions from each of the collection sites than collecting from newer sites.

5.10 Study of population structure and phylogeny among Indian wheat landraces

5.10.1 Model based population structure

Population structure analysis plays a key role in determining population stratification and nonfunctional spurious associations (admixtures) based on an unequal distribution of alleles within groups. As revealed by Evanno plot analysis (Figure 4.6), the K-value was four, suggesting the presence of four subpopulations in a panel of 618 Indian wheat accessions. The number of groups deciphered by STRUCTURE analysis (Figure 4.7) is consistent with number of clads revealed by cluster analysis (Figure 4.3 and Figure 4.4). However, as expected, grouping of members in model-based population structure is not same as that in clustering revealed by dendrogram.

The subpopulation-A was the largest, with 155 accessions having pure ancestry and 21 accessions belonging to admixture, out of total 176 accessions in the group (Table 4.13). On the other hand, the smallest was the subpopulation-C having a total of 135 accessions segregated into 130 pure and 5 admixtures. The higest number of admixture ancestry of 23 accessions was observed in subpopulation-B that had 121 pure accessions out of a total number of 144 accessions. The lowest number of 2 admixture accessions was recorded, in subpopulation-D, along with 161 accessions with pure ancestry out of total 163 accessions.

Threshold value of membership-probability 0.80, used in this study was higher than that (0.60) that reported by Chen *et al*. (2012), suggesting the high level of authenticity of the results obtained in the present analysis.

5.10.2 Principal coordinate analysis (PCoA)

Principal coordinate analysis (PCoA) is an important platform that plays two roles: one, it helps to determine the robustness of the results obtained from STRUCTURE analysis and dendrogram; and, the other, it shows a comparative genetic relationship among individual genotypes through spatial illustration (Liu *et al*., 2013; Khan *et al*., 2015).

In the present study, three-dimensional scaling by principal coordinate analysis (PCoA) separated a panel of 618 Indian wheat landrace accessions into 4 major groups, confirming the result of model-based population structure analysis; and, it was visualized in two-dimensional scatter-plot (Figure 4.8). Further, it was found that the first, second, and third axes explained 4.14 %, 3.34 % and 2.77%, of the total genetic variation, respectively. These results indicated the robustness of the grouping methods used in the present study.

Similar strategy of confirming the dendrogram and STRUCTURE results using PCoA analysis is reported by previous workers. In a study conducted by Sardouie-Nasab *et al*. (2013), all the genotypes were assigned into 3 distinct clusters by UPGMA-based dendrogram and it was corroborated by principal coordinates analysis (PCoA). In another study, Najaphy *et al.* (2012) reported that both marker-derived dendrogram and principal coordinate analyses (PCoA) clustered the studied set of genotypes into five groups. Abouzied *et al*. (2013), in their study of genetic diversity among group of cultivars comprised of *Triticum aestvium* and *Triticum durum*, reported that all the most similar genotypes were grouped into the same cluster; with genotypes from the four distinct regions grouping into two main clusters, duram and bread wheat varieties, as inferred by dendrogram and reiterated by the principal coordinate analysis (PCoA). Xu *et al.* (2016) in their recent study, further clarified that STRUCTURE analysis and multivariate analyses (Cluster and PCoA) were useful means to partition the accessions into subpopulations, with some admixtures, if any, across populations.

Therefore, the results of the present study on cluster and model-based grouping of accessions are useful for exploitation of studied set of 618 Indian accessions in wheat improvement programs.

5.10.3 Phylogenetic relationship among Indian wheat landraces

Phylogenetic analysis of a set of Indian wheat accessions is crucial to identify the evolutionary relationship and genetic relatedness among the genotypes. Results of the present study indicated that the studied set of the wheat accessions were broadly grouped into 4 major phylogenetic clads: A1, A2, A3 and A4, each diverging from an unknown common ancestor (Figure 4.3, Figure 4.4 and Table 4.11). Clad A2 had a single genotype having accession number IC-118758 of unknown origin. On the other hand, clad A4 had the largest number of 579 accessions that were highly diverse, and the dissimilarity coefficient 0.7 grouped them further into 11 clusters. Clad A1 had 20 wheat accessions comprising of: 6 accessions, IC-78937, IC-75327, IC-75328, IC-75339, IC-78096 and IC-78860, from Uttar Pradesh; 3 accessions, IC-212145, IC-28974 and IC-73493, belonging to Rajasthan; 2 accessions, IC-32520 and IC-31496, of Madhya Pradesh origin; 3 accessions, IC-73205, IC-73198 and IC-73201, from Karnataka; and, 6 accessions, IC-532887, IC-35070, IC-73232, IC-64238, IC-78933 and IC-78877 of unknown origin. Clad A3 had 18 wheat accessions of which majority (10) of genotypes, *viz*., IC-78987, IC-78982, IC-78981, IC-78990, IC-79091, IC-79090, IC-79077, IC-78996, IC-78991 and IC-82303, belonging to Himachal Pradesh, grouped together with two accessions, IC-78968 and IC-79085, of Uttar Pradesh origin, along with 6 accessions, *viz*., IC-79056, , IC-79065, IC-79068, IC-138898, IC-82338, and IC-82286, of unknown origin.

However, the grouping pattern had no connection with geographical origin of accessions, suggesting that there was an extensive exchange of genetic material across the geographical regions representing the origins of accessions. Accessions from Uttar Pradesh, Rajasthan and Karnataka, for instance, shared common ancestors and were phylogenetically more related among themselves. Similarly, accessions from Himachal Pradesh were more closely related to those from Uttar Pradesh and distantly related to accessions from Rajasthan as well as those from Madhya Pradesh.

Information of phylogenetic relationship elucidated in the present study is highly useful for understanding the ancestral relationship among Indian wheat landraces. For instance, a landrace with accession number IC-118758 of unknown origin grouped singly in a phylogenetic clad A2, suggesting that it is phylogenetically more divergent from every member of the most diverse and the largest clad 4, having 11 clusters, with a total of 579 accessions. Thus, IC-118758 can be utilized to hybridize

with every member of clad A2 for broadening of genetic base of Indian wheat accessions for utilizing them in wheat improvement programs.

5.11 Genetic diversity among Indian wheat landraces based on morphological traits

Genetic diversity can be estimated based on various criteria: pedigree analysis (Barrett *et al.,* 1998), morphological traits (Schut and Stam., 1997; Maric *et al.,* 1998; Casadesus *et al.,* 2007; Zarkti *et al.,* 2012; Malik *et al.,* 2014) as well as biochemical markers (Cox *et al.,* 1985; Metakovsky and Branlard, 1998) or molecular markers (Rao and Riley, 1994; Karp *et al.,* 1996; Gupta *et al.,* 1999; Manifesto *et al.,* 2001; Pagnotta *et al.,* 2005).

In the present study, diversity of Indian wheat landraces was assessed using morphological traits, in addition to molecular/DNA markers.

5.11.1 Study of agro-morphological diversity

5.11.1.1 Extent and pattern of genetic diversity

Large natural variations were found among the Indian wheat accessions for all the investigated traits. This is an indication of wider phenotypic diversity among the accessions as compared to modern high yielding varieties.

Results of the present study showed higher coefficients of variation (CVs) for 1000-kernels weight (CV, 42.78%), tillering capacity (CV, 55.69%) and number of seeds per spike (CV, 42.78%) (Table 4.14) than those reported by Ali *et al.* (2008) and Sabaghnia *et al*. (2014). Majority of the accessions were of spreading type with medium plan height (90 to 105 cm), intermediate tillering capacity (5 to 10 cm), medium spike length (5 to 10 cm), having 15 to 20 spikelets per spike, single kernel per floret, 25 to 50 seeds per spike, awnedness, glume color purple to black, pubescent glume, intermediate type of spike density, red seed color, intermediate seed size, intermediate seed shriveling, early emergence (12 to 15 days), days to tillering medium (20 to 25 days), days to booting late (75 to 90 days), days to spike emergence early (60 to 75 days), very early flowering ($\leq$ 70 days), early physiological maturity (90 to 105 days), and low thousand kernel weight (20 to 30 g). Two accessions, IC_534777 and IC_534787, recorded the highest 1000-kernel weights of more than 50 gram each.

Estimated phenotypic diversity (H') for individual traits ranged from 0.93, for glume hairiness, to 0.49, for maximum number of seeds per spike (Table 4.15). Analyzed accessions showed high phenotypic diversity (H'>0.60), except for NSS (0.59) and NSSL (0.49), indicating high-levels of polymorphism among studied accessions for evaluated traits.

Values of most of the agro-morphological characters analyzed in the present study were negatively skewed and biased towards higher values of phenotypes. Negatively skewed distributions for these traits might have presumably arisen from rapid pyramiding of genes that conditioned these traits, due to environmental and/or man mediated selection during domestication (Moragues *et al*., 2005; Zarkti *et al*., 2012).

High genetic variability was reported for plant height (Hirachand, 1988), ears per plant (Baisakh and Nayak, 1991), thousand grain weight (Ehdaie and Waines 1989), number of seeds per spike (Amin *et al.*, 1992). In contrary to above, low genetic variability has also been reported for number of seeds per spike (Pawar and Patil, 1989), plant height, days to heading, and days to maturity (Ehdaie and Waines, 1989). Therefore, our findings are of practical significance that can be utilized in wheat genetic improvement programs.

5.11.1.2 Correlation among the 23 morphological traits

Breeding strategies and selection procedure are greatly influenced by relationships among traits of interest (Mishra *et al.*, 2015). Selection would be easier if all the traits of economic importance were positively correlated, as it helps for indirect selection of all correlated traits while selecting for a single trait. Similarly, it would not be too difficult if all the traits are either positively correlated or independently inherited (Mohammadi and Amri, 2011). True challenges for the breeder is the fact that there often exist negative genetic correlations between traits of breeding objectives (Yan and Wallace, 1995; Lewis, 2006). Therefore, Information on correlation coefficient for agro-morphological traits is a matter of great concern for the breeders, as, for instance, it helps indirect selection of positively correlated traits.

Results of the present study showed a negative correlation of thousand kernels weight with days to booting ($r = -0.28$, $p < 0.001$), days to emergence ($r = -0.19$, $p < 0.001$), days to spike emergence ($r = -0.29$, $p < 0.001$), days to tillering ($r = -0.27$,

$p < 0.001$), days to flowering ($r = -0.29$, $p < 0.001$), days to physiological maturity ($r = -0.29$, $p < 0.001$), and degree of seed shriveling ($r = -0.84$, $p < 0.001$) (Table 4.16 and Figure 4.9). Highly significant positive correlations existed between days to spike emergence and days to booting ($r = 0.99$, $p < 0.001$); days to spike emergence and days to flowering ($r = 0.99$, $p < 0.001$); and, between days to booting and days to flowering ($r = 0.98$, $p < 0.001$), all of which are related to duration of the wheat crop. Also, plant height showed positive correlation with 1000-kernels weight, days to spike emergence and days to flowering; and, the comparable results were reported by Zarkti *et al.* (2012) for plant height with days to spike emergence and days to flowering. However, the results of the present study showed positive correlation of thousand kernel weight with seed size, which is consistent with reports of Sabaghnia *et al.* (2014) who suggested that seed size is the main component of the yield and it continues to be the major breeding target. Reynolds *et al.* (2002) suggested that 1000-kernels weight best explained the genotype-by-environment interactions for wheat grain yield. Selection of superior genotypes regarding high grain yield would be as effective as selection for its components, namely NSS and TKW (Peltonen-Sainioa *et al.,* 2007; Sabaghnia *et al.,* 2014). Although there existed a positive relationship between plant height and grain yield, its magnitude was less. Days to anthesis was negatively associated with 1000-kernel weight which is in line with the observation of Mishra *et al.* (2015) that this might be due to the shortened grain filling duration. Number of seeds per spike showed significant ($p<0.01$) positive correlation with spike length ($r=0.3$), number of spikelets per spike ($r=0.33$), number of florets per spikelets ($r=0.51$), and number of seeds per spikelets ($r=0.93$); and, negative association with seed color; suggesting that accessions with long spike length and white seed color possess relatively more number seeds per spike (NSS) compared to those with shorter spike length.

Thus, results on Carl Pearson's coefficients of correlation for agro-morphological traits of studied set of Indian wheat accessions, obtained in the present study, are of practical utility for genetic improvement of wheat through conventional as well as molecular breeding.

5.11.2 Principal component analysis (PCA) of the traits

Number of principal components to be retained for further analysis can be decided based on the criterion that Eigen values for them is more than one (Mishra *et al.,* 2015). Accordingly, four components were retained in the present study. Therefore,

in this analysis, the first, second, third and fourth components, PC1, PC2, PC3 and PC4, had Eigen values of 5.577, 2.821, 1.646 and 1.394, respectively (Table 4.17 and Figure 4.10). The first principal component (PC1) accounted for 37.25% of total variance and had a high contributing factor loading from all the phenological traits, namely, days to emergence, days to tillering, days to booting, days to emergence, days to flowering, and days to physiological maturity; and, low variance for degree of seed shriveling. The second principal component (PC2), accounting for further 18.8% of total variability, indicated primarily the patterns of variation in number of florets per spikelet, number of seeds per spikelet, and number of seeds per spike. The third principal component (PC3) accounting for additional 11% of the total variation, had lower loading values except for thousand kernels weight. PC4, showing another 9.3% of the total variability, had factor loading from spike length, number of spikelets per spike and plant height. Of total variance, 76.35% was accounted for by the first four principal components, PC1 to PC4.

Therefore, except maximum number of seeds per floret, all the quantitative traits, analyzed in the present study, contributed high proportion of variance amounting to 76.35% of the total variance observed in 618 Indian wheat accessions.

5.11.3 Clustering of the 618 wheat accessions based on morphological traits

Two distinct landraces may arise either independently from two different parents *ab initio* or may diverge from each other after an event of hybridization between the parents. To benefit from transgressive segregation, genetic distance between parents is necessary (Joshi *et al.* 2004) that can be estimated by Euclidean distance (Hoque and Rahman, 2006). Though the dendrogram (Figure 4.11), broadly indicates presence of two clusters, based on the following criteria for determining optimal number of clusters, 618 accessions were grouped into five clusters.

5.11.3.1 Criteria for determining number of clusters

Statistically optimal number of clusters in a dendrogram can be determined using criteria described by Millign and Cooper (1985). Accordingly, in the present study optimal number of clusters was determined based on the concurrence of these three criteria: one, cubic clustering criteria (CCC) curve should have the first peak with sudden fall; two, pseudo F curve should be declining slowly; and, three, pseudo T-squared curve should be showing a steady phase. An incision line was drawn such

that all the three criteria are satisfied, and the number of clusters was determined to be five; based on the interception of this incision line on the x-axis (Figure 4.12).

5.11.3.2 GT-biplot and Minimum spanning tree

GT-biplot analysis offers an opportunity for visualization of genetic correlation among genotypes and understanding of genotype-by-trait relationships (Ma *et al.* 2004; Yan and Fregeau-Reid, 2008; Mishra *et al.,* 2015). Pictorial display of clustering of 618 accessions using GT-biplot (Figure 4.13) and minimum spanning tree (Figure 4.14) further confirmed the presence of five clusters.

5.12 Analysis of intra-cluster variability

The information on intra-cluster variability is highly useful in making core collections from the evaluated set of accessions. Five clusters showed intra-cluster variability for different investigated traits, as revealed by their values of coefficient of variations (CV) (Table 4.18). Cluster-I did not vary much for any of the traits; cluster-II varied highly for days to booting (CV, 51.51), and slightly for days to tillering (CV, 24.63) and days to emergence (CV,41.47); cluster-III showed high CVs for glume color (CV, 87.25), days to flowering (CV, 64.54), and days to physiological maturity (CV, 60.46), and was slightly variable for degree of seed shriveling (CV, 23.87); cluster-IV, varied highly for awnedness (CV, 89.38) and moderately for growth habit (CV, 31.74); and, cluster V exhibited high variability for number of florets per spikelet (CV, 53.64) and number of productive florets per spikelet (CV, 51.30). All the five clusters exhibited comparable variability for spike length, number of spikelets per spike, number of seeds per spike, glume hairiness, spike density, thousand kernel weight, seed size, plant height, tillering capacity, and days to emergence.

It was interesting to note that cluster-IV was not at all variable for days to physiological maturity, suggesting that there is no scope for selecting in search of crop duration in this cluster, while such opportunity is offered by cluster-III, for instance. Thus, the results of the present study can be utilized in making core collections from the evaluated set of accessions, listed according to their cluster-membership (Table 4.19).

5.13 Comparison of molecular- and morphological data-based grouping

Since there is no single bioinformatic tool that can handle both molecular and morphological data using the same computational method and algorithm, it is not

prudent to compare every aspect of the study. However, comparison in terms of overall number of major groups revealed that molecular marker-based method grouped the 618 Indian accessions into 4 major groups (Table 4.20), while that derived using morphological data revealed 5 major groups (Table 4.21); indicating that there was no correlation between molecular genotyping- and morphological data-based cluster analysis. The dissimilarity between molecular genotyping- and morphological data-based results, obtained in this study, arises from the fact that while molecular data directly reflected only the allelic variations of markers loci used in the analysis, morphological data indirectly reflected variations in much of genomic regions that correspond to the assessed morphological traits.

The results of the present study are in agreement with previous reports by Salem *et al*. (2008) and Najaphy *et al*. (2012). However, contradictory results were reported by Soriano *et al*. (2016). Salem *et al*. (2008), while studying genetic diversity of 7 wheat varieties using 9 morphological traits and 48 SSR marker loci, found that there was no similarity between the molecular genotyping- and morphological data-based clustering. However, they suggested that the information on grouping of the genotypes based on morphological and SSR marker data is useful for wheat breeders to plan crosses for positive traits and that the wheat microsatellite primers can be used to distinguish the genotypes and estimate their genetic diversity. Similarly, Najaphy *et al*. (2012) combined 20 morphological traits, including agro-morphological and phenological, with DNA-marker (10 ISSR markers) data to study genetic diversity of 30 wheat cultivars and advanced breeding lines, and found that there was no correlation between the two methods. They suggested that information on genetic diversity among 30 wheat genotypes, as revealed by 10 ISSR and 20 morphological traits, would be useful in traditional as well as molecular breeding programs. On the other hand, Soriano *et al*. (2016), upon analyses of genetic structure among 172 durum wheat landraces, utilizing both 14 agro-morphological traits (including yield, physiological and phenological traits) and 44 SSR marker loci data, demonstrated that morphological data-based population structure had similarity with that derived from DNA marker data, indicating that they were connected by geographic origin of the landraces.

While morphological, biochemical and DNA markers can be deployed for diversity studies (Mohammadi and Prasanna, 2003; Sudre *et al*., 2007; Goncalves *et al*., 2009), the DNA markers are advantageous over the other type as they are not affected

by environmental factors and gene interactions. In addition, as DNA markers involve advanced methods of multiplexing and high-resolution genotyping (Binneck *et al*., 2002; Antonio *et al*., 2004; Saker *et al*., 2005; Goncalves *et al*., 2008; Souza *et al,* 2008), they are useful for detailed study of genetic diversity involving a larger number of genotypes or accessions. Thus, molecular and agro-morphological study of genetic diversity, population structure and phylogenetic relationship among Indian wheat landraces augment each other; as they reveal additional information, though not the same. Hence, the dissimilarity between molecular genotyping- and morphological data-based results, obtained in this study, indicates that while molecular data directly reflect only the allelic variations of marker loci used in the analysis, morphological data indirectly reflect variations in much of genomic regions that contribute to the analyzed morphological traits.

Since there was no correlation between molecular and morphological data-based clustering, the results of both the methods are important and will be useful for identification of core set, conservation and utilization of studied set of 618 Indian wheat landraces in wheat improvement programs, to achieve sustainable wheat production, in the face of ever-changing climate and never receding human populations.

CHAPTER 6

SUMMARY & CONCLUSION

6. SUMMARY AND CONCLUSION

Wheat (*Triticum aestivum L.*) is the most important food crop of the world. The success of the crop's cultivation is due to its adaptation to diversified environments, ranging from upland sites in the tropics to the great plains of North America and steppes of Russia and the Ukraine. It is grown in almost all the continents except Antarctica.

Narrow genetic base of the breeding material limits the genetic gain during varietal improvement. Knowledge of genetic diversity among the donor lines of agronomic traits is a crucial factor in adopting the systematic and effective breeding strategies for genetic improvement of wheat**.** Assessment of genetic diversity is essential for germplasm characterization, utilization and conservation. The continued development and utilization of the genetically diverse wheat germplasm resources to constantly expand and enrich the genetic basis of the breeding parents are extremely important for the development and promotion of wheat breeding and production. With an aim of furthering the understanding of extent and pattern of diversity in a set of six hundred and eighteen Indian wheat landraces at genetic (using morphological traits) and genomic (using DNA markers) levels, the current study entitled "**Study of population structure and phylogeny in Indian wheat (*Triticum* spp.) landraces using DNA markers"** was undertaken with the following objectives:

1. Analysis of genetic diversity of Indian wheat landraces using DNA markers.
2. Study of population structure and phylogeny among the Indian wheat landraces.

The outcomes and salient findings of the present study are summarized as follows:

1. Forty-two SSR markers deployed to genotype a panel of 618 Indian wheat landraces revealed a total of 24779 amplicons and 771 size-based polymorphic alleles, with an average of 18.36 per locus. The highest number of 26 alleles per locus was observed in SSR marker xgwm136 located on the short arm of chromosome 1 in subgenome-A. Therefore, a large number of total amplicons as well as polymorphic alleles entails the utility of SSR markers and fluorescent-labelling, deployed in the present study, in assessment of diversity and phylogenetic relationship among Indian wheat landraces.
2. Based on the frequency of their occurrence among 618 Indian wheat accessions, the 771 polymorphic alleles were further partitioned into 139 major- and 632

minor alleles, constituting 17.89% and 82.1%, respectively (Table 4.3). Minor alleles (< 5%) found in this study represent rare alleles. Out of 632 rare alleles, 59 were unique as each of them occurred in only one of the 618 genotypes.

3. In the present study, on an average, each of the 618 genotypes had two null alleles, constituting 4.76% of the total 42 expected amplicons per genotype. On the other hand, each locus had an average of 28 (4.53%) null alleles among a total of 618 expected alleles per locus.
4. Among 42 SSR markers loci surveyed in a set of 618 Indian wheat landrace accessions, 35 loci showed a total of 59 unique alleles with number of alleles per locus ranging from 4 to 0. Indian wheat accession IC-57983 had the highest number of 14 unique alleles. Information on unique alleles entails applications of SSR markers, used in the present study, in DNA-finger printing of wheat landraces.
5. Markers loci in homoeologous group 1 exhibited the highest number of alleles per locus, varying in the range of 15-26, while the lowest number of 12 alleles per locus was detected in the homoeologous group 6. Forty-two SSR loci had polymorphic alleles that were distributed among all the 7 homoeologous groups, ranging from 16 to 23 per group, and they occurred among more than 5% of the 618 Indian wheat landrace accessions. Homoeologous group 2 had the highest number of 23 out of 138 major alleles which was above the group average (19.7). Group 4 had the highest average number of 5.72 major alleles per locus.
6. Each of the seven homoeologous groups had an high degree of polymorphism information content (PIC) ranging from 0.910 in group 1 to 0.873 in group 2, with an average of 0.893 per group. High PIC values distributed among 7 homoeologous groups suggest the usefulness of the 42 SSR markers, deployed in the present study, in deciphering and exploiting the genetic diversity of Indian wheat landraces for genetic improvement of wheat.
7. The comparative analysis of gene diversity among the seven homoeologous groups, each analyzed with 6 marker loci, revealed the existence of a high level of genetic diversity, with an average gene diversity index of 0.901, for all the 7 homoeologous groups. Group 3 showed the highest genetic diversity of 0.912, while group 2 had the lowest (0.883). Further, homoeologous groups 3 (0.912), 6 (0.91) and 1 (0.907) are more diverse than the groups 4 (0.895), 5 (0.893), and 2 ((0.883). In addition, the average genetic diversity indices for homoeologous

groups 3 (0.912), 6 (0.91), 1 (0.907) and 7 (0.905) were higher than the group average of 0.901, whereas those for 4 (0.895), 5 (0.893), and 2 (0.883) were lower than the group average.

8. Upon investigating each of the seven homologous groups with 2 SSR marker loci, one each on their short- and long arms, homologous group 1A showed the highest number of 28 polymorphic alleles. Chromosome 1A was the richest in number of polymorphic alleles (25.5). Chromosomes 3A and 5A had the highest number of 6 unique alleles each, while 2D and 6D had none. While the highest number of 43 rare alleles were detected in chromosome 1A, chromosome 6D had 21. While 5B had no null-allele, 1B had 170. The remaining 19 chromosomes harboured at least one null allele each.
9. Although each of the homologous chromosomes 2A, 3A and 6A had the highest average polymorphism information content (PIC, 0.920) for two markers, Chromosome 6A alone exhibited the highest gene diversity (GDI, 0.925), followed by 2A and 3A with GDI of 0.920 each. On the other hand, chromosome 6A had a greater number of null-alleles (101) than 2A and 3A, which showed 25 each suggesting that the null-alleles also contributed to the decipherment of genetic diversity.
10. Most of the homologous groups had the genetic diversity above the group average (GDI, 0.9), while it was below the average for six chromosomes: 2B (GDI, 0.875), 2D (GDI, 0.855), 4B (GDI, 0.895), 4D (GDI, 0.89), 5B (GDI, 0.875) and 6D (GDI, 0.895).
11. In the subgenome A, there were a total of 320 polymorphic-, 54.62 major-, 265.38 minor-, 32 unique- and 532 null-alleles, that constituted 41.5 %, 39.72 %, 41.89 %, 54.23 % and 45.43 % of alleles recorded in the total genome (ABD), respectively. The average PIC value of 14 SSR marker loci located in subgenome-A was 0.907, while the average genetic diversity index (GDI) was 0.912. This suggests that 14 markers explained the 91.2% of the genetic diversity of subgenome-A corresponding to 14 markers loci among 618 Indian wheat landraces.
12. Subgenome-B had a total of 250 polymorphic-, 46.29 major-, 203.71 minor-, 18 unique- and 365 null-alleles, that constituted 32.4 %, 33.67 %, 32.15 %, 30.5 % and 31.67 % of alleles detected in the total genome (ABD), respectively. The average PIC value for 14 SSR markers loci located in subgenome-B was 0.887,

while the average genetic diversity index (GDI) was 0.895. Therefore, 89.5% of genetic diversity of subgenome-B, with respect to the 14 loci, was revealed by 14 SSR markers amplified among 618 Indian wheat landrace accessions.

13. In total, subgenome-D had 201 polymorphic-, 36.59 major-, 164.41 minor-, 9 unique- and 294 null-alleles, that comprised 26.07 %, 26.61%, 25.95%, 15.25% and 22.9% of those detected in the whole genome (ABD), respectively. Both average PIC value as well as average GDI for 14 SSR markers loci across the subgenome-D was 0.89. Thus, 14 SSR markers explained 89% of the genetic diversity of subgenome-D, with respect to 14 markers loci, in a panel of 618 Indian wheats.
14. The subgenome-A had the highest genetic diversity index (GDI) of 0.912 ranging from 0.88-0.93, followed by subgenomes B (0.896) and D (0.890), where GDI varied from 0.85-0.92 and 0.84-0.92, respectively.
15. A panel of 618 Indian wheat accessions were grouped into of 14 clusters (or groups), at genetic dissimilarity coefficient of 0.7. Clusters were coded as A_1, A_2, A_3, C_1, D_1, E_1, F_1, G_1, H_1, I_1, J_1, K_1, L_1 and L_2, and had membership strengths of 20, 1, 13, 24, 1, 1, 94, 169, 68, 13, 80, 55, 23 and 50, respectively. Among them, three clusters A_2, D_1 and E_1 had a single member each, *viz*., IC-118758 of unknown origin, IC-30276 belonging to state of Jammu and Kashmir, and IC-212185 from state of Rajasthan, respectively.
16. Three accessions (IC-79095, IC-78999 and IC-82342) belonging to Himachal Pradesh, four accessions (IC82180, IC-104537, IC-82127 and IC-82161) from Rajasthan, and two accessions (IC-118729 and IC-73215) of Maharashtra were grouped into one cluster along with one accession (IC-35069) belonging to Karnataka, and two accessions (IC-78832 and IC-78707) from Uttar Pradesh.
17. Cluster L_1 had a total of 23 accessions of which the most (20) of them belonged to state of Uttar Pradesh, two, IC-79066 and IC-79079, were of Himachal Pradesh, and one, IC-78941, was of unknown origin.
18. Indian landrace accessions; IC-118758 of unknown origin, IC-30276 belonging to state of Jammu and Kashmir, and IC-212185 from state of Rajasthan; were grouped into separate distinct clusters, suggesting that they can be utilized in broadening the genetic base of Indian wheat landrace collections for their exploitation in wheat breeding programs in the face of ever-changing climate and never receding human population.

19. Analysis of molecular variance (AMOVA) partitioned the total diversity into variability within each population and among populations. The molecular genetic variance within populations was 90%, whereas it was 10% among populations and the partitioning pattern was uniform across all the three subgenomes (A, B and D) as well as in the whole genome (ABD).
20. The Model based population structure analysis revealed the presence of four subpopulations at threshold value of membership-probability 0.80. This is consistent with the results of dendrogram analysis where 618 accessions were clustered into four major groups or clads. Three-dimensional scaling by principal coordinate analysis (PCoA) also separated a panel of 618 Indian wheat landrace accessions into 4 major groups, the first, second and third axes explained 4.14 %, 3.34 % and 2.77%, of the total genetic variation, respectively.
21. The Phylogenetic analysis showed that a set of 618 Indian wheat accessions were grouped into 4 major phylogenetic clads: Clad-A1, Clad -A2, Clad -A3 and Clad-A4, each diverging from an unknown common ancestor.
22. The accessions from Uttar Pradesh, Rajasthan and Karnataka, for instance, shared common ancestors and were phylogenetically more related among themselves. Similarly, accessions from Himachal Pradesh were more closely related to those from Uttar Pradesh and distantly related to accessions from Rajasthan as well as those from Madhya Pradesh.
23. A landrace with accession number IC-118758 of unknown origin grouped singly in a phylogenetic Clad A2, suggesting that it was genetically more divergent from every member of the most diverse and the largest Clad 4, having 11 subgroups and a total of 579 accessions. Thus, IC-118758 can be utilized to hybridize with every member of Clad A2 for broadening of genetic base of Indian wheat accessions for utilization in wheat improvement programs.
24. High coefficient of variation (CV) was recorded for 1000-kernel weight (CV, 42.78%), tillering capacity (CV, 55.69%) and number of seeds per spike (CV, 42.78%).
25. Among 618 Indian wheat accessions, majority were of spreading type with medium plan height (90 to 105 cm), intermediate tillering capacity (5 to 10 cm), medium spike length (5 to 10 cm), having 15 to 20 spikelets per spike, single kernel per floret, 25 to 50 seeds per spike, awned, glume color purple to black, pubescent glume, intermediate type of spike density, red seed color,

intermediate seed size, intermediate seed shriveling, early emergence (12 to 15 days), days to tillering medium (20 to 25 days), days to booting late (75 to 90 days), days to spike emergence early (60 to 75 days), very early flowering (≤ 70 days), early physiological maturity (90 to 105 days), and low thousand kernel weight (20 to 30 g).

26. Two accessions, IC_534777 and IC_534787, recorded the highest 1000-kernel weights of more than 50 gram each.
27. Estimated phenotypic diversity (H') for individual traits ranged from 0.93, for glume hairiness, to 0.49, for maximum number of seeds per spike. Analyzed accessions showed high phenotypic diversity (H'>0.60), except for NSS (0.59) and NSSL (0.49), indicating high-levels of polymorphism among studied accessions for evaluated traits.
28. Results of the present study showed a negative correlation of thousand kernel weight with days to booting ($r = -0.28$, $p < 0.001$), days to emergence ($r = -0.19$, $p < 0.001$), days to spike emergence ($r = -0.29$, $p < 0.001$), days to tillering ($r = -0.27$, $p < 0.001$), days to flowering ($r = -0.29$, $p < 0.001$), days to physiological maturity ($r = -0.29$, $p < 0.001$) and degree of seed shriveling ($r = -0.84$, $p < 0.001$). Plant height showed positive correlation with 1000-kernel weight, days to spike emergence, and days to flowering.
29. Four components, PC1, PC2, PC3 and PC4 accounted for 37.25%, 18.8%, 11% and 9.3% of total variance, respectively. Therefore, except maximum number of seeds per floret, all the quantitative traits, analyzed in the present study, contributed to the high proportion of variance amounting to 76.35% of the total variance observed in 618 Indian wheat accessions.
30. Morphological and phenological traits-based cluster analysis grouped 618 Indian wheat accessions into five major groups. Pictorial display of clustering of 618 accessions using GT-biplot and minimum spanning tree further confirmed the presence of five clusters.
31. Five clusters showed intra-cluster variability for different investigated agro-morphological traits, as revealed by their values of coefficient of variation (CV). Cluster-I did not vary much for any of the studied traits; cluster-II varied highly for days to booting (CV, 51.51), and slightly for days to tillering (CV, 24.63) and days to emergence (CV, 41.47); cluster-III showed high CV for glume color (87.25), days to flowering (64.54), and days to physiological maturity (60.46),

whereas there was a slight variability for degree of seed shriveling (CV, 23.87); cluster-IV varied highly for awnness (CV, 89.38) and moderately for growth habit (CV, 31.74); and, cluster V exhibited high variability for number of florets per spikelet (CV, 53.64) and number of productive florets per spikelet (CV, 51.30). All the five clusters exhibited comparable variability for spike length, number of spikelets per spike, number of seeds per spike, glume hairiness, spike density, thousand kernel weight, seed size, plant height, tillering capacity, and days to emergence.

32. There was no correlation between molecular and morphological data-based clustering. Therefore, results of both the methods augment each other and can be used for identification of core set, conservation and utilization of studied set of 618 Indian wheat landraces in wheat improvement programs.

CONCLUSION

In conclusion, paucity of knowledge of available genetic variation within Indian wheat accessions limits its effective utilization; and, the earlier studies with limited number of genotypes are inadequate to depict the complete picture of wheat genetic resources of India. Deployment of high-resolution genotyping method, involving fluorescent-labelled SSR markers and capillary gel electrophoresis, to investigate a set of 618 Indian wheat accessions, revealed a high number of polymorphic alleles, high polymorphism information content (PIC) values, and high gene diversity indices across all the three homoeologous groups, all the 7 homologous groups, and in all the three subgenomes. Based on SSR genotyping data, there were 4 major groups (or clades) diversging from a common ancestor. Molecular variance was more within populations than among populations, suggesting that there was selective adaptation of the landraces towards the local growth conditions specific to the geopolitical state of their origin. In addition, estimates of coefficients of variation and phenotypic diversity index revealed a high degree of polymorphism for studied productivity-related traits and wide spectrum of diversity in a set of 618 Indian wheat accessions. Further, PCA cluster analysis and GT-biplot grouped the accessions into five main clusters. Since there was no correlation between molecular and morphological data-based clustering, the results of both the methods are important and will be useful for identification of core set, conservation and utilization of studied set of 618 Indian wheat landraces in wheat genetic improvement programs.

FUTURE SCOPE

Even though the findings of the present study are promising, suggesting the utility of Indian wheat landrace accessions and the markers, used in the present study, in molecular and traditional wheat breeding; further understanding of diversity; population structure and phylogeny of Indian wheat landraces with detailed passport data, by using large number of markers and genotypes and involving advanced system of markers such as SNP and DArT; is required to mine allele-rich Indian wheat landrace accessions for useful traits of agronomic importance that my be utilized in the breeding programs boosting the sustainable wheat production in future to meet the growing demand, in the face ever changing climate, on one hand, and never receding human populations, on the other hand.

ABSTRACT (ENGLISH)

Study of population structure and phylogeny in Indian wheat (*Triticum* spp.) landraces using DNA markers

ABSTRACT

Since genetic diversity among genotypes offers prospects for improving the plant characteristics, Indian wheat genetic diversity study is a key to wheat improvement programs in India. To assess extent and pattern of genetic diversity to further our understanding of population structure and phylogeny in Indian wheat landraces, 42 fluorescence-labelled SSR markers loci, evenly distributed throughout the hexaploid wheat genome, were amplified among 618 accessions, and an high-resolution genotyping was performed using capillary gel electrophoresis. It was found that there was a total of 771 polymorphic alleles ranging from 12 to 26 alleles per locus and a mean gene richness of 18.36. Substantial variations in allelic richness were detected by 42 SSR markers in all the 21 homologous chromosomes, 7 homoeologous groups, and the three subgenomes (A, B, and D) among the genotypes. The average genetic diversity index was 0.901 with three genomes having values ranging from 0.84 to 0.93. Analysis of variance (ANOVA) revealed that subgenomes A and D had the highest (0.912) and the lowest (0.890) genetic diversity, respectively. There were four major groups based on STRUCTURE, UPGMA cluster and principal coordinate (PCoA) analyses. STRUCTURE analysis showed that the model-based population was not correlated with the ecological origin of accessions; yet, an high amount of diversity was present as indicated by ANOVA. Phylogenetic analysis revealed that the studied set of 618 accessions belonged to four clads that diverged from a common unknown ancestor. Analysis of molecular variance (AMOVA) partitioned the total genetic variance, estimated in this study, into 90% within groups and 10% among groups, both at individual subgenome and whole genome levels, suggesting that there was selective adaptation of the landraces towards the local growth conditions specific to the geopolitical state of their origin. Agro-morphological and phenological trait-based UPGMA tree revealed the existence of five clusters, confirmed by principal component analysis (PCA). However, there was no correlation between molecular and morphological data-based clustering, suggesting that the results of both the methods are important and will be useful

for identification of core set, conservation and utilization of Indian wheat germplasm in wheat breeding programs.

In conclusion, evidences from this study suggest that Indian wheat landrace accessions have high genetic richness and diversity and, therefore, they can be useful sources of genes for various traits of agronomic importance that can be used by breeders in different wheat hybridization programs. In addition, genetically diverse accessions identified in this study can be used as parents for increasing the usefulness of the collections by broadening the genetic base. Findings of the present study also indicate that fluorescent labelled microsatellite markers permit high-resolution genotyping by generating a large set of genotypic data that can be used for assessing genetic diversity and population structure, facilitating the discovery of all the alleles present in the studied set of genotypes.

Finally, results of this study would be useful in traditional and molecular breeding programs involving Indian wheat landraces, analysed in this study, so that sustainable wheat production can be achieved in the face of ever changing climate and never receding human population.

Key words: *Triticum*, Indian wheat, Landrace accessions, Microsatellite (SSR) markers, ANOVA, AMOVA, STRUCTURE, Cluster, PCoA, PCA, Diversity, Population structure, Phylogeny, Molecular data, Agro-morphological, Wheat breeding

---

ABSTRACT (HINDI)

डी एन ए चिह्नकों का उपयोग कर भारतीय गेहूँ (ट्रिटीकम प्रजातियों) भूप्रजातियों में पापुलेक्ष्न संरचना एवं जातिवृत्त का अध्ययन

सार

भारत में गेहूँ सुधार कार्यक्रमों के लिए, भारतीय गेहूँ में आनुवंशिक विविधता का अध्ययन करना अत्यावश्यक है अतः जीनप्ररूपों के बीच आनुवंशिक विविधता, पादप गुणों में सुधार करने की संभावना उपलब्ध कराती है, । भारतीय गेहूँ भूप्रजातियों में पापुलेक्ष्न संरचना एवं जातिवृत्त संरचना को और अधिक समझने के लिए, हेक्साप्लॉयड गेहूँ जीनोम में समान रूप से वितरित 42 प्रतिदीप्तिशील–लेबल्ड एस एस आर चिह्नक एलओसीआइ का 618 वंशक्रमों में प्रवर्धन किया गया और कैपिलरी जैव इलेक्ट्रोफोरेसिस का उपयोग कर उच्च–वियोजन जीनप्ररूपण किया गया। यह पाया गया कि 18.36 की एक औसत आनुवंशिक समृद्धता सहित प्रतिलोकस 12 से 26 युग्मविकल्पियों की सीमा में कुल 771 बहुरूपी युग्मविकल्पी थे। जीनप्ररूपों के बीच तीन उपजीनोम (ए, बी एवं डी), 7 समजात समूह तथा सभी 21 समजात गुणसूत्रों में एस एस आर चिह्नकों द्वारा युग्मविकल्पी संबंधी समृद्धता में काफी अधिक परिवर्तनों की पहचान की गई। 0.84 से 0.93 की सीमा में मान रखने वाले तीन जीनोम के साथ औसत आनुवंशिक विविधता घातांक 0.901 था। प्रसरण विश्लेषण (एएनओवीए) ने दर्शाया कि ए एवं डी जीनोम में क्रमशः अधिकतम (0.912) एवं न्यूनतम (0.890) आनुवंशिक विविधता थी। स्ट्रक्चर, यूपीजी एमए समुच्चय एवं प्रिंसिपल कोऑर्डिनेर (पी सी ओ ए) विश्लेषणों के आधार पर, चार उपसमूह विद्यमान पाए गए। स्ट्रक्चर विश्लेषण ने दर्शाया कि मॉडल आधारित आबादी का वंशक्रमों की पारिस्थितिक उत्पत्ति के साथ संहसंबंध नहीं था; हालाँकि बहुत अधिक जैवविविधता विद्यमान थी, जैसाकि एएनओवीए ने दर्शाया। जातिवृत्तीय विश्लेषण ने दर्शाया कि अध्ययन किया गया 618 वंशक्रमों का सेट 4 क्लेड्स के अंतर्गत आता है जिनका एक ही अज्ञात पूर्वज से विपथन हुआ। आव्विक प्रसरण विश्लेषण (एएमओवीए) ने इस अध्ययन में आकलित कुल आनुवंशिक प्रसरण को उपजीनोम एवं सम्पूर्ण जीनोम, दोनों स्तरों पर, समूहों के भीतर 90% और समूहों के बीच 10% प्रसरण में विभक्त किया जो यह दर्शाता है कि भूप्रजातियों का वृद्धि परिस्थितियों की ओर एक चयनित अनुकूलन था। कृषि–आकारिकीय एवं लक्षण विज्ञान संबंधी गुण–आधारित यूजीजीएमए वृक्ष ने पांच समुच्च्यों की उपस्थिति दर्शायी जिनकी प्रमुख घटक विश्लेषण (पीसीए) द्वारा पुष्टि की गई। तथापि, आण्विक एवं आकारिकीय डाटा–आधारित समुच्चयन के बीच कोई सहसंबंध नहीं था जिससे सुझाव मिलता है कि दोनों विधियों के परिणाम महत्वपूर्ण हैं और वे गेहूँ प्रजनन कार्यक्रमों में कोर सैट की पहचान करने, संरक्षण एवं उपयोग में लाभप्रद रूप से सहायक होंगे।

निष्कर्ष के रूप में कहा जा सकता है कि इस अध्ययन के परिणाम दर्शाते हैं कि भारतीय भूप्रजाति वंशक्रमों में उच्च आनुवंशिक समृद्धता एवं विविधता विद्यमान है और इसलिए, वे सस्यविज्ञान की दृष्टि से

महत्वपूर्ण कई जीनों के उपयोगी स्रोत हो सकते हैं और प्रजनकों द्वारा अपने विभिन्न गेहूँ–संकरण कार्यक्रमों में उनका उपयोग किया जा सकता है। भविष्य में, कोर संग्रह की पहचान करने और भारतीय गेहूँ भूप्रजाति संग्रहों में साहचर्य अध्ययनों हेतु, इस अध्ययन के परिणाम महत्वपूर्ण सूचना उपलब्ध कराते हैं। इसके अतिरिक्त, इस अध्ययन में पहचाने गए आनुवंशिक विविधता युक्त वंशक्रमों का उपयोग जनकों के रूप में किया जा सकता है ताकि आनुवंशिक आधार को व्यापार बनाकर संग्रहों की उपयोगिता में बढ़ोतरी जी जा सके। प्रस्तुत अध्ययन यह भी दर्शाता है कि प्रतिदाप्तिशील लेबल्ड सूक्ष्मअनुषंगी चिह्नक उच्च–वियोजन जीनप्ररूपण द्वारा जीनप्ररूपों का एक बड़ा सैट बना सकते हैं जिनका भारतीय गेहूँ की भूप्रजातियों में आनुवंशिक विविधता एवं पापुलेक्ष्न संरचना के आकलन में उपयोग हो सकता है और इस प्रकार से ये अध्ययन किए गए जीनप्ररूपों में उपस्थिति सभी युग्मविकल्पियों की खोज में सहायक होंगे।

अंत में, आण्विक चिह्नकों (42 एस एस आर लोकाई) एवं 23 आकारिकीय गुणों का उपयोग कर 618 भारतीय गेहूँ भूप्रजाति वंशक्रमों के बीच आनुवंशिक विविधता, पापुलेक्ष्न संरचना एवं जातिवृत्तीय विश्लेषणों के संबंध में इस अध्ययन के परिणाम, इस अध्ययन में उपयोग किए गए भारतीय गेहूँ भूप्रजातियों के समावेश वाले पारम्परिक एवं आण्विक, दोनों प्रकार के प्रजनन कार्यक्रमों में उपयोगी होंगे ताकि दिन–प्रतिदिन बढ़ती मानव जनसंख्या के लिए और प्रतिकूल जैव एवं अजैव कारकों का सामना करने के लिए, टिकाऊ गेहूँ उत्पादन प्राप्त किया जा सके।

मुख्य शब्दः ट्रिटीकम, भारतीय गेहूँ भूप्रजातियां, वंशक्रम, सूक्ष्म अनुषंगी (एस एस आर), चिह्नक एएनओवीए, एएमओवीए, संरचना, समुच्चय, पीसीओए, पीसीए विविधता, पापुलेक्ष्न संरचना, जातिवृत्त, आण्विक, कृषि–आकारिकीय, गेहूँ प्रजनन

...

BIBLIOGRAPHY

BIBLIOGRAPHY

Abouzied, H.M., Samah, M.M., Eldemery and Abdellatif, K.F. (2013). SSR- Based genetic diversity assessment in tetraploid and hexaploid wheat populations. *British Biotechnology Journal*, **3**(3): 390-404.

Adams, M.D., Kelley, J.M., Gocayne, J.D., Dubrick, M., Polymeropoulos, M.H., Xiao, H., *et al*. (1991). Complementary DNA sequencing: Expressed sequence tags and human genome project. *Science*, **210**: 1651-1656.

Adams, M.W. (1997). An estimation of homogeneity in crop plants, with special reference of genetic variability in the dry bean, *Phaseolus vulgar*. *Euphytica*, **6**: 665-679.

Adhikari, T.B., Tang, X., Cavelettoo, J.R., Hu, X., Buechley, G., Ohm, H.W., *et al*. (2004). Molecular mapping of Stb1, a potentially durable gene for resistance to septoria tritici blotch in wheat. *Theoretical and Applied Genetics*, **109**(5): 944-53.

Aghaee, M., Mohammadi, R. and Nabovati, S. (2010). Agro-morphological characterization of durum wheat accessions using pattern analysis. *Australian Journal of Crop Sciences*, **4**: 505-514.

Akfirat, F.S. and Uncuoglu, A.A. (2013). Genetic diversity of winter wheat (*Triticum aestivum* L.) revealed by SSR markers. *Biochemical Genetics*, **51**: 223-229.

Akkaya, M.S., Bhagwat, A.A. and Cregan, P.B. (1992). Length polymorphism of simple sequence repeats DNA in soybean. *Genetics*, **132**: 1131-1139.

Akram, A., Khan, A.I., Awan, S.F., Rehman, A.U., Ahmad, F., Malik, A., *et al*. (2012). Genetic Diversity in Indian sub-continental landrace cultivars of the genus *Triticum* L. *African Journal of Biotechnology*, **11**(44): 10170-10175.

Ali, Y., Atta, B.M., Akhter, J., Monneveux, P. and Lateef, Z. (2008). Genetic variability, association and diversity studies in wheat (*Triticum aestivum* L.) germplasm. *Pakistan Journal of Botany*, **40**: 2087-2097.

Allard, R.W. (1996). Principles of Plant Breeding, 2nd edition. John Wiley and Sons, New York. **pp**. 254.

Alsaleha, A., Shehzad, F., Balochb, M.N. and Hakan, O. (2016). Phenotypic and genotypic intra-diversity among Anatolian durum wheat "Kunduru" landraces. *Biochemical Systematics and Ecology*, **65**: 9-16.

Alsammanl, A. M., Adawyl, S. S., Ibrahim1, S. D., Hussein, B. A. and Hussein, E. H. A. (2017). Selective amplification of start codon polymorphic Loci (SASPL). *Plant Omics Journal*, **10**(02): 64-77.

Amin, M.R., Barma, N. and Razzaque, M.S. (1992). Variability, heritability, genetic advance and correlation study in some quantitative characters in durum wheat. *Rachis*, **11**(1-2): 30-32.

Ammiraju, J.S.S., Dholakia, B.B., Jawdekar, G., Santra, D.K., Gupta, V.S., Roder, M.S., *et al*. (2002). Inheritance and identification of DNA markers associated with yellow berry tolerance in wheat (*Triticum aestivum* L.). *Euphytica*, **123**: 229-233.

Andersen, J.R. and Liberstedt, T. (2003). Functional markers in plants. *Trends in Plant Science*, **8**: 554-560.

Andersen, P. and Cohen, M.J. (2001). Rich and poor country perspectives on biotechnology. In: The future of food: Biotechnology markets and policies in an international setting. Pardey, P.G. (ed.). International Food Policy Research Institute, Washington, D.C**. pp**. 17-47

Anderson, G.A., Ogihara, Y., Sorrels, M.E. and Tanksley, S.D. (1992). Development of chromosome arm map for wheat based on RFLP markers. *Theoretical and Applied Genetics*, **83**: 1035-1043.

Anderson, J.A., Stack, R.W., Liu, S., Waldron, B.L., Fjeld, A.D., Coyne, C., *et al*. (2001). DNA markers for Fusarium head blight resistance QTLs in two wheat populations. *Theoretical and Applied Genetics*, **102**: 1164–1168.

Anderson, J.A., Churchill, G.A., Autrique, J.E., Tanksley, S.D. and Sorrells, M.E. (1993). Optimizing parental selection for genetic linkage maps. *Genome*, **36**: 181-186.

Angers, B. and Bernatchez, L. (1997). Complex evolution of a salmonid microsatellite locus and its consequences in inferring allelic divergence from size information. *Molecular Biology and Evolution*, **7**(14): 230-238.

Antonio, A.F., Garcia, A.A.F., Benchimol, L.L., Barbosa, A.M.M., Geraldi, I.O., Souza-Jr., *et al*. (2004). Comparison of RAPD, RFLP, AFLP and SSR markers for diversity studies in tropical maize inbred lines. *Genetics and Molecular Biology*, **27**: 579-588.

Arcot, S.S., Wang, Z., Weber, J.L., Deininger, L. and Batzer, M.A. (1995). Alu repeats: A source for the genesis of primate micro-satellites. *Genomics*, **29**: 136-144.

Armour, J.A.L., Alegre, S.A., Miles, S., Williams, L.J. and Badge, R.M. (1999). Minisatellites and mutation processes in tandemly repetitive DNA. In: Goldstein, D.B. and Schlotterer, C. (eds.) Microsatellites: evolution and applications. Oxford University Press, Oxford. **pp**. 24-33.

Arora, A., Kundu, S., Dilbaghi, N., Sharma, I. and Tiwari, R. (2014). Population structure and genetic diversity among Indian wheat varieties using microsatellite (SSR) markers. *Australian Journal of Crop Sciences*, **8**(9): 1281-1289.

Aruna, R. (2003). Morphological and molecular diversity analysis, and genetic enhancement of pigeonpea (*Cajanus cajan* L) for pod borer resistance through wide hybridization. Ph. D thesis (unpublished) submitted to Osmania University, Hyderabad, India.

Ayala, M., Guzmán, C., Pena, R.J. and Alvarez, J.B. (2015). Diversity of phenotypic (plant and grain morphological) and genotypic (glutenin alleles in Glu-1 and Glu-3 loci) traits of wheat landraces (*Triticum aestivum*) from Andalusia (Southern Spain). *Genetic Resources and Crop Evolution*, **63**(3): 465-475.

Badakhshan, H., Mohammadi, S.A., Ahari Zad, S., Moghaddam, M., Jalal-Kamali, M.R. and Khodarahmi, M. (2008). Quantitative trait loci in bread wheat (*Triticum aestivum* L.) associated with resistance to stripe rust. *Biotechnology and Biotechnological Equipment*, **22**(4): 901-906.

Bai, D.P., Scoles, G.J. and Knott, D.R. (1999). Transfer of leaf rust and stem rust resistance genes from (*Triticum Piaristatum)* to durum and bread wheat and their molecular cytogenetic localization. *Genome*, **37**(3): 410-418.

Baisakh, B. and Nayak, S.K. (1991). Genetic variability and correlation studies of yield contributing characters in wheat. *Environment and Ecology*, **9**(3): 624-696.

Baranger, A., Aubert, G., Arnau, G., Laine, A.L., Deniot, G., Potier, J., *et al*. (2004). Genetic diversity within *Pisum sativum* using protein- and PCR-based markers. *Theoretical and Applied Genetics*, **108**: 1309–1321.

Barkley, N.A., Dean, R.E., Pittman, R.N., Wang, M.L., Holbrook, C.C. and Pederson, G.A. (2007). Genetic diversity of cultivated and wild type peanuts evaluated with M13-tailed SSR markers and sequencing. *Genetic Research*, **89**: 93-106.

Barkley, N.A., Roose, M.L., Krueger, R.R. and Federici, C.T. (2006). Assessing genetic diversity and population structure in a citrus germplasm collection utilizing simple sequence repeat markers (SSRs). *Theoretical and Applied Genetics*, **112**: 1519–1531.

Barkley, N.A., Newman, M.L., Wang, M.L., Hotchkiss, M.W. and Pederson, G.A. (2005). Assessment of the genetic diversity and phylogenetic relationships of a temperate bamboo collection by using transferred EST-SSR markers. *Genome*, **48**. 731-7.

Barrett, B.A., Kidwell, K.K. and Fox, P.N. (1998). Comparison of AFLP and pedigree-based genetic diversity assessment methods using wheat cultivars from the pacific northwest. *Crop Sciences*, **38**: 1271-1278.

Bartual, R., Carbonell, E.A. and Green, D.E. (1985). Multivariate analysis of a collection of soyabean cultivars for southeastern Spain. *Euphytica*, **34**: 113-123.

Beckmann, J. and Soller, M. (1983). Restriction fragment length polymorphisms in genetic improvement of agricultural species. *Euphytica*, **35**: 111-124.

Beckmann, J.S. and Soller, M. (1990). Towards a unified approach to genetic mapping of eukaryotes based on sequence tagged microsatellite sites. *Biotechnology*, **8**:930-932.

Beckmann, J.S. and Weber, J.L. (1992). Survey of human and rat microsatellites. *Genomics*, **12**: 627-631.

Belderok, B., Mesdag, H. and Dingena, A. (2000). Bread-Making Quality of Wheat. Parts 1 and 2. Springer-Verlag, New York. **pp**. 126-138

Bennett, E. (1970). Adaptation in wild and cultivated plant populations. In: Genetic resources in plants: Their exploitation and conservation. Frankel O.H. and Bennett, E. (eds.). Blackwell Scientific Publishers, Oxford, UK., **pp**. 115-129.

Bennett, M.D. and Smith, J.B. (1976). Nuclear DNA amounts in angiosperms. *Philosophical Transaction of Royal Society, London*, **274**: 227-274.

Bevan, M.W., Uauy, C., Wulff, B.B.H., Zhou, J., Krasileva, K., Matthew, D., *et al*. (2017). Genomic innovation for crop improvement. *Nature*, **543**: 346-354.

Binneck, E., Nedel, J.L. and Dellagostin, O.A. (2002). RAPD analysis on cultivar identification: a useful methodology. *Revista Brasileira de Ginecologia e Obstetrícia*, **24**: 183-196.

Boder, P., Deak, T., Bacso, R., Velich, I., Bisztray, G.D., Fascar, G., *et al*. (2006). Morphological and genetic investigation of medieval grape seeds. *Acta Horticulture (ISHS)*, **pp**. 713-718.

Bohn. M., Utz, H.F. and Melchinger, A.E. (1999). Genetic similarities between winter wheat cultivars determined on the basis of RFLPs, AFLPs, and SSRs and their use for predicting progeny variance. *Crop Science*, **39**: 228-237.

Bonierbale, M., Beebe, S., Tohme, J. and Jones, P. (1997). Molecular genetic techniques in relation to sampling strategies and the development of core collection. In: Molicular Genetic Techniques for Plant Genetic Resources. Ayad, H .G., Hodgkin, T., Jaradat, A. and Rao, V. R. (Eds.). Report of an IPGRl workshop, 9-11 October 1995, Rome, Italy, International Plant Genetic Resources Institute, Rome, Italy. **pp**. 98-102.

Borlaug, N.E. *(*1988). Challenges for global food and fiber production. *Journal of Research on Swedish Academy of Agriculture and Forestry*, **21**(Suppl)*:* 15–55.

Borner, A., Schumann, E., Furste, A., Coster, H., Leithold, B., Roder, S., *et al*. (2002). Mapping of quantitative trait loci determining agronomic important characters

in hexaploid wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics*, **105**(6-7): 921-936.

Botstein, D., White, R.L., Skolnick, M. and Davis, R.W. (1980). Construction of genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics*, **32**: 314-331.

Bowcock, C. (1994). Association between molecular markers and quantitative traits in oat germplasm pool: can we infer linkage. *Journal of Agriculture and Genetics*, **3**: 11-18.

Bowden, W.M. (1959). The taxonomy and nomenclature of the wheats, barleys and tyres and their wild relatives. *Canadian Journal of Biology*, **37**: 967-684.

Branlard, G., Autran, J.C. and Monneveus, P. (1989). High molecular weight glutenin subunit in durum wheat (*T. durum*). *Theoretical and Applied Genetics*, **78**: 353-358.

Breen, J., Wicker, T., Kong, X., Zhang, J., Ma, W., Paux, E., *et al*. (2010). Highly conserved island of three genes on chromosome 3B of hexaploid wheat: diverse gene function and genomic structure maintained in a tightly linked block. *BMC Plant Biology*, **10**(98): 1471-2229.

Breiman, A. and Graur, D. (1995). Wheat Evaluation. *Israel Journal of Plant Science*, **43**: 58-95.

Breseghello, F. and Sorrells, M.E. (2006). Association analysis as a strategy for improvement of quantitative traits in plants. *Crop Science*, **46**: 1323-1330.

Bretting, P.K. and Widrlechner, M.P. (1995). Genetic markers and horticultural germplasm management. *Horticultural Sciences*, **30**: 1349-1356.

Brohele, J. and Ellegren, H. (1999). Microsatellite evolution: polarity of substitution within repeats and neutrality of flanking sequences. *Proceedings of Royal Society of London Biological Sciences*, **266**: 825-833.

Brookfield, J.F.Y. (1992). DNA Fingerprint in clonal organisms. *Molecular Ecology*, **1**: 21-26.

Bryan, G.J., Collins, A.J., Stephenson, P., Orry, A., Smith, J.B. and Gale, M.D. (1997). Isolation and characterization of microsatellite from hexaploid bread wheat. *Theoretical and Applied Genetics*, **94**: 557-563.

Buerstmayr, H., Lemmens, M., Hartl, L., Doldi, L., Steiner, B., Stierschneider, M., *et al*. (2002). Molecular mapping of QTLs for *Fusarium* head blight resistance in spring wheat. I. Resistance to fungal spread (Type II resistance). *Theoretical and Applied Genetics*, **104**: 84-91.

Bustos, A., Soler, C. and Jouve, N. (1999). Analysis by PCR-based markers using designed primers to study relationships between species of *Hordeum* (Poaceae). *Genome*, **42**: 129-38.

Caetano-Anolles, G., Bassam, B.J. and Gresshoff, P.M. (1991). DNA amplification fingerprinting using very short arbitrary oligonucleotide primers. *Biotechnology*, **9**: 553-557.

Casadesus, J., Kaya, Y., Bort, J., Nachit, M.M., Araus, J.L., Amor, S., *et al*. (2007). Using vegetation indices derived from conventional digital cameras as selection criteria for wheat breeding in water-limited environments. *Annals of Applied Biology*, **150**: 227-236.

Chakraborty, R., Kimmel, M., Stivers, D.N., Davison, L.J. and Deka, R. (1997). Relative mutation rates at di-, tri-, and tetranucleotide microsatellite loci. *Proceedings of National Academy of Sciences (USA)*, **94**: 1041-1046.

Chambers, G.K. and Mac-Avoy, E.S. (2000). Microsatellites: consensus and controversy. *Comparative Biochemistry and Physiology*, **126**: 455-476.

Chao, S., Sharp, P.J., Worland, A.J., Warham, E.J., Koebner, R.M.D. and Gale, M.D. (1989). RFLP-based genetic maps of wheat homologous group 7 chromosomes. *Theoretical and Applied Genetics*, **78**: 495-504.

Chao, S., Zhang, W., Dubcovsky, J. and Sorrells, M. (2007). Evaluation of genetic diversity and genome-wide linkage disequilibrium among U.S. wheat (*Triticum aestivum* L.) germplasm representing different market class. *Crop Sciences*, **47**: 1018–1030.

Chao, S.M., Zhang, W.J., Akhunov, E., Sherman, J., Ma, Y.Q., Luo, M.C., *et al*. (2009). Analysis of gene-derived SNP marker polymorphism in US wheat (*Triticum aestivum* L.) cultivars. *Molecular Breeding*, **23**: 23-33.

Charlesworth, B., Sniegowski, P. and Stephan, W. (1994). The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature,* **371**: 215-220.

Chen, J. and Gustafson, J.P. (1997). Chromosomal rearrangement of wheat (*T. aestivum*) chromosome 4A by *in situ* hybridization. *Chinese Journal of Genetics*, **24**: 39–47.

Chen, X. (2013). High-temperature adult-plant resistance: A key for sustainable control of stripe rust. *American Journal of Plant Sciences*, **4**: 608-627.

Chen, X. and Sullivan, P.F. (2003). Single nucleotide polymorphism genotyping: biochemistry, protocol, cost and throughput. *Pharmacogenomics Journal*, **3**: 77-96.

Chen, X., Min, D., Yasir, T.A. and Hu, Y.G. (2012). Genetic diversity, population structure and linkage disequilibrium in elite Chinese winter wheat investigated with SSR markers. *PloS One*, **7**(9): e44510.

Ciaffi, M., Lafiandra, D., Porceddu, E. and Benedettelli, S. (1993). Storage protein variation in wild emmer wheat (*Triticum turgidum* ssp. *Coccoides*) from Jordan and Turkey. I. Electrophoretic characterization of genotypes. *Theoretical and Applied Genetics*, **86**: 474-480.

Clegg, M., Brown and Whitefield, P.R. (1984). Chloroplast DNA diversity in wild and cultivated barley: implications for genetic conservation. *Genetic Research*, **43**: 339-343.

Close, P.S., Shoemaker, R.C. and Keim, P. (1989). Distribution of restriction site polynorphysm within the chloroplast genome of the genus Glycine subgenus Soja. *Theoretical and Applied Genetics*, **77**: 768-776.

Cnops, G., denBoer, B., Gerats, A., VanMontagu and VanLijsebettens (1996). X-Chromosome landing at the *Arabidoipsis* TORNADO1 locus using an AFLP based strategy. *Molecular Genomes and Genetics*, **253**: 32-41.

Coates, D.J. and Byrne, M. (2005). Genetic variation in plant populations: assessing cause and pattern. In: Plant diversity and evolution: Genotypic and phenotypic variation of higher plants. Henry, R. J., (ed.). CAB International, Wallingford. **pp**. 139–164.

Coe, E., Cone, K., McMullen, M., Chen, S.S., Davis, G., Gardiner, J., *et al*. (2002). Access to the maize genome: an integrated physical and genetic map. *Plant Physiology*, **128**: 9–12.

Condit, R. and Hubell, S.P. (1991). Abundance and DNA sequences of two base repeat regions in tropical tree genomes. *Genome*, **34**: 66-71.

Corazza-Nunes, M.J., Machado, M.A., Nunes, W.M.C., Cristofani, M. and Targon, M.L.P.N. (2002). Assessment of genetic variability in grape fruits *(Citrurpradise* Mscf.) and pummelos (*C. maxima* Burm.) Merr.) using RAPD and SSR markers. *Euphytica*, **126**: 169-176.

Cordeiro, G.M., Casu, R., McIntyre, C.L., Manners, J.M. and Henry, R.J. (2001). Microsatellite markers from sugarcane (*Saccharum* spp.) ESTs across transferable to *Erianthus* and *Sorghum. Plant Science*, **160**: 1115-1123.

Cox, T.S., Lookhart, G.L., Walker, D.E., Harrell, L.G., Albers, L.D. and Rodgers M.D. (1985). Genetic relationships among hard red winter wheat cultivars as evaluated by pedigree analysis and gliadin polyacrylamide -gel electrophoretic patterns. *Crop Sciences*, **25**: 1058-1063.

Cristopher, W.Y., Kafton, D. and Wilkes, G. (1984). Plant Genetic resources: A conservation imperative. AAAS Selected Symp. 87. Westview Press, Inc., Boulder, Colorado. **87**: 131-164.

Crossa, J., Burgueno, J., Dreisigacker, S., Vargas, M., Herrera, F.S.A., Lillemo, M., *et al*. (2007). Association analysis of historical bread wheat germplasm using additive genetic covariance of relatives and population structure. *Genetics*, **177**: 1889-1913.

Cuc, L.M., Mace, E.S., Crouch, J.H., Quang, V.D., Long, T.D. and Varshney, R.K. (2008). Isolation and characterization of novel microsatellite markers and their application for diversity assessment in cultivated groundnut (*Arachis*

hypogaea). BMC Plant Biology, **8**: 55. doi: https://doi.org/10.1186/1471-2229-8-55.

Dadong, Z., Guihua, B., Chengsong, Z., Jianming, Y. and Brett, F.C. (2010). Genetic diversity, population structure, and linkage disequilibrium in U.S. elite winter wheat. *The Plant Genome*, **3**: 117–127.

Dale, M.F.B., Ford-Lloyd, B.V. and Amold, M.H. (1985). Variation in some agonomically important characters in a germplasm collection of beet *(Beta vulgaris* L.). *Euphytica*, **34**: 449-455.

Deu, M., Hamon, P., Chantereau, J., Dufour, P., D' Hont, A. and Lanaud, C. (1995). Mt. DNA diversity in wild and cultivated sorghum. *Genome*, **38**: 635-645.

Devos, K.M., Moore, G. and Gale, M.D. (1995). Conservation of marker synteny during volution. *Euphytica*, **85**: 367-372.

Dholakia, B.B., AmmiRaju, J.S.S., Santra, D.K., Singh, H., Katti, M.V., Lagu, M.D., *et al*. (2001). Molecular marker analysis of protein content using PCR-based markers in wheat. *Biochemical Genetics*, **3**(9-10): 325–33.

Dice, L.R. (1945) Measures of the Amount of Ecologic Association between Species. *Ecology*, 26: 297-302.

Dieringer, D. and Schlotterer, C. (2003). Two distinct modes of microsatellite mutation processes: evidence from the complete genomic sequences of nine species. *Genome Research*, **13**: 2242-2251.

Dikshit, N., Sivaraj, N. and Das, A.B. (2014). Genotypic variations in agro-morphological and physico-chemical traits in ethnic rice (*Oryza sativa*) landracs of orissa. *Indian Journal of Agricultural Sciences*, **84**(8): 987-992.

Dograr, N., Akin-Yalin, S. and Akkaya, M.S. (2000). Discriminating durum wheatcultivars using highly polymoprhic simple sequence repeat DNA markers. *Plant Breeding*, **119**: 360-382.

Doll, H.; Brown, A. H. D. (1979). Hordein variation in wild (*Hordeum spontaneum*) and cultivated (*H. vulgare*) barley. *Canadian Journal of Genetics and Cytology*, **21**: 391-404.

Donmez, E., Sears, R.G., Shroyer, J.P. and Paulsen, G.M. (2001). Genetic Gain in Yield attributes of winter wheat in the Great Plains. *Crop Science*, **41**: 1412-1419.

Doyle, J.J. and Doyle, J.L. (1990) Isolation of plant DNA from fresh tissue. *Focus*, **12**: 13–15.

Dreisigacker, S., Zhang, P., Warburton, M.L., Skovmand, B., Hoisington, D. and Melchinger, A.E. (2005). Genetic diversity among and within CIMMYT wheat landrace accessions investigated with SSRs and implications for plant genetic resources management. *Crop Science*, **45**: 653-661.

Dubcovsky, J. and Dvorak, J.M. (2007). Genome plasticity: A key factor in the success of polyploid wheat under domestication. *Science Magazine*, **316**: 58-33.

Dubcovsky, J., Luo, M.C., Zhong, G.Y., Bransteiter, R., Desai, A., Kilian, A., *et al*. (1996). Genetic map of diploid wheat, (*Triticum monococcum* L.), and its comparison with maps of *Hordeum vulgare* L. *Genetics*, **143**: 983-999.

Dvorak, J. and Appels, R. (1982a). Chromosome and nucleotide sequence differentiation in genomes of polyploid *Triticum* species. *Theoretical and Applied Genetics*, **63**: 349-380.

Dvorak, J. and Zhang, H.B. (1982b). Reconstruction of the phylogeny of the genus *Triticum* from variation in repeated nucleotide sequences. *Theoretical and Applied Genetics*, **84**: 419-429.

Dvorak, J., Tetizi, P., H.B. and Resta, P. (1993). The evolution of polyploid wheats. Identification of the A genome donor species. *Genome*, **36**: 21-31.

Ehdaie, B. and Waines, J. (1989). Genetic variation, heritability and path-analysis in landraces of bread wheat from southwestern Iran. *Euphytica*, **41**(3): 183-190.

Ehrlich, H.A., Celfand, G.H. and Sninsky, J.J. (1991). Recent in the polymerase chain reaction. *Science*, **252**: 1643-1651.

Ellegren and Hans (2004). Microsatellites: simple sequences with complex evolution. *Nature Reviews*, **5**: 435-445.

Emon, R.M., Gustafson, J.P., Nguyen, H., Musket, T., Jahiruddin, M., Islam, M.A., *et al*. (2010). Molecular marker-based characterization and genetic diversity of

wheat genotypes in relation to Boron use efficiency. *Indian Journal of Genetics*, **70**(4): 339-348.

Ender, A., Schwenk, K., Stadler, T., Streit, B. and Schierwater, B. (1996). Rapid identification of microsatellites in Daphnia. *Molecular Ecology,* **5**: 437-447.

Epplen, J.T. (1992). Electrophoresisi. In: Advances in electrophoresis, Volume 5. Inchrambach, A., Dunn, N. J. and Radola, B.J. (eds). VCH Cambridge. *Vol* 5. **pp**. 59-112.

Erayman, M., Ilhan, E., Hakan, A., Eren, H., Gungor and Batuhan, A.K. (2016). Diversity analysis of genetic, agronomic, and quality characteristics of bread wheat (*Triticum aestivum* L.) cultivars grown in Turkey. *Turkish Journal of Agriculture and Forestry*, **40**: 83-94.

Erschadi, S., Haberer, G., Schoniger, M. and TorresRuiz, R.A. (2000). Estimating genetic diversity of *Arabidopsis thaliana* ecotypes with amplified fragment length polymorphisms (AFLP). *Theoretical and Applied Genetics*, **100**: 633-640.

Eticha, F., Bekele, E., Belay, G. and Borner, A. (2005). Phenotypic diversity in durum wheat collected from bale and wello regions of Ethiopia. *Plant Genetic Resources*, **3**(1): 35-43.

Evanno, G., Regnaut, S. and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**: 2611–2620.

Excoffier, L., Smouse, P. E. and Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distance among haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**: 479-491.

Fahima, T., Roder, M.S., Grams, A. and Nevo, E. (1998). Microsatellite DNA polymorphism divergence in *Triticum diccocoides* accessions highly resistant to yellow rust. *Theoretical and Applied Genetics*, **96**: 187-195.

FAO. (2002). http://www.fao.org/docrep/006/y4011e/y4011e08.htm (accessed on June 26, 2013)

FAO. (2010). http://www.fao.org/ wsfs/ world-summit/ wsfs-challenges/ en/ (accessed on June 26, 2013)

FAOSTAT. (2014). http://www.fao.org/faostat/en/#data/QC (accessed on December 14, 2016)

Feldman, M. (2001). The origin of cultivated wheat. In: The wheat book: A History of wheat breeding. Benjean, A. P. and Angus, W.J. (eds.). Lavoisier Publishing, Paris. **pp**. 3-56.

Feldman, M. and E.R. Sears, (1981). The wild gene resources of wheat. *Scientific American*, **44**: 102-12.

Flachowsky, H., Schumann, E., Weber, W.E. and Peil, A. (2001). Application of AFLP for the detection of sex-specific markers in hemp. *Plant Breeding*, **120**: 305-309.

Flavell, R.B. and Smith, D.B. (1976). Nucleotide sequence organization in the wheat genome. *Heredity*, **37**: 231-252.

Frankel, O.H. and Hawkes, (Eds.) (1975). Crop genetic resources for today and tomorrow. International Biological Programme, Vol.2, Cambridge University Press, Cambridge. **pp**. 150-205.

Fu, Y.B. (2005). Impact of plant breeding on genetic diversity of agricultural crops: searching for molecular evidence. *Theoretical and Applied Genetics*, **110**: 1505-1516.

Gadaleta, A., Giancaspro, A., Zacheo, S., Nigro, D., Giove, S.L., Colasuonno, P., *et al*. (2011). Comparison of genomic and EST-derived SSR markers in phylogenetic analysis of wheat. *Plant Genetic Resources: Characterization and Utilization*, **9**(2): 243–246.

Gale, M.D. and Youssefian, S. (1985). Dwarfing genes in wheat 1. In: Progress in Plant Breeding. Russell, G. E. (ed.). Butterworths and Co, London. **pp**.1-35.

Gawel, N.I., Jarret, R.L. and Whittemore, A. (1992). Restriction fragment length polymorphism (RFLP)- based phylogenetic analysis of Musa. *Theoretical and Applied Genetics*, **84**: 286-290.

Gepts, P. (1995). Genetic markers and core collection. In: Core collections of plant genetic resources. Hodgkin, Brown, T., A.H.D., Th, J.C., Hintum, V. and Morales, E. A. V. (eds.). John Wiley and Sons, (UK). **pp**. 127-146.

Ghareyazie, B., Huang, N., Second, G., Bennet, J. and Khush, G.S. (1995). Classification of rice germplasm I. Analysis using ALP and PCR-based RFLP. *Theoretical and Applied Genetics*, **91**: 218-227.

Gibson, L. and Benson, G. (2002). Origin, history, and uses of oat (*Avena sativa*) and wheat (*Triticum aestivum*). Iowa State University, Department of Agronomy. **pp**.1-4. Online article accessed on 15.12.2016 (URL http://www.agron.iastate.edu/courses/agron212/Readings/Oat_wheat_history.htm)

Gill, K.S. (2004). Gene distribution in cereal genomes. In: Cereal genomics. Gupta., P. K. and Varshney, R.K. (eds). Kluwer Academic Publisher, Dordrecht, **pp**. 361-384.

Gill, K.S., Gill, B.S., Endo, T.R. and Boyko, E.V. (1996a). Identification of high-density mapping of gene-rich regions in chromosome group 5 of wheat. *Genetics*, **143**: 1001-1012.

Gill, K.S., Gill, B.S., Endo, T.R. and Taylor, T. (1996b). Identification and high-density mapping of gene-rich regions in chromosome group 1 of wheat. *Genetics*, **144**: 1883-1891.

Godfray, H.C.J., Beddington, J.R., Crute, I.R., Haddad, L., Lawrence, D., Muir, J.F., *et al*. (2010). Food security: the challenge of feeding 9 billion people. *Science*, **327**: 812-818.

Godt, M.J. and Hamrick, J.L. (1999). Population genetic analysis of *Elliottia racemosa* (Ericaceae), a rare Georgia shrub. *Molecular Ecology*, **8**: 75-82.

Goldstein, D.B. and Pollock, D.D. (1997). Launching microsatellites: a review of mutation processes and methods of phylogenetic inference. *Journal of Heredity*, **88**: 335-342.

Goldstein, D.B. and Schlotterer, C. (1999). Microsatellites: Evolution and Applications. Oxford Press, Oxford, UK.

Goncalves, L.S., Rodrigues, R., Maral, A. and Karasawa, M. (2008). Comparison of multivariate statistical algorithms to cluster tomato heirloom accessions. *Genetics and Molecular Research*, **7**: 1289-1297.

Goncalves, J.P., Madeira, S.C. and Oliveira, A.L. (2009). BiGGEsTS: Integrated environment for biclustering analysis of time series gene expression data. *BMC Research Notes,* **2**: 124

Grabau, E.A., Davis, W.H., Phelps, N.D. and Gengenbach, B.G. (1992). Classification of soybean cultivars based on mitochondrial DNA restriction fragments length polymorphisms. *Crop Sciences*, **32**: 271-274.

Grimaldi, M.C. and Crouau-Roy, B.B. (1997). Microsatellite allelic homoplasy due to variable flanking sequences. *Journal of Molecular Evolution*, **44**: 336-340.

Grodzicker, T., Williams, J., Sharp, P. and Sambrook, J. (1975). Physical mapping of temperature sensitive mutants of adenovirus. Cold Spring Harbor Symposia on Quantitative Biology, **39**: 439-446.

Gu, W. K., Waden, N. F., Yu, J. and Wallace, D. H. (1995). Large-scale, cost effective screening of PCR products in marker-assisted selection applications. *Theoretical and Applied Genetics,* **91**: 465–70.

Gupta, M., Chawla, V., Garg, P., Yadav, N., Munjal, R. and Sharma, B. (2015). Genetic analysis of yield and heat stress related traits in wheat (*Triticum aestivum* L. em. Thell) using microsatellite markers. *Journal of Applied and Natural Science*, **7**(2): 739-744.

Gupta, M., Chyi, Y.S., Romero-Severson, J. and Owen, J.L. (1994). Amplification of DNA markers from evolutionarily diverse genomes using single primers of simple-sequence repeats. *Theoretical and Applied Genetics,* **89**: 998-1006.

Gupta, P.K. and Varshney, R.K. (2000). The development and use of microsatellite markers from genetic analysis and plant breading with emphasis on bread wheat. *Euphytica,* **113**: 165-185.

Gupta, P.K., Kulwal, P.L. and Rustga, S. (2005). Wheat cytogenetic in the genomic era and its relevance to breeding. *Cytogenetic and genome research*, **109**: 315-327.

Gupta, P.K., Roy, J.K. and Prasad, M. (2001). Single nucleotide polymorphisms: A new paradigm for molecular marker technology and DNA polymorphism detection with emphasis on their use in plants. *Current Science*, **80**: 524-535.

Gupta, P.K., Varshney, R.K., Sharma, P.C. and Ramesh, B. (1999). Molecular markers and their applications in wheat breeding. *Plant Breeding*, **118**: 369-390.

Gupta, P.K., Rustgi, S., Sharma, S., Sing, R., Kumar, N. and Balyan, H.S. (2003). Transferable EST-SSR markers for the study of polymorphism and genetic diversity in bread wheat. *Molecular Genetics and Genomics*, **270**: 315-323.

Gurung, S., Mamidi, S., Bonman, J.M., Xiong, M., Brown-Guedira, G. and Adhikari, T.B. (2014). Genome-wideassociation study reveals novel quantitative trait loci associated with resistance to multiple leaf spot diseases of spring wheat. *PloS One*, **9**(9): e108179.

Hafida, Z., Hassa, Q., Sripada, M.H., Fatima, G. and Abdarraouf, H. (2012). Agro-morphological variability in durum wheaat landraces of morocco. *Australian Journal of Crop Sciences*, **6**(7): 1172-1178.

Hai, L., Guo, H.J., Wagner, C., Xiao, S.H. and Friedt, W. (2008). Genomic regions for yield and yield parameters in Chinese winter wheat (*Triticum aestivum* L.) genotypes tested under varying environments correspond to QTL in widely different wheat materials. *Plant Science*, **175**(3): 226-232.

Hailu, F., Merker, A. and Johnsson, E. (2010). Patterns of phenotypic diversity for phenologic and qualitative traits in ethiopian durum wheat germplasm. *Genetic Resource and Crop Evolution*, **53**: 1089-1098.

Halaieva, M.V., Fait, V.I., Chebotar, S.V., Halaiev, O.V. and Syvolap, I. (2013). Association of microsatellite loci alleles of the group-5 chromosomes with frost resistance of winter wheat. *Tsitol Genetics*, **47**(5): 3-11.

Hamada, H., Petrino, M.C. and Takugana, T. (1982). A novel repeated element with Z-forming potential is widely found in evolutionarily diverse eukaryotic genomes. *Proceedings of National Academy of Science (USA)*, **79**: 6465-6469.

Hammer, K., Filatenko, A.A. and Korzun, V. (2002). Microsatellite markers - A new tool for distinguishing diploid wheat species. *Genetic Resources and Crop Evolution*, **47**(5): 497-505.

Hamrick, J.L. and Godt, M.J. (1990). Allozyme Diversity in plant Species. In: Plant Population Genetics, Breeding and Genetic Resources, Brown, A.H.D., Clegg, M.T., Kanler, A.L. and Weir, B.S. (eds.). Sin over Associates Inc. Massachusetts, **pp**. 43-63.

Hancock, J.M. (1995). The contribution of slippage like process to genome evaluation. *Journal of Molecular Evolution*, **41**: 1038-1047.

Hancock, J.M. (1999). Microsatellites and other simple sequences: genomic context and mutational mechanisms. In: Microsatellites: evolution and applications Goldstein, D.B. and Schlotterer, C. (eds). Oxford University Press, Oxford. **pp**. 1-9.

Hao, Y., Chen, Z., Wang, Y., Bland, D., Buck, J., Brown-Guedira, G., *et al*. (2011). Characterization of a major QTL for adult plant resistance to stripe rust in US soft red winter wheat. *Theoretical and Applied Genetics*, **123**: 1401-1411.

Harding, R.M., Boyce, A.J. and Clegg, J.B. (1992). The evolution of tenderly repetitive DNA: recombination rules. *Genetics*, **132**: 847-859.

Hardon, J.J., Vosmon, B. and Hintum, T.J.L.V. (1994). Identifying genetic resources and their origin. The capabilities and limitation of modem biochemical and legal systems. CPGR Back ground Paper No. 4. FAO, Rome.

Harlan, J.R. (1981). The early history of wheat: earliest traces to the sack of Rome. In: Wheat Science - Today and Tomorrow. L.T. Evans, W.J. Peacock (eds.). Cambridge University Press, Cambridge, **pp**. 1-19.

Hartl, L., Mohler, V., Zeller, F.J., Hsam, S.L.K. and Schweizer, G. (1999). Identification of AFLP markers closely linked to the powdery mildew resistance genes Pmlc and Pm4a in common wheat (*Triticum aestivum* L). *Genome*, **42**: 322-329.

Hearne, C.M., Ghosh, S. and Todd, J.A. (1992). Microsatellites for linkage analysis of genetic traits. *Trends in Genetics*, **8**(8): 288-294.

Heath, D.D., Iwarma, G.K. and Devin, R.H. (1993). PCR primed with VNTR core sequences yields species specific patterns and hypervariable probes. *Nucleic Acids Research*, **21**: 5782-5785.

Heidari, B., Saeidi, G., Tabatabaei, B.E.S. and Suenaga. K. (2012). QTLs involved in plant height, peduncle length and heading date of wheat (*Triticum aestivum* L.). *Journal of Agricultural Science and Technology*, **14**: 1093-1104.

Helentjaris, T., Slocum, M., Wright, S., Schaefer, A. and Nienhuis, J. (1986). Construction of genetic linkage maps in maize and tomato using restriction fragment length polymorphisms. *Theoretical and Applied Genetics*, **61**: 650-658.

Heslop-Harrison, J.S., Leitch, A.R., Scwarzacher, T. and Anamthawat-Jonsson, K. (1990). Detection and characterization of 1B/1R translocations in hexaploid wheat. *Heredity*, **65**: 385-392.

Heun, M., Schafer-Pregl, R., Klawan, D., Castagna, R., Accerbi, M., Borghi, B., *et al*. (1997). Site of einkorn wheat domestication identified by DNA fingerprinting. *Science*, **278**: 1312-1314.

Hiebert, C., Thomas, J. and McCallum, B. (2005). Locating the broad-spectrum wheat leaf rust resistance gene Lr52 (LrW) to chromosome 5B by a new cytogenetic method. *Theoretical and Applied Genetics*, **110**: 1453-1457.

Hill, M., Witsenboer, H., Zabeau, M., Vos, P., Kesseli, R. and Michelmore R. (1996). PCR-based fingerprinting using AFLPs as tool for studying genetic relationship in Lactuca spp. *Theoretical and Applied Genetics*, **93**: 1202-1210.

Hillis, D.M. (1987). Molecular Versus Morphological Approaches To Systematics. *Annual Reviews in Ecology and Systematics*, **18**(1): 23-42.

Hirachand, Srivastava, K.S. and Metha, S.K. (1978). Genetic variability and correlations coefficient analysis in wheat (*Triticum aestivum* L.). *Madras Agricultural Journal*, **65**: 17-37.

Hohmann, U., Endo, T.R., Gill, K.S. and Gill, B.S. (1994). Comparison of genetic and physical maps of group 7 chromosomes form *Triticum aestivum* L. *Molecular Genetics and Genetics*, **245**: 644-653.

Hoisington, D., Khairallah, M., Reeves, T., Jean-Marcel, R., Skovmand, B., Suketoshi, T., *et al*. (1999). Plant genetic resources: What can they contribute toward increased crop productivity. *Proceedings of National Academy of Sciences (USA)*, **96**: 5937-5943.

Hongtrakul, L. (1997). AFLP analysis of relationships among cassava and other mammihot species. *Theoretical and Applied Genetics*, **95**: 741-750.

Hoque, M. and Rahman, L. (2006). Estimation of euclidean distance for different morpho-physilogical characters in some wild and cultivated rice genotypes(*Oryza sativa* L.). *Pakistan Journal of Science*, **1**: 77-79.

Hoshino, A.A., Bravo, J.P., Morelli, K.A. and Nobile, P.M. (2012). Microsatellites as tools for genetic diversity analysis. In: genetic diversity in microorganisms. INTECH Open Access Publisher. **pp**. 149-170.

Hoshino, A.A., Bravo, J.P., Nobile, P.M. and Morelli, K.A. (2012). Microsatellites as tools for genetic diversity analysis. In: Genetic diversity in microorganisms. Caliskan, M. (ed.). In Tech, Janeza Trdine, Rijeka, Croatia, **pp**. 149-170.

Hu, Z., Zhang, D., Zhang, G., Kan, G., Hong, D. and Yu, D. (2014). Association mapping of yield-related traits and SSR markers in wild soybean (*Glycine soja* Sieb. and Zucc.). *Breeding Science*, **63**: 441-449.

Huang, Q., Borner, A., Roder, S. and Ganal, W. (2002). Assessing genetic diversity of wheat (*Triticum aestivum* L.) germplasm using microsatellite markers. *Theoretical and Applied Genetics*, **105**: 699–707.

Huang, S., Sirikhachornkit, A., Su, X., Gill, B., Haselkorn, R. and Gornicki, P. (2002). Genes encoding plastid acetyl-CoA caboxylase and 3-phosphoglycerate kinase of the *Tritoicum/Aegilops* complex and the evolutionary history of polyploid wheat. *Proceedings of National Academy of Sciences* (*USA)*, **99**: 8133-8138.

Huang, X.Q., Kempf, H., Canal, M.W. and Roder, M.S. (2004). Advanced backcross QTL analysis in progenies derived from a cross between a German elite winter wheat variety and synthetic wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics*, **109**: 933-943.

Huem, M., SchaferPergl, R., Klawan, D., Castagna, R., Accerbi, M., Borghi, B., *et al*. (1997). Site of einkorn wheat domestication identified by DNA fingerprinting. *Science*, **278**: 1312-1314.

Hunter, R.L. and Markert, C.L. (1957). Histochemical demonstration of enzymes separated by zone electrophoresis in starch gels. *Science*, **125**: 1294-1295.

IBPGR/IPGRI. (1985). Descrptors for wheat and barley. International Board for Plant Genetic Resources, Rome, Italy.

Idrees, M. and Irshad, M. (2014). Molecular markers in plants for analysis of genetic diversity: A review. *European Academic Research*, **2**(1): 1513-1540.

Imran, J., Shahid, I.A., Hafiz, M.A. and Aurangzeb, R. (2016). Assessment of genetic diversity in wheat synthetic double haploids for yield and drought related traits through factor and cluster analyses. *Bio Publisher*, **7**: 1-9.

International Wheat Genome Sequencing Consortium. (2014). A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. *Science*, **345** (doi: 10.1126/ science.1251788).

Jain, S.K. (1977). Genetic diversity of weedy rye populations in Califomia. *Crop Science*, **17**: 480-482.

Jain, S.K. and Singh, R.S. (1972). Population biology of *Avena* U. Isoenzyme polyrnorphisms in populations of Mediterranean region and central California. *Theoretical and Applied Genetics*, **41**: 79-84.

Jain, S.K., Qualset, C.O., Bhatt, G.M. and Wu, K.K. (1975). Geographical patterns of phenotypic diversity in a world collection of dipyletic origin of polyploid wheats. *Chromosome Research*, **2**: 59-64.

Jakupciak, J.P. and Wells, R.D. (2000). Gene conversion (recombination) mediates expansions of CTG.CAG repeats. *Journal of Biological Chemistry*, **275**: 4003-4013.

Jamalirad, S., Mohammadi, S.A. and Toorchi, M. (2012). Assessing genetic diversity in a set of wheat (*Triticum aestivum*) genotypes using microsatellite markers to improve the yellow rust resistant breeding programs. *African Journal of Agricultural Research*, **7**(48): 6447-6455.

Jarne, P. and Lagoda, P.J.L. (1996). Microsatellites, from molecules to populations and back. *Tree*, **11**: 424-429.

Jarret, R.L., Gawel, N. and Whittemore, A. (1992). Phylogenetic relationships of the sweet potato (*Ipomoea batatas* (L.) Lam.). *Journal of American Society for Horticultural Sciences*, **117**: 633–637.

Jarvis, P., Lister, C., Szabo, V. and Dean, C. (1994). Integration of CAPs markers into the RFLP map generated using recombinant inbred lines of *Arabidopsis thaliana*. *Plant Molecular Biology*, **24**: 685-667.

Jarwan, A.P. and Wells, R.A. (1989). Hypervariable minisatellites: Recombinators or innocent bystanders? *Trends in Genetics*, **5**: 367-371.

Jeffery, A.J., Wilson, V. and Thein, S.L. (1985a). Hyper variable mini satellite regions in human DNA. *Nature*, **314**: 67-73.

Jeffreys, A.J., Wilson, V. and Thein, S.L. (1985b). Individual- specific fingerprints of human DNA. *Nature*, **316**: 76-79.

Jena, K.K. and Kochert, G.C. (1991). Restriction fragment length polymorphism analysis of CCDD genome species of the genus *Oryza* L. *Plant Molecular Biology*, **16**: 831-839.

Jiang, J. and Gill, B.S. (1994). Different species-specific chromosome translocations in *Triticum timopheevi* and *Triticum turgidum* support the diphyletic origin of polyploid wheats. *Chromosome Research*, **2**: 59-64.

Jin, L., Macaubas, C., Hallmayer, J., Kimura, A. and Mignot, E. (1996). Mutation rate varies among alleles at a microsatellite locus: Phylogenetic evidence. *Proceedings of National Academy of Sciences* (USA), **93**: 15285-15288.

Jin, X.J., Sun, D.F., Li, H.Y., Yang, Y. and Sun, G.L. (2013). Characterization and molecular mapping of a dwarf mutant in wheat. *Genetic and Molecular Research*, **12**(3): 3555-3565.

Jing Runchun, Johnson Richard, Seres Andrea, Kiss Gyorgy, Ambrose Mike J., Knox Maggie R., Ellis T. H. Noel and Flavell Andrew J. (2007). Gene-based sequence diversity analysis of field pea (*Pisum* spp.). *Genetics*, **177**: 2263–2275.

Jones, H., Roger, A. Leigh, R.A., Tomos, A.D. and Jones, R.G.W. (1987). The effect of abscisic acid on cell turgor pressures, solute content and growth of wheat roots. *Planta*, **170**(2): 257-262.

Joshi, A.K., Mishra, B., Chatrath, R., Ferrara, O.G. and Singh, R.P. (2007). Wheat improvement in India: Present status, emerging challenges and future prospects. *Euphytica*, **157**(3): 431-446.

Joshi, B.K., Mudwari, A., Bhatta, M.R. and Ferrara, G.O. (2004). Genetic diversity in Nepalese wheat cultivars based on agromorphological traits and coefficient of parentage. *Nepal Agricultural Research Journal*, **5**: 7-17.

Joshi, C.P. and Nguyen, H.T. (1993). Application of random amplified polymorphic DNA technique for the detection of polymorphism among wild and cultivated tetraploid wheats. *Genome*, **36**: 602-609.

Joshi, M.G. and Singh, B. (1979). Genetic divergence among tetraploid Triticum species. *Indian Journal of Genetics* (The), **39**:188-193.

Kalender, R., Grob, T., Regina, M., Suoniemi, A. and Schulman, A. (1999). IRAP and REMAP: two new retrotransposon-based DNA fingerprinting techniques. *Theoretical and Applied Genetics*, **98**: 704-711.

Kalia, R.K., Rai, M.K., Kalia, S., Singh, R. and Dhawan, A.K. (2011). Microsatellite markers: an overview of the recent progress in plants. *Euphytica*, **177**: 309-334.

Kandel, Y.R., Glover, K.D., Osborne, L.E. and Gonzalez-Hernandez, J.L. (2015). Mapping quantitative resistance loci for bacterial leaf streak disease in hard red spring wheat using an identity by descent mapping approach. *Euphytica*, **201**(1): 53-65.

Kantety, R.V., Rota, M.L., Matthews, D.E. and Sorrells, M.E. (2002). Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. *Plant Molecular Biology*, **48**: 501-510.

Karp, A. and Edwards, K.J. (1997). DNA markers: a global overview. In: DNA markers: protocols, application and overviews. Caetano-Anolles, G. and P.M. Greshoff. (eds.). Wiler-VCH, New York. **pp**. 1-13.

Karp, A., Kresovich, S., Bhat, K.V., Ayad, W.G. and Hodgkin, T. (1997). Molecular tools in plant genetic resources conservation: A guide to the technologies. IPGRI Technical Bulletin, No. 2, IPGRI, Rome.

Karp, A., Seberb, O. and Buiatti, M. (1996). Molecular techniques in the assessment of botanical diversity. *Annals of Botany*, **78**: 143-149.

Keystone-Center (1991). Biological diversity on federal lands: report of a keystone policy Dialogue. Keystone center. Keystone CO. **pp**. 96.

Khaled, A.G.A. and Hamam, K.A. (2015). Association of Molecular Markers with Phenotypic traits of bread wheat genotypes. *Egyptian Journal of Genetics and Cytology*, **44**: 115-130.

Khan, M.K., Pandey, A., Choudhary, S., Hakki, E.E., Akkaya, M.S. and Thomas, G. (2014). From RFLP to DArT: molecular tools for wheat (*Triticum* spp.) diversity analysis. *Genetic Resources and Crop Evolution*, **61**: 1001–1032.

Khan, M.K., Pandey, A., Thomas, G., Akkaya, M.S., Kayis, S.A., Ozsensoy, Y., *et al*. (2015). Genetic diversity and population structure of wheat in India and Turkey. *AoB Plants*, **7**: plv083; doi:10.1093/aobpla/plv083.

Khan, O.U., Ijaz, F. and Muhammad, T.S. (2015). Heritability Analysis for Yield Associated Traits in Wheat (*Triticum aestivum* L.). *Science, Technology and Development*, **34**(4): 260-264.

Khodadadi, M., Fotokian, M.H. and Miransari, M. (2011). Genetic diversity of wheat (*Triticum aestivum* L) genotypes based on cluster and principal component analyses for breeding strategies. *Australian Journal of Crop Science*, **5**(1): 17-24.

Kihara, C. (2012). Physical mapping of DNA-based molecular markers and genome analysis in bread wheat (*Triticum aestivum* L. em Thell.). Ph.D. thesis (unpublished) submitted to Ch. Charan Singh University, Meerut-250004, State of Uttar Pradesh, India.

Kihara, H. (1924). Cytologische and gentische Studien bei wichtigen. Getreidearten mitbesondener Rucksicht auf das Verhalten der Chromosmen und die Steriitat

in den Bastarden. Memoirs of the College of Science; Kyoto Imperial University. 1B 1-200. Kyoto, Japan. (English version accessed on 10.3.2012).

Kihara, H. (1930). Genomanalyse bei *Triticum* and *Aegilops* II. *Cytologia*, **2**: 106-156.

Kihara, H. (1954). Consideration on the evolution and distribution of *Aegilops* species based on the analyser method. *Cytologia*, **19**: L336-357.

Kihara, H. (1983). In: Proceedings of the 6th International Wheat Genetic Symposium. S. Sakamoto, (eds), Plant Germplasm Institute, University of Kyoto, Japan.

Kim, K.H., Jung, J.U., Kang, C.S., Shin, S.H., Park, J.C., Son, J.H., *et al*. (2014). Mapping of QTL for Pre-harvest sprouting resistance and seed dormancy in a doubled haploid population. *Plant Breeding and Biotechnology*, **2**(3): 247-256.

Kimber, G. (1971). The rationale of measuring chromosomes. *Seiken Ziho*, **22**: 5-8.

Kimber, G. and Feldman, M. (1987). Wild wheat - An introduction, special Report 353, Univ. Missouri-Columbia, USA.

Kliebenstein, D.J., Gershenzon, J. and Mitchell-Olds, T. (2001). Comparative quantitative trait loci mapping of aliphatic, indolic and benzylic glucosinolate production in *Arabidopsis thaliana* leaves and seeds. *Genetics*, **159**: 359-370.

Kojima, T., Nagaonka, T., Noka, K. and Ogijhara, Y. (1998). Genetic linkage map of ISSR and RAPD markers in eikorn wheat in relation to that of RFLP markers. *Theoretical and Applied Genetics*, **96**: 37-45.

Koniecyzn, A. and Ausubel, F.M. (1993). A procedure for mapping Arabidopsis mutations using co-dominant ecotype-specific PCR-based markers. *Plant Journal*, **4**: 403–410.

Kresovich, S., Mc Pherson, J.R. and Westman, A.L. (1997). Using molecular markers in genebanks: identity, duplication, contamination and regeneration. In: Proceedings of the IPGRl workshop on molecular genetic techniques for plant genetic resources, October 9-11, 1995. Ayad, W. G., Hodgkin, T., Jaradat, A. and Rao, V. R. (eds.). IPGRI, Rome, Italy. **pp**. 23-38.

Krull, C.F. and Borlaug, N.E. (1970). The utilization of collection in plant breeding and production. In: Genetic Resources in Plants: Their Exploitation and Conservation. Frankel, O. H. and Bennet, E. (eds.). IBP Hand Book No.11. Blackwell Scientific Publisher, Oxford, United Kingdom, **pp**. 427-439.

Kumar, P., Gupta, V.K., Misra, A.K., Modi, D.R. and Pandey, B.K. (2009). Potential of molecular markers in plant biotechnology. *Journal of Plant Omics*, **2**: 141-162.

Kumar, P., Yadava, R.K., Kumar, S. and Kumar, P. (2016a). Molecular diversity analysis in wheat genotypes using SSR markers. *Electronic Journal of Plant Breeding*, **7**(2): 464-468.

Kumar, S., Archak, S., Tyagi, R.K., Kumar, J., VK, V., Jacob, S.R., *et al*. (2016b) Evaluation of 19,460 Wheat accessions conserved in the Indian national genebank to identify new sources of resistance to rust and spot blotch diseases. *PLoS ONE*, **11**(12): e0167702.

Kumar, S., Kumar, V., Kumari, P., Kirti, Singh, A.K. and Singh, R. (2016c). DNA fingerprinting and genetic diversity studies in wheat genotypes using SSR markers. *Journal of Environmental Biology*, **37**: 319-326.

Kumar, S., Singh, S.S., Mishra, C.N., Saroha, M., Gupta, V., Sharma, P., *et al*. (2015). Assessment of tiller inhibition (tin) gene molecular marker for its application in marker-assisted breeding in wheat. *National Academy of Science Letters*, **38**(6): 457–460.

Kunert, A., Naz, A.A., Dedeck, O., Pillen, K. and Le´on, J. (2007). AB-QTL analysis in winter wheat: I. Synthetic hexaploid wheat (*T. turgidum* ssp. *dicoccoides* X *T. tauschii*) as a source of favorable alleles for milling and baking quality traits. *Theoretical and Applied Genetics*, **115**: 683-695.

Kuzmanovic, L., Gennaro, A., Benedettelli, S., Dodd, I.C., Quarrie, S.A. and Ceoloni, C. (2014). Structural–functional dissection and characterization of yield-contributing traits originating from a group 7 chromosome of the wheatgrass species Thinopyrum ponticum after transfer into durum wheat. *Journal of Experimental Biology*, **65**(2): 509-525.

Kwon, S., Hong, S., Son, J., Lee, J.K., Cha, Y., Eun, M., *et al*. (2006). CACTA and MITE transposon distributions on a genetic map of rice using F15 RILs derived from Milyang 23 and Gihobyeo hybrids. *Molecules and Cells*, **21**: 360-366.

Lagercrantz, U., Ellegren, H. and Andreson, L. (1993). The abundance of various polymorphic microsatellite motifs differ between plants and vertebrates. *Nucleic Acids Research,* **21**: 1111-1115.

Laido, G., Marone, D., Russo, M.A., Colecchia, S.A., Mastrangelo, A.M., Vita, P.D., *et al*. (2014). Linkage disequilibrium and genome-wide association mapping in tetraploid wheat (*Triticum turgidum* L.). *Plos One*, **9**(4): e95211.

Law, J.R., Dornini, P., Koebner, R.M.D. and Reeves, J.C. **(**1998). DNA profiling and plant variety registration. The statistical assessment of distinctness in wheat using Amplified Fragment Length Polyrnorphisms. *Euphytica,* **102**: 335-342.

Lee, D., Reeves, J.C. and Cooke, R.J. (1995). The use of DNA-based markers for distinctness, uniformity and stability testing in oilseeds, rape and Barley. UPOV working group on biochemical and molecular techniques and DNA profiling in particular. UPOV, Geneva, Switzerland. **pp**. 3-4.

Leigh, F., Lea, V., Law, J., Wolters, P., Powell, W. and Donini, P. (2003). Assessment of EST- and genomic microsatellite markers for variety discrimination and genetic diversity studies in wheat. *Euphytica*, **133**: 359-366.

Leonardo, S., Vanzetti, Yerkovich, N., Chialvo, E., Lombardo, L., Vaschetto, L., *et al*. (2013). Genetic structure of Argentinean hexaploid wheat germplasm. *Genetics and Molecular Biology*, **36**(3): 391-399.

Lerceteau, E. and Szmidt, A.E. (1999). Properties of AFLP markers in inheritance and genetic diversity studies of *Pinups sylvestries* L. *Heredity,* **82**: 252-260.

Levinson, G. and Gutman, G.A. (1987). Slipped strand mispairing: A major mechanism for DNA sequence evolution. *Molecular and Biology Evolution*, **4**: 203-221.

Levy, A.A. and Feldman, M. (1998). Ecological distribution of HMW glutenin alleles in population of wild tetraploid wheat *Triticum turgidum* var. *dicoccoides. Theoretical and Applied Genetics,* **75**:651-658.

Levy, A.A. and Fieldman, M. (2002). The impact of polyploidy on grass genome evolution. *Plant Physiology*, **130**: 1587-1593.

Levy, A.A., Galili, A. and Feldman, M. (1998). Polymorphism and genetic control of high-molecular-weight glutenin subunits in wild tetraploid wheat *Triticum turgidum* var. *dicoccoides. Heredity*, **61**:63-72.

Lewis, R.S. (2006). Identification of germplasm of possible value for confronting an unfavorable inverse genetic correlation in tobacco. *Crop Sciences*, **46**: 1764-1771.

Li, G.H.S., Bucklin, K., Ribes, V. and Hedgecock, D. (2003). Characterization of 79 microsatellite DNA markers in the Pacific oyster *Crassostrea gigas*. *Molecular Ecology Notes*, **3**: 228-232.

Li, H.M., Liang, H., Li, Z., Tang, Z.X., Fu, S.L., Geng, Y.Y., *et al*. (2015). Dynamic QTL analysis of protein content and glutamine synthetase activity in recombinant inbred wheat lines. *Genetic and Molecular Research*, **14**(3): 8706-8715.

Li, W. and Gill, B.S. (2002). The collinearity of the Sh2/A1 orthologous region in rice sorghum and maize is interrupted and accompanied by genome expansion in the Triticale. *Genetics*, **160**: 1153–1162.

Li, W., Zhang, B., Li, R., Chang, X. and Jing, R. (2015). Favorable alleles for stem water-soluble carbohydrates identified by association analysis contribute to grain weight under drought stress conditions in wheat. *PLoS ONE*, **10**(3): e0119438.

Li, Y.C., Fahima, T., Bieles, A., Korol, A.B. and Navo, E. (1999). Microclimatic stress and adaptative DNA differentiation in wild emmer wheat, *Triticum dicoccoides*. *Theoretical and Applied Genetics*, **98**: 873-883.

Li, Y.C., Fahima, T., Roder, M.S., Kirzhner, V.M., Beiles, A., Korol, A.B., *et al*. (2003b). Genetic effects on micro-satellite diversity in wild emmer wheat, *Triticum dicoccoides*, at Yehudiyya microsite. *Heredity*, **90**: 150-156.

Li, Y.C., Korol, A.B., Fahima, T. and Nevo, E. (2004). Microsatellites within genes: structure, function, and evolution. *Molecular Biology and Evolution*, **21**: 991-1007.

Li, Y.C., Korol, A.B., Fahima, T., Beiles, A. and Nevo, E. (2002). Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Molecular Ecology*, **11**: 2453-2465.

Li, Y.C., Roder, M.S., Fahima, T., Kirzner, V.M., Beiles, A., Korol, A.B., *et al*. (2000). Natural selection causing microastellite divergence in wild emmer wheats at the ecological variabile microsite at Ammiad, Isreal. *Theoretical and Applied Genetics*, **100**: 985-999.

Linman, M.J., Yu, H., Chen, X. and Cheng, Q. (2009). Fabrication and characterization of a sialoside-based carbohydrate microarray bio interface for protein binding analysis with surface Plasmon resonance imaging. *merican Chemical Society*, **1**: 1755-76.

Linnaeus, C. (1953). Species Plantarum Laurentii Salvii, Holmiae, The Hagu, Litt, M. and Luty, J.A. (1989). A hypervariable microsatellite revealed by *in vitro* amplification of a dinucleotide repeat within the cardiac muscle actin gene. *American Journal of Human Genetics*, **44**: 397-401.

Liu, J. (2017). PowerMarker V3.0 Manual. Accessed on Apri 10, 2017. URL: http://www.powermarker.net

Liu, J., Shi, S., Chang, E., Yang, W. and Jiang, Z. (2013). Genetic diversity of the critically endangered Thuja sutchuenensis revealed by ISSR markers and the implications for conservation. *International Journal of Molecular Sciences*, **14**: 14860–14871.

Liu, K. and Muse, S.V. (2005). Power Marker: An integrated analysis environment for genetic marker analysis. *Bioinformatics*, **21**: 2128–2129.

Liu, M., Smith, M. and Gill, S. (2002). Identification of microsatellite markers linked to Russian wheat aphid resistance genes *Dn4* and *Dn6*. *Theoretical and Applied Genetics*, **104**(6-7): 1042-1048.

Liu, X.M., Gill, B.S. and Chen, M.S. (2005). Hessian fly resistance gene H13 is mapped to a distal cluster of resistance genes in chromosome 6DS of wheat. *Theoretical and Applied Genetics*, **111**: 243-249.

Lu, H.J., Fellers, J.P., Friesen, T.L., Meinhardt, S.W. and Faris, J.D. (2006). Genomic analysis and marker development for the *Tsn1* locus in wheat using bin-mapped ESTs and flanking BAC contigs. *Theoretical and Applied Genetics*, **112**: 1132-1142.

Luo, M., Yang, Z., You, F., Kawahara, T., Waines, J. and Dvorak, J. (2007). The structure of wild and domesticated emmer wheat populations, gene flow between them, and the site of emmer domestication. *Theoretical and Applied Genetics*, **114**: 947–959.

Lyamichev, V., Brow, M.A.D. and Dahlberg, J.E. (1993). Structure-specific endonucleolytic cleavage of nucleic acids by eubacterial DNA polymerase. *Science*, **260**:778-783.

Ma, B.L., Yan, W., Dwyer, L.M., Fregeau-Reid, J., Voldeng, H.D., Dion, Y., *et al*. (2004). Graphic analysis of genotype, environment, nitrogen fertilizer, and their interactions on spring wheat yield. *Journal of Agronomy*, **96**: 169-180.

Ma, Z.Q. and Lapitan, N.L.V. (1998). A comparison of amplified and restriction fragment length polymorphism in wheat. *Cereal Research (Common)*, **26**: 7-13.

Maccaferri, M., Stefanelli, S., Rotondo, F., Tuberosa, R., Sanguineti, M.C. (2007). Relationships among durum wheat accessions. I. Comparative analysis of SSR, AFLP, and phenotypic data. *Genome*, **50**(4): 373-384.

MacKill, D.J., Zhang, Z., Redona, E.D. and Colowit, P.W. (1996). Level of polymorphism and genetic mapping of AFLP markers in rice. *Genome*, **39**: 969-977.

Maestra, R. and Naranjo, T. (1998). Homologous relationships of *Aegilops speltoides* chromosomes of bread wheat. *Theoretical and Applied Genetics*, **97**: 181-186.

Maity, S. and Das, S. (2015). Morphological Characterization of Bread Wheat (*Triticum aestivum* L.) based on Distinctness, Uniformity and Stability (DUS) Test. *Journal of Agricultural Technology*, **2**(1-2): 101-106.

Malik, R., Sharma, H., Sharma, I., Kundu, S., Verma, A., Sheoran, S., *et al*. (2014). Genetic diversity of agro-morphological characters in Indian wheat varieties using GT biplot. *African Journal of Crop Science*, **8**(9): 1266-1271.

Malik, S., Rahman, M.U. and Malik, T.A. (2015). Genetic mapping of potential qtls associated with drought tolerance in wheat. *The Journal of Animal and Plant Sciences*, **25**(4): 1032-1040.

Malla, S., Ibrahim, A.M.H., Yen, Y., Glover, K.D. and Berzonsky, W.A. (2011). Association of Fhb1 and Qfhs.ifa-5A in spring versus winter growth habits in Bread Wheat (*Triticum aestivum* L.). *Journal of Agricultural Sciences*, **4**: 39-48.

Malyshev, S.V. and Kartel, N.A. (1997). Molecular markers in mapping of plant genomes. *Molecular Biology*, **31**: 163-171.

Manen, J.F., Bouby, L., Dalnoki, O., Marinval, P., Turgay, M. and Schlumbaum, A. (2003). Microsatellites from archaeological *Vitis vinifera* seeds allow a tentative assignment of the geographical origin of ancient cultivars. *Journal of Archaeological Science*, **30**: 721-729.

Manifesto, M.M., Schlatter, A.R., Hopp, H.E., Suarez, E.Y. and Dubcovsky, J. (2001). Quantitative evaluation of genetic diversity germ plasm using molecular markers. *Crop Science*, **41**: 682-690.

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research*, **27**: 209-220.

Maric, S., Bede, M., Martincic, J. and Guberac, V. (1998). Variability of some winter wheat traits from breeding process. *Sjemenarstvo Seed Science Journal*, **15**: 421-433.

Martin, G.B., Williams, J.G.K. and Tanksley, S.D. (1991). Rapid identification of markers linked to a *Pseudomonas* resistance gene in tomato by using random primers and near isogenic lines. *Proceedings of National Academy of Sciences (USA)*, **88**: 2336-2340.

Martinez, E.A., Destombe, C., quilled, M.C. and Valero, M. (1999). Identification of random amplified polymorphic DNA (RAPD) markers highly linked to sex

determination in the red alga *Gracilaria gracilis*. *Molecular Ecology*, **8**: 1533-1538.

Maryam, G., Ahmad, A., Mohammad, M.M.S.A., Ebrahim, A.T.B. and Abolghasem, M.S. (2009). Detection of SSR markers linked with gene(s) controlling components of yield traits in durum wheat under drought stress and non-stress conditions. Abstract published in proceedings of a conference on 19th International Triticeae mapping Initiative, August 31-September 04, 2009. 3rd COST Teratogen, Clermont-Ferrand, France.

Matus, I.A. and Hayes, P.M. (2002). Genetic diversity in three groups of barley germplasm assessed by simple sequence repeats. *Genome*, **45**: 1095-1106.

Maughan, P.J. Sanghai-Maroof, M.A. Buss, G.R. and Huestis, G.M. (1996). Amplified fragment length polymorphisms (AFLP) in soybean: species diversity, inheritance, and near- isogenic lines analysis. *Theoretical and Applied Genetics*, **93**(3): 392-401.

Maurya, R. (2014). Development and applicationof microsatellite markers for diversity analysis in *Jatropa curcas* L. Ph. D. Thesis (unpublished) submitted to University of Lucknow, Lucknow-226 007, India.

Maxwell, J.J., Lyerly, J.H., Srnic, G., Murphy, J.P., Cowger, C., Parks, R., *et al*. (2012). MlNCD1: A Novel *Aegilops tauschii*-derived powdery mildew resistance gene identified in common wheat. *Crop science*, **52**(3): 1162-1170.

Mba, C., Guimaraes, E.P. and Ghosh, K. (2012). Re-orienting crop improvement for the changing climatic conditions of the 21st century. *Agriculture and Food Security*, **1**: 7.

McMaster, G.S., White, J.W., Hunt, L.A., Jamieson, P.D., Dhillon, S.S. and Ortiz-Monasterio, J.I. (2008). Simulating the influence of vernalization, photoperiod and optimum temperature on wheat developmental rates. *Annals of Botany*, **102**: 561-569.

Meglecz, E., Petenian, F., Danchin, E., Coeur-d'Acier, A., Rasplus, J.Y. and Faure, E. (2004). High similarity between flanking regions of different microsatellites

detected within each of two species of lepidoptera: *Parnassius apollo* and *Euphydryas aurinia*. *Molecular Ecology*, **13**(6): 1693-700.

Melchinger, A.E. (1999). Genetic diversity and heterosis. In: The genetic diversity and exploitation of heterosis in crops. Coors, J.G. (ed). Madison, W153711, USA. **pp**. 99-109.

Melchinger, A.E., Lee, M., Lamkey, K.R., Hallauer, A.R. and Woodman, W.L. (1991). Genetic diversity for restriction fragment length polymorphisms and heterosis for two diallel sets of maize inbreds. *Theoretical and Applied Genetics*, **80**: 488-496.

Mengistu, D.K., Kidane, Y.G., Catellani, M., Frascaroli, E., Fadda, C., Enrico, M.P.E., *et al*. (2016). High-density molecular characterization and association mapping in Ethiopian durum wheat landraces reveals high diversity and potential for wheat breeding. *Plant Biotechnology Journal*, **14**: 1800–1812.

Mengistu, D.K., Kiros, A.Y. and Pe, M.E. (2015). Phenotypic diversity in Ethiopian dururm wheat (*Triticum turgidum* var. durum) landraces. *The Crop Journal*, **3**: 190-199.

Messier, W., Li, S.H. and Stewart, C.B. (1996). The birth of microsatellites. *Nature*, **381**(6582): 483.

Metakovsky, E.V. and Branlard, G. (1998). Genetic diversity of french common wheat germplasm based on gliadin alleles. *Euphytica*, **96**: 209-218.

Meyer, W., Freedman, E.L. and Vilgalys, R. (1993). Hybridization probes for conventional DNA finger printing used as single primers in polymerase chain reaction to distinguish strains of *Cryptococcus neoformans*. *Journal of Clinical Microbiology*, **31**:2274-2280.

Michelmore, R.W., Paran, I. and Kesseli, R.V. (1991). Identification of markers linked to disease resistance genes by bulked sergeant analysis a rapid method to detect markers in specific genomic regions by using segregation populations. *Proceedings of National Academy of Sciences (USA)*, **88**: 9828-9832.

Miesfeld, R., Krystal, M. and Arnheim, N. (1981). A member of a new repeated sequence family which is conserved throughout the eukaryotic evolution is found between the human globin genes. *Nucleic Acids Research*, **9**: 5931-5947.

Miller, J.C. and Tanksley, S.D. (1990). RFLP analysis of phylogenetic relationships and genetic variation in the genus. *Lycopersicon. Theoretical and Applied Genetics*, **80**: 437-448.

Miller, T.E. (1987). Systematics and evolution. In: Wheat breeding: its scientific basis. Lupton, F. G. H (ed.) Chapman and Hell, London. **pp**. 1-30.

Millign, G.W. and Cooper, M.C. (1985). An examination of procedures for determining the number of clusters in a data set. *Psychometrica*, **50**(2): 159-179.

Mir, R.R., Kumar, J., Balyan, H.S. and Gupta, P.K. (2011). A study of genetic diversity among Indian bread wheat (*Triticum aestivum* L.) cultivars released during last 100 years. *Genetic Resources and Crop Evolution*, **59**(5): 717-726.

Mishra, C.N., Tiwari, V., Satish-Kumar, V., Gupta, A., Kumar, A. and Sharma, I. (2015). Genetic diversity and genotype by trait analysis for agromorphological and phhysiological traits of wheat (*Triticum aestivum*). *SABRAO Journal of Genetics and Plant Breeding*, **47**(1): 40-48.

Moeykens, C.A., Mackenzie, S.A. and Shoemaker, R.C. (1995). Mitochondrial genome diversity in soybean: Repeats and rearrangements. *Plant Molecular Biology*, **29**(2): 245-254.

Mohammadi, R. and Amri, A. (2011). Graphic analysis of trait relations and genotype evaluation in durum wheat. *Journal of Crop Improvement*, **25**(6): 680-696.

Mohammadi, S.A. and Prasanna, B.M. (2003). Analysis of genetic diversity in crop plants-salient statistical tools and considerations. *Crop Sciences*, **43**: 1235-1248.

Moragues, M., Garcia, del Moral, L.F., Moralejo, M. and Royo, C. (2005). Yield formation strategies of durum wheat landraces with distinct pattern of dispersal with in the mediterranean basin. *Field Crop Research*, **95**: 194-205.

Morgante, M. and Olivieri, A.M. (1993). PCR amplified microsatellites as markers in plant genetics. *Plant Journal*, **3**: 175-182.

Morgante, M. and Vogel, J. (1994). Compound microsatellite primers for the detection of genetic polymorphisms. US Patent Application, 08/326456.

Morgante, M., Rafalski, A., Biddle, P., Tingey, S. and Olivieri, A.M. (1994). Genetic mapping and variability of seven *soybean* simple sequence repeat loci. *Genome*, **37**: 763-769.

Morgante, M., Hanafey, M. and Powell, W. (2002). Microsatellites are preferentially associated with no repetitive DNA in plant genomes. *Nature Genetics*, **30**: 194-200.

Mori, N., Liu, Y.G. and Tsunewaki, K. (1995). Wheat phylogeny determined by RFLP analysis of nuclear DNA 2. Wild tetraploid wheats. *Theoretical and Applied Genetics*, **90**: 129-134.

Mori, N., Miyashita, N.T., Terachi, T. and Nakamura, C. (1997). Variation in *coxll* intron in the wild ancestral species of wheat. *Hereditas*, **126**: 281-288.

Moseman, J.G. and Craddock, J.C. (1976). Genetic basis for collecting, evaluating and maintaining bartley germplasm. In: Barley Genetics 111. Gaul, H. (ed.). Proceedings of International Barley Genetics Symposium. **pp**. 51-57.

Mujeeb-Kazi, A., Gul, A., Farooq, M., Rizwan, S. and Ahmad, I. (2008). Rebirth of synthetic hexaploid with global implications for wheat improvement. *Australian Journal of Agricultural Research*, **59**: 391-398.

Mukai, Y. and Gill, B.S. (1991). Detection of barley chromatin added to wheat by genomic *in situ* hybridization. *Genome*, **34**: 448-452.

Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G. and Erlich, H. (1986). Specific enzymatic amplification of DNA *in vitro*: The polymerase chain reaction. *Cold Spring Harbor Symposia on Quantitative Biology*, **51**: 263-273.

Munck, L., Karlson, K.E., Hagberg, W. and Eggum (1970). Gene for improved nutritional value in barley seed protein. *Science*, **168**: 985-987.

Murray, H.G. and Thompson, W.F. (1980). Rapid isolation of high molecular weight DNA. *Nucleic Acids Research*, **8**: 4321-4325.

Mwale, V.M., Tang, X. and Chilembwe, E. (2016). Assessment of genetic diversity among sixty bread wheat (*Triticum aestivum*) cultivars using microsatellite markers. *African Journal of Biotechnology*, **15**(21): 960-973.

Nadir, E., Margalit, H., Gallily, T. and Ben-Sasson, S.A. (1996). Microsatellite spreading in the human genome: evolu-tionary mechanisms and structural implications. *Proceedings of the National Academy of Sciences,* (*USA),* **93**: 6470-6475.

Nagaoka, T. and Ogihara, Y. (1997). Applicability of inter-simple sequence repeat olymorphisms in wheat for use as DNA markers in comparison to RFLP and RAPD markers. *Theoretical and Applied Genetics*, **94**: 597–602.

Najaphy, A., Parchin, R.A. and Farshadfar, E. (2012). Comparison of phenotypic and molecular characterizations of some important wheat cultivars and advanced breeding lines. *Australian Journal of Crop Sciences*, **6**(2): 326-332.

Nakamura. Y., Leppert, M., O'Connell, P., Wolff, R., Holm, T., Culver, M., *et al*. (1987). Variable number of tandem repeat (VNTR) markers for human gene mapping. *Science*, **235**: 1616-1622

Naranjo, T., Roca, P., Goicoechea, P.G. and Giraldez, R. (1987). Arm homoeology of wheat and rye chromosomes. *Genome*, **29**: 873–882.

Nelson, J.C., Sorrells, M.E., Devnze, A.E.V., Lai, Y.H., Alkinson, M., Bernard, M., *et al*. (1995a). Molecular mapping of wheat major genes and rearrangements in homologous groups 4, 5, and 7. *Genetics*, **141**: 721-731.

Nelson, J.C., Devyzne, A.E.V., Antique, E., Sorrells, M.E., Lu, Y.H., Negre, S., *et al*. (1995b). Molecular mapping of wheat homeologous group 3. *Genome*, **38**: 525-533.

Nesbitt, M. and Samuel, D. (1998). Wheat domestication: archaeobotanical evidence. *Science*, 279:1431.

Nevo, E. and Payne, P.I. (1987). Wheat storage protein diversity of HWM glutenin subunits in wild emmer from Isreal. 1. Geographical patterns and eco geographical predictability. *Theoretical and Applied Genetics*, **74**: 827–36.

Nevo, E., Johary, D., Brown, A.H.D. and Haber, M. (1986). Genetic diversity and environmental associations of wild barley, *Hordeum spontaneum* in Israel. *Evolution*, **33**: 815-835.

Nevo, E., Korol, B., Lerles, A. and Fahima, T. (2002). Evolution of wild emmer and wheat improvement. Springer-Verlag, Heidelberg GmbH, Berlin.

Newbury, H.J. and Ford-Lloyd, B.V. (1997). Estimation of genetic diversity. In: Plant genetic conservation: The *in situ* approach. Maxted, N., Ford-Lloyd, B.V. and Hawkers, J.G. (eds.). Chapman and all, London. **pp**: 192-206.

Nielsen, D.M. and Weir, B.S. (1999). A classical setting for associations between markers and loci affecting quantitative traits. *Genetic Research*, **74**: 271-277.

Nielsen, N.H., Backes, G., Stougaard, J., Andersen, S.T. and Jahoor, A. (2014). Genetic Diversity and Population Structure Analysis of European Hexaploid Bread Wheat (*Triticum aestivum* L.) Varieties. *PloS One*, **9**(4): e94000.

Nikiforov, T.T., Rendle, R.B., Goelet, P., Rogers, Y.H., Kotewicz, M.L., Anderson, S., Trainor, G.L. and Knapp, M.R. (1994). Genetic bit analysis a solid phase method for typing single nucleotide polymorphisms. *Nucleic Acids Research*, **22**: 4167-4175.

Oliveira, E.J., Pdua, J.G., Zucchi, M.I., Vencovsky, R. and Vieira, M.L.C. (2006). Origin, evolution and genome distribution of microsatellites. *Genetics and Molecular Biology*, **29**: 294-307.

Orita, M., Suzuki, Y., Seklya, T. and Hayashi, K. (1989). Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. *Genomes*, **5**: 874-877.

Pagnotta, M., Mondini, L. and Atallah, M. (2005). Morphological and molecular characterization of Italian emmer wheat accessions. *Euphytica*, **146**: 29-37.

Palacios, C., Kresovich, S. and Gonzales-Candelas, F. (1999). A population genetic study of the endangered plant species *Limonium dufourii* (Plumbaginaceae) based on amplified fragment length polymorphism (AFLP). *Molecular Ecology*, **8**: 645-657.

Paran, I. and Michelmore, R.W. (1993). Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. *Theoretical and Applied Genetics*, **85**: 985-933.

Parasnis, A.S., Gupta, V.S., Tamhankar, S.A. and Ranjekar, P.K. (2000). A highly reliable sex diagnostic PCR assay for mass screening of papaya seedlings. *Molecular Breeding*, **6**: 337-344.

Pardue, M.L. and Gall, J.G. (1970). Chromosomal localization of mouse satellite DNA. *Science*, **168**: 1356-1358.

Park, Y.J., Ju, K.L. and Nam-Soo, K. (2009). Simple sequence repeats polymorphisms (SSRPs) for evaluation of molecular diversity and germplasm classification of minor crops. *Molecules*, **14**: 4546-4569.

Parker, G.D., Chalmers, K.J., Rathjen, A.J. and Langridge, P. (1998). Mapping loci associated with flour colour in wheat *(Triticum aestivum* L.). *Theoretical and Applied Genetics*, **97**: 238-245.

Paterson, A.H., Lim, Y.R., Li, X.K., Schertz, K.F., Doebley, J.F., Pinson, S.R.M., *et al*. (1995). Convergent domestication of cereal crops by independent mutations at corresponding genetic loci. *Science*, **269**: 1714-1718.

Pawar, V. and Patil, R.B. (1989). Variability and heritance of yield components in crosses of wheat. *Journal of Maharashtra Agricultural University*, **14**(2): 143-145.

Peakall, R. and Smouse, P.E. (2012). GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research: An update. *Bioinformatics*, **28**: 2537–2539.

Pearson, C.E., Edamura, N.K. and Cleary, J.D. (2005). Repeat instability: Mechanisms of dynamic mutations. *Nature Reviews Genetics*, **6**: 729-742.

Peeters, J.P. and Martinelli, J.A. (1989). Hierarchical cluster analysis as a tool to manage variation in germplasm collections. *Theoretical and Applied Genetics*, **78**: 42-48.

Peltonen-Sainioa, P., Kangas, A., Salo, Y. and Jauhiainen, L. (2007). Grain number dominates grain weight in temperate cereal yield determination: Evidence based on 30 years of multilocation trials. *Field Crop Research*, **100**: 179-188.

Plaschke, J., Ganal, M.W. and Röder, M.S. (1995). Detection of genetic diversity in closely related bread wheat using microsatellite markers. *Theoretical and Applied Genetics,* **91**: 1001–1007.

Polignano, L. and SpagnolettiZeuli (1985). Polymorphisms of human DNA by gel electrophoresis as single - strand conformation polymorphisms. *Proceedings of National Academy of Sciences (USA)*, **86**: 2766-2770.

Porceddu, E. (1976). Variation for agronomical traits in a world collection of durum wheat. Z. *Pflanzenzuchtg*, **77**: 314-329.

Powell, W., Machray, G. and Provan, J. (1996). Polymorphism revealed by simple sequence repeats. *Trends in Plant Sciences*, **1**: 215-222.

Prevosti, A., Ocana, J. and Alonzo, G. (1975). Distances between populations for Drosophila Subobscura based on chromosome arrangement frequencies. *Theoretical and Applied Genetics*, **45**: 231-241.

Pritchard, J.K., Stephens, M. and Donelly, P. (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**: 945-959.

Provan, J., Powell, W. and Hollingsworth, P.M. (2001). Chloroplast microsatellites: new tools for studies in plant ecology and evolution. *Trends in Ecology and Evolution,* **16**: 142–147.

Pujar, S., Tamhankar, S.A., Gupta, V.S., Rao, V.S. and Ranjekar, P.K. (2002). Diversity analysis of Indian tetraploid wheat using intersimple sequence repeat markers reveals their superiority over random amplified polymorphic DNA Markers. *Biochemical Genetics*, **40**: 63-69.

Pujar, S., Tamhankar, S.A., Rao, V.S., Gupta, V.S., Naik, S. and Ranjekar, P.K. (1999). Arbitrarily primed-PCR based diversity assessment reflects hierarchical groupings of Indian tetraploid wheat genotypes. *Theoretical and Applied Genetics*, **99**: 868–876.

Qi Lili, Benjamin Echalier, Bernd Friebe, Bikram, S. and Gill. (2003). Molecular characterization of a set of wheat deletion stocks for use in chromosome bin mapping of ESTs. *Functional and Integrated Genomics*, **3**: 39-55.

Qi, L.L., Echalier, B., Lazo, C., Butler, S.G.R., Anderson, G.E.O.D., Akhunov, E.D., *et al*. (2004). A chromosome bin map of 16,000 expressed sequence tag loci and distribution of genes among the three genomes of polyploidy wheat. *Genetics*, **168**: 701–712.

Qualset, C.O. and Puri, Y.P. (1947). Sources of earliness and winter habit in durum wheat. *Wheat Information Series*, **38**: 13-15.

Qualset, C.O. and Puri, Y.P. (1974). Heading time in world collection of durum wheat photo and thermal sensitivity related to latitudinal responses and geographical origins. In: International Symposium on Genetics and Breeding of Durum Wheat. Scarascia Mugnozza (ed). **pp**. 166-178.

Quarrie, S.A., Quarrie, S.P., Radosevic, R., Rancic, D., Kaminska, A., Barnes, J.D., *et al*. (2006). Dissecting wheat QTL for yield present in a range of environments: from the QTL to candidate genes. *Journal of Experimental Botany*, **57**(11): 2627-2637.

Quarrie, S.A., Steed, A., Calestani, C., Semikhodskii, A., Lebreton, C., Chinoy C., *et al*. (2005). A high-density genetic map of hexaploid wheat (*Triticum aestivum* L.) from the cross Chinese Spring X SQ1 and its use to compare QTLs for grain yield across a range of environments. *Theoretical and Applied Genetics*, **110**: 865-880.

Rachovska, G., Dimova, D. and Bojinov, B. (2002). Application of cluster analysis and principal component analysis for evaluation of common winter wheat genotypes. *Proceedings of the scientific session of jubilee*, **3**: 68-72.

Rafalski, J.A. and Tingey, S.V. (1993). Genetic diagnostic in plant breeding, RAPDs, microsatellites and machines. *Trends in Genetics*, **9**: 275-280.

Rafalski, J.A. and Vogel (1996). RAPD markers a new technology for genetic molecular mapping and Plant breeding. *Agricicultural Biotechnology News Information*. **pp**. 3545-648.

Rahman, M., Barma, N.C.D., Biswas, B.K., Khan, A.A. and Rahman, J. (2016). Study on morpho-physiological traits in spring wheat (*Triticum aestivum* L.) under rain fed condition. *Bangladesh Journal of Agricultural Research*, **41**(2): 235-250.

Raju, A., Dholakia, J.S.S., Santra, B.B., Singh, D.K., Lagu, H., Tamhankar, M.D., *et al*. (2001). Identification of inter-simple sequence repeat (ISSR) markers associated with seed size in wheat. *Theoretical and Applied Genetics*, **102**: 726-732.

Ranjekar, P.K., Pallota, D. and Lafontaine, J.G. (1976). Analysis of the genome of plants II. Characterization of repetitive DNA in barley (*Hordeum vulgare*) and what (*Triticum aestivum*). *Acta Biochemistry and Biophysics*, **425**: 30-40.

Ranjekar, P.K., Pallota, D. and Lafontaine, J.G. (1978). Analysis of plant genomes III. Denaturation and reassociation properties of cryptic satellite DNAs in barley (*Hordeum vulgare*) and wheat (*Triticum aestivum*). *Acta Biochemistry and Biophysics*, **520**: 103-110.

Rao, A.C., Maya, M.M., Duque, M.C., Tohme, J., Allem, A.C. and Bonierable, M.W. (1997). AFLPs analysis of relationship among cassava other *Manihot* species. *Theoretical and Applied Genetics*, **95**: 741-750.

Rao, V.R. and Riley, K.W. (1994). The use of biotechnology for conservation and utilization of plant genetic resources. *Plant Genetic Resources Newsletter*, **97**: 3-20.

Reif, J.C., Zhang, P., Dreisigacker, S., Warburton, M.L., Ginkel, M., Van-Hoisington, D., *et al*. (2005). Wheat genetic diversity trends during domestication and breeding. *Theoretical and Applied Genetics*, **110**: 859–864.

Ren, J., Sun, D., Chen, L., You, F.M., Wang, J., Peng, Y., *et al*. (2013). Genetic diversity revealed by single nucleotide polymorphism markers in a worldwide germplasm collection of durum wheat. *International Journal of Molecular Sciences*, **14**: 7061-7088.

Ren,Y., Li, Z., He, Z., Wu, L., Bai, B., Lan, C., *et al*. (2012). QTL mapping of adult-plant resistances to stripe rust and leaf rust in Chinese wheat cultivar Bainong 64. *Theoretical and Applied Genetics*, **125**(6): 1253-1262.

Reynolds, M.P., Trethowan, R., Crossa, J., Vargas, M. and Dayre, K.D. (2002). Physiological factors associated with genotype by environment interaction in wheat. *Field Crop Research*, **75**: 139-160.

Richard, G.F. And Paques, F. (2000). Mini- and microsatellite expansions: the recombination connection. *European Molecular Biology Organization Reports*, **11**: 122-126.

Richardson, T., Cato, S., Ramser, J., Kahl, G. and Weising, K. (1995). Hybridization of microsatellite to RAPD: a new source of polymorphic marker. *Nucleic Acids Research*, **23**(18): 3798-3799.

Riley, R. and Chapman, V. (1958). Genetic control of the cytologically diploid behavior of hexaploid wheat. *Nature*, **182**: 713-715.

Riley, R., Unrau, J. and Chapman, V. (1958). Evidence on the origin of the B genome of wheat. *Journal of Heredity*, **49**: 91-98.

Robert, T., Lexpinasse, R., Pemess, J. and Sarr, A. (1991). Gametophyte competition as influencing gene flow between wild and cultivated forms of pearl millet (*Pennisetum typhoides*). *Genome*, **34**: 195-200.

Roder, M.S., Korzun, V., Wendehake, K., Plaschke, J., Tixier, M.H., Leroy, *et al*. (1998). A microsatellite map of wheat. *Genetics*, **149**: 2007–2023.

Roder, M.S., Plaschke, J., Koing, S.U., Borner, A., Sorrells, M.E., Tanksley, S.D., *et al*. (1995). Abundance variability and chromosomal location of microsatellite was in wheat. *Molecular Genomes and Genetics*, **246**: 327-333.

Rohlf, F.J. (2000). NTSYS-PC. Numerical taxonomy and multivariate analysis system, version 2.02e. Exeter Software, New York.

Rongwen, J., Akkaya, M.S., Bhagawat, A.A., Lavi, U. and Cregan, P.V. (1995.) The use of microsatellite DNA markers for soyabean genotype identification. *Theoretical and Applied Genetics*, **90**: 43-48.

Ronning, C. M., Stegalkina, S.S., Ascenzi, R.A., Bougri, O., Hart, A.L., Utterbach, T. R., Vanaken, S. E., Riedmuller, S. B., White, J. A. and Cho, J. (2003). Comparative analyses of potato expressed sequence tag libraries. *Plant Physiology*, **131**: 419-429.

Rose, O. and Falush, D. (1998). A threshold size for microsatellite expansion. *Molecular Biology and Evolution*, **15**: 613-615.

Rosenbaum, V. and Riesner, D. (1987). Temperature-gradient gel electrophoresis: Thermodynamic analysis of nucleic acids and proteins in purified form and in cellular extracts. *Biophysical Chemistry*, **26**(2-3): 235-246.

Roussel, V., Koening, J., Beckert, M. and Balfourier, F. (2004). Molecular diversity in French bread wheat accessions related to temporal trends and breeding programs. *Theoretical and Applied Genetics*, **108**(5): 920-930.

Routray, P., Basha, O., Garg, M., Singh, N.K. and Dhalikal H.S. (2007). Genetic diversity of landraces of wheat (*Triticum aestivum* L.) from hilly areas of Uttaranchal, India. *Genetic Resources and Crop Evolution*, **54**: 1315-1326.

Russel, J.R. (2000). A retrospective analyses of spring barley germplasm development from 'foundation genotypes' to currently successful cultivars. *Molecular Breeding*, **6**: 553-568.

Sabaghnia, N., Janmohammadi, M. and Segherloo, A.B. (2014). Evaluation of some agro-morphological traits diversity in Iranian bread wheat genotypes. *Annales Universitatis Mariae Curie-sklodovska*, **69**(1): 79-92.

Saiki, R.K., Gelfand, D.H., Stoffei, S., Scharf, S.J., Higuchi, R., Horn, G.T., *et al*. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Nature*, **239**: 487-497.

Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A., *et al*. (1985). Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science*, **230**: 1350–54.

Sakamura, T. (1918). Kurze Mitterlund uber die chromosomenzahlen und dle Verwandschftsverhalnisse der. *Triticum-Arten. Botanical Magazine (Tokyo)*, **32**: 150-153.

Saker, M., Nagchtigall, M. and Kuehne, T.A. (2005). Comparative assessment of DNA fingerprinting by RAPD, SSR and AFLP in genetic analysis of some barley genotypes. *Egyptian Journal of Genetics and Cytogenetics*, **34**: 81–97.

Salehi, S., Gholami, S., Menati, M., Soraya, N. and Parvin, T. (2016). Effects of morphological traits on qualitative and quantitative yield of bread wheat (*Triticum aestivum* L.) cultivars. *Biosciences Biotechnology Research Asia*, **13**(2): 711-714.

Salem, K.F.M., El-Zanaty, A.M. and Esmail, R.M. (2008). Assessing Wheat (*Triticum aestivum* L.) Genetic diversity using morphological characters and microsatallite markers. *World Journal of Agricultural Sciences*, **4**(5): 538-544.

Sandhu, D., Champoux, J.S., Bondareva, S.N. and Gill, K.S. (2001). Identification and physical localization of useful genes and markers to a major gene-rich region on wheat group 1S chromosomes. *Genetics*, **157**: 1735-1747.

Sardouie-Nasab, S., Nejad, G.H.M. and Nakhoda, B. (2013). Assessing genetic diversity of promising wheat (*Triticum aestivum L.*) lines using microsatellite markers linked with salinity tolerance. *Journal of Plant Molecular Breeding*, **1**(2): 28-39.

Sarfraz, Z., Shah, M.M. and Iqbal, M.S. (2016). Estimation of components of variation of morphological traits and yield in A-genome wheat populations. *SABRAO Journal of Plant Breeding and Genetics*, **48**(2): 105-109.

Sarkar, G., Cassady, J., Bottema, C.D.K. and Sommer, S.S. (1990). Characterisation of polymerase chain reaction amplification of specific alleles. *Analytical Biochemistry*, **186**: 64-68.

SAS Institute Inc. 2011. Base sas©9.3 procedures guide. Cary, NC, USA.

Scarascia-Mugnozza, G.T. and Porceddu, E. (1974). Genetics studies on coleoptile length of durum wheat. In Sears E.R. Sears M.R. (Eds.) Proceedings of. 4th wheat genetics symposium, Columbia. **pp**. 301-308.

Schlegal, R. and Mettin, D. (1978). Bemerkungen zur Karyotypanalyse beim Saatweizen (*Triticum aestivum* L.). *Archeologica Zuechtungsforsch* (English), **8**: 333-344.

Schlotterer, C. (1998). Microsatellites. In: Molecular genetic analysis of populations: A practical approach. Hoelzel, A.R., (ed.). IRL Press, Oxford. **pp**. 237-261.

Schlotterer, C. (2004). The evolution of molecular markers- just a matter of fashion? *Genetics*, **5**: 63-6.

Schoen, D.J. and Brown, A.H.D. (1991). Intraspecific variation in population gene diversity and effective population size correlates with the mating system in plants. *Proceedings of National Academy of Sciences (USA)*, **88**: 4494-4497.

Schuler, S.F., Bacon, R.K. and Edward, E.G. (1994). Kernel and spike character influence on test weight of soft red winter wheat. *Crop Science*, **34**: 1309-1313.

Schut, J.W., Qi, X. and Stam, P. (1997). Association between relationship measures based on AFLP markers, pedigree data and morphological traits in barley. *Theoretical and Applied Genetics*, **95**: 1161-1168.

Schwarzacher, T. (1996). The physical organization of Triticeae chromosomes. *Journal of Experimental Biology*, **7**: 71-75.

Schwarzacher, T., Anamthawat-Johnson, K., Harrson, Q.E., Islam, A.K.M.R., Jia, J.S., King, I.P., *et al*. (1992). Genomic *in situ* hybridization to identify alien chromosomes and chromosome segments in wheat. *Theoretical and Applied Genetics*, **84**: 779-786.

Scott, K.D., Eggler, P., Seaton, G., Rossetto, M., Ablett, E.M., Lee, L.S., *et al*. (2000). Analysis of SSRs derived from grape ESTs. *Theoretical and Applied Genetics*, **100**: 723-726.

Sehgal, S.A., Tahir, R.A. and Nawaz, M. (2012). Molecular characterization of wheat genotypes using SSR markers. *International Journal of Bio automation*, **16**(2): 119-128.

Semagn, K., Bjornstad, A. and Ndjiondjop, M.N. (2006). An overview of molecular marker methods for plants. *African Journal of Biotechnology*, **5**: 2540-2568.

Sharma, A. and Chauhan, R.S. (2011). Repertoire of SSRs in the castor bean genome and their utilization in genetic diversity analysis in *Jatropha curcas*. *Comparative and Functional Genomics*, **2011**: Article ID 286089.

Sharma, H.K., Sarkarb, M., Choudharya, S.B., Kumara, A.A., Maruthi, R.T., Mitra, J., *et al*. (2016). Diversity analysis based on agromorphological traits and microsatellite-based markers in global germplasm collections of roselle (*Hibiscus sabdariffa* L.). *Industrial Crops and Products*, **89**: 303-315.

Sharma, R. (2016). Molecular characterization of wheat Varieties suitable for plains of Jammu. M. Sc. Thesis (unpublished) submitted to Sher-e-Kashmir University of Agricultural Sciences & Technology of Jammu, Jammu-180 009, India.

Sharma, S.K. Knox, M.R. and Ellis, T.H.N. (1996). AFLP analysis of the diversity and phylogeny of *Lens* and its comparison with RAPD analysis. *Theoretical and Applied Genetics*, **93**(5-6): 751-758.

Shewry, P. R. (2009). Wheat. *Journal of Experimental Botany*, **60**(6)**:** 1537–1553.

Shi, G.J., Friesen, T.L., Saini, J., Xu, S.S., Rasmussen, J.B. and Faris, J.D. (2015). The Wheat *Snn7* Gene Confers Susceptibility on Recognition of the *Parastagonospora nodorum* Necrotrophic Effector SnTox7. *Plant Genome*, **8**(2): 1-10.

Shiferaw, B., Smale, M., Braun, H.J., Duveiller, E., Reynolds, M. and Muricho, G. (2013). Crops that feed the world 10. Past successes and future challenges to the role played by wheat in global food security. *Food Security*, **5**: 291–317.

Shin, J. H., Kwon, S. J., Lee, J. K., Min, H. K. and Kim, N.S. (2006). Genetic diversity of maize kernel starch-synthesis genes with SNAPs. *Genome*, **49**: 1287–1296.

Simmonds, N.W. (Eds.), (1976). Evolution of crop plants. Longman, London.

Singh, K.B., Ocampo, B. and Robertson, L.D. (1999). Diversity for abiotic and biotic stress resistance in wild annual *Cicer* species. *Genetic Resources and Crop Evolution*, **45**: 9-17.

Singh, N.K., Dalal, V., Batra, K., Singh, B.K., Chitra, G., Singh, A., *et al*. (2007). Single-copy genes define a conserved order between rice and wheat for understanding differences caused by duplication, deletion, and transposition of genes. *Functional and Integrated Genomics*, **7**: 17–35.

Singh, R.P., Hodson, D.P., Huerta-Espino, J., Jin, Y., Bhavani, S., Njau, P., *et al*. (2011). The emergence of *Ug99* races of the stem rust fungus is a threat to world wheat production. *Annual Review of Phytopathology*, **49**: 465–481.

Singh, R.P., Hodson, D.P., Huerta-Espino, J., Jin, Y., Njau, P., Wanyera, R., *et al*. (2008). Will stem rust destroy the world's wheat crop? *Advances in Agronomy*, **98**: 271–309.

Sinha, S. and Kumaravadivel, K. (2016). Understanding genetic diversity of sorghum using quantitative traits. *Scientifica* (http://dx.doi.org/10.1155/2016/3075023, accessed on May 28, 2016).

Sivaramakrishnan, S., Seetha, K., Kamala, V., Reddy, L.J., Rai, K.N., Reddy, B.V.S., *et al*. (2001). RFLP analysis of mitochondrial DNA for the identification of cytoplasmic male sterile systems and detection of genetic variability in Pearl millet, Sorghum and Pigeon pea. Plant and Animal Genome VIII Conference, January 9-12, 2001. San Deigo, CA,.USA.

Slager, S.L. and Schaid, D.J. (2001). Evaluation of candidate genes in case-control studies: A statistical method to account for related subjects. *American Journal of Human Genetics*, **68**: 1457-1462.

Slavov, G.T., Howe, G.T., Gyaourova, A.V., Birkes, D.S. and Adams, W.T. (2005). Estimating pollen flow using SSR markers and paternity exclusion: Accounting for mistyping. *Molecular Ecology*, **14**: 3109-3121.

Smale, M.I., Ortiz-Monsterio, M., Warburton, B., Skovmand, M., Reynolds, J., Crossa, R., *et al*. (2002). Dimensions of diversity in modern spring bread wheat in developing countries from 1965. *Crop Sciences*, **42**: 1766–79.

Smith, O.S., Smith, J.S.C., Brown, S.L., Tenborg, R.A. and Wall, S.J. (1991). Similarities among a group of elite maize inbreeds as measured by pedigree, F, grain yeild, heterosis, and RFLPs. *Theoretical and Applied Genetics*, **80**: 833-840.

Somers, D.J., Isaac, P. and Edwards, K. (2004). A high-density microsatellite consensus map for bread wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics*, **109**: 1105–1114.

Somers, D.J., Zhou, Z., Bebeli, P. and Gustafson, J.P. (1996). Repetitive genome specific probes in wheat (*Triticum aestivum* L.) amplified with minisatellite core sequences. *Theoretical and Applied Genetics*, **93**: 983-989.

Song, Q.J., Shi, J.R., Singh, S., Fickus, E.W., Costa, J.M., Lewis, J., *et al*. (2005). Development and mapping of microsatellite (SSR) markers in wheat. *Theoretical and Applied Genetics*, **110**: 550–560.

Soriano, J.M., Villegas, D., Aranzana, M.J., Moral, L.F.G. and Royo, C. (2016). Genetic structure of modern durum wheat cultivars and Mediterranean landraces matches with their agronomic performance. *PloS One*, **11**(8): e0160983.

Southern, E.M. (1975). Deletion of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology*, **96**: 503-517.

Souza, C.M.P., Viana, A.P., Ferreiram C.F., Silva, S.O., Carvalho, A.J.C., Berbert, P.A. and Souza, E.F. (2008). Evaluation of genetic dissimilarity in banana genotypes (*Musa* spp.) *via* RAPD markers. *Revista Brasileira de Fruticultura*, **30**: 419-424.

Spanic, V., Buerstmayr, H. and Drezner, G. (2012). Assessment of genetic diversity of wheat genotypes using microsatellite markers. *Periodicum Biologorum*, **114**(1): 37-42.

Spritz, R.A. (1981). Duplication/deletion polymorphism 5'- to the human beta globin gene. *Nucleic Acids Research*, **9**: 5037-5047.

Stansfield, W.D. (1986). Theory and problems of genetics. McGraw-Hill Book Company, New York. Stein, N., Prasad, M., Scholz, U., Thiel, T., Zhang, H., Wolf, M., *et al*. (2007). A 1,000-loci transcript map of the barley genome: new anchoring points for integrative grass genomics. *Theoretical and Applied Genetics*, **114**: 823-839.

Stebbins, G.L. (1956). Taxonomy and evolution of genera with special reference to the family Gramineae. *Evolution*, **10**: 35-36.

Stephan, W. and Kim, Y. (1998). Persistence of microsatellite arrays in finite populations. *Molecular Biology and Evolution*, **15**: 1332-1336.

Stuber, S.A., Krishna, T.G. and Bhatia, C.R. (1992). Purification and characterization of protease inhibitors from pigeonpea *(Cajonus cajan* L. Millsp.) *Journal of Food Science and Agriculture*, **64**: 331-335.

Subirana, J. and Messeguer (2008). Structural families of genomic microsatellites. *Gene*, **408**: 124-132.

Sudre, C., Leonardecz, E., Rodrigues, R. and Junior, A. (2007). Genetic resources of vegetable crops: a survey in the Brazilian germplasm collections pictured through papers published in the journals of the Brazilian Society for Horticultural Science. *Horticultura Brasileira*, **25**: 496-503.

Sun, J.X., Mullikin, J.C., Patterson, N. and Reich, D.E. (2009). Microsatellites are molecular clocks that support accurate inferences about history. *Molecular Biology and Evolution*, **26**: 1017-1027.

Sun, Q., Ni, Z., Litu, Z., Gao, J. and Huang, T. (1998). Genetic relationships and diversity among Tibetian wheats, common wheat and European spelt wheat revealed by RAPD markers. *Euphytica*, **99**: 205-211.

Surwenshi, A., Kumar, V., Shanwad, U.K. and Jalageri, B.R. (2011). Critical review of diversity in *Jatropha curcas* for crop improvement: A candidate biodiesel crop. *Research Journal of Agricultural Sciences*, **2**: 193-198.

Suvarna, T.T. (2001). Molecular analysis of wheat genome using ISSR and RAPD markers. M.Sc. thesis (unpublished), submitted to University of Pune, Pune-411008, State of Maharashtra, India.

Szalma, S.J., Buckler, E.S., Snook, M.E. and McMullen, M.D. (2005). Association analysis of candidate genes for maysin and chlorogenic acid accumulation in maize silks. *Theoretical and Applied Genetics*, **110**: 1324-1333.

Tamura, K., Dudley, J., Nei, M. and Kumar, S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, **24**:1596-1599.

Tang, S. and Knapp, S.J. (2003). Microsatellites uncover extraordinary diversity in native American land races and wild populations of cultivated sunflower. *Theoretical and Applied Genetics*, **106**: 990–1003.

Tang, Z.X., Fu, S.L., Ren, Z.L. and Zou, Y.T. (2009). Rapid evolution of simple sequence repeat induced by allopolyploidization. *Journal Molecular Evolution*, **69**: 217. doi:10.1007/s00239-009-9261-2.

Tanto, T. and Demissie (1996). A comparative genetic diversity study for four major crops managed under Ethiopian condition, Institute of Biodiversity Conservation and Research; Ethiopia: **pp**. 1-9 online accessed on July 15.2017 (URL http://archive.unu.edu/env/plec/cbd/Montreal/papers/Tanto.pdf)

Tar, M., Purnhauser and Csosz, M. (2008). Identification and localization of molecular markers linked to the *Lr52* leaf rust resistance gene of wheat. *Cereal Research (Common)*, **36**(3): 409–415.

Tautz, D. (1989). Hyper variability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Research*, **17**: 6463-6471.

Tautz, D. and Renz, M. (1984). Simple sequence repeats are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acid Research*, **12**: 4127-4138.

Tautz, D. and Schlotterer, C. (1994). Simple Sequences. *Current Opinion in Genetics Development*, **4**: 832-837.

Temnykh, S., DeClerck, G., Lukashova, A., Lipovich, L., Cartinhour, S. and McCouch, S. (2001). Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): Frequency, length variation, transposon associations, and genetic marker potential. *Genome Research*, **11**: 1441-1452.

Tesfaye, T., Getachew, B. and Worde, M. (2008). Morphological diversity in tetraploid wheat landrace populations from central highlands of Ethiopia. *Heredity*, **114**(2): 171-176.

Thiel, T., Michalek, W., Varshney, R.K. and Graner, A. (2003). Exploiting EST databases for the development of cDNA derived microsatellite markers in barley (*Hordeum vulgare* L.). *Theoretical and Applied Genetics*, **106**: 411-422.

Thoren, P.A., Paxton, R.J. and Estoup, A. (1995). Unusually high frequency of (CT)n and (GT)n microsatellite loci in a yellow jacket wasp, *Vespula rufa* (L.) (Hymenoptera: Vespidae). *Insect Molecular Biology*, **4**: 141-148.

Thuillet, A.C., Bru, D., David, J., Roumet, P., Santomi, S., Sourdille, P., *et al*. (2002). Direct estimation of mutation rate for 10 microsatellite loci in durum wheat, *Triticum turgidum* (L.) Thell. ssp durum desf. *Molecular Biology and Evolution*, **19**: 122-125.

Tilahun, D., Shiferaw, E., Johansson, E. and Hailu, F. (2015). Genetic variability of Ethiopian bread wheat genotypes (*Triticum aestivum* L.) using agromorphological traits and their gliadin content. *African Journal of Agricultural Research*, **11**(5): 330-339.

Toth, G., Gaspari, Z. and Jurka, J. (2000). Microsatellites in different eukaryotic genomes: survey and analysis. *Genome Research*, **10**: 967-981.

Tozuka, A., Fukushi, H., Hirata, T., Ohara, M., Kanazawa, A., Mikami, T., Abe, J. and Shimamoto, Y. (1998). Composite and clinal distribution of Glycine soja in Japan revealed by RFLP analysis of mitochondrial DNA. *Theoretical and Applied Genetics*. **96**: 170-176.

Trethowan, R.M. and Mujeeb-Kazi, A. (2008) Novel germplasm resources for improving environmental stress tolerance of hexaploid wheat. *Crop Science*, **48**: 1255-1265.

Trkulja, D., Kondic-Spika, A., Brbaklić, L., Kobiljski, B. and Hristov, N. (2012). Marker-trait associations for spike-related characters in a doubled haploid population of wheat. *Romanian Journal of Agricultural Research*, **29**: 9-16.

Van-Deynze, A.E., Dubcovsky, J., Gill, K.S., Nelson, J.C., Sorrels, M.E., Dvorak, J., *et al*. (1995). Molecular genetic maps for group 1 chromosomes of *Triticeae* species and their relation to chromosomes in rice and oat. *Genome*, **38**: 45-59.

Varshney, R.K., Chabane, K., Hendre, P.S., Aggarwal, R.K. and Graner, A. (2007). Comparative assessment of EST-SSR, EST-SNP and AFLP markers for evaluation of genetic diversity and conservation of genetic resources using wild, cultivated and elite barleys. *Plant Science*, **173**: 638–649.

Varshney, R.K., Graner, A. and Sorrells, M.E. (2005). Genetic microsatellite markers in plants: features and applications. *Trends in Biotechnology*, **23**: 48-55.

Varshney, R.K., Thiel, T., Stein, N., Langridge, P. and Graner, A. (2002). *In silico* analysis on frequency and distribution of microsatellites in ESTs of some cereal species. *Cellular and Molecular Biology Letter*, **7**: 537-546.

Vellve, R. (1993). The decline of diversity in European agriculture. *Ecologist*, **23**: 48-63.

Vierling, A.R. and Nguyen, H. (1992). Use of RAPD markers to determine the genetic diversity of diploid, wheat genotypes. TAG. *Theoretical and Applied Genetics*, **84**: 835-838.

Virk, P.S., Ford-Lloyd, B.V., Jackson, M.T. and Newbury, J.H. (1995). Use of RAPD for the study of diversity within plant germplasm collections. *Heredity*, **74**: 170-179.

Vishwakarma, M.K., Mishra, V.K., Gupta, P.K., Yadav, P.S., Kumar, H. and Joshi, A.K. (2014). Introgression of the high grain protein gene Gpc-B1 in an elite wheat variety of Indo-Gangetic Plains through marker assisted backcrossbreeding. *Current Plant Biology*, **1**: 60–67.

Voorrips, R.E., Jongerius, M.C. and Kanne, H.I. (1997). Mapping of two genes for resistance to clubroot *(Plasmodiophora brassicae)* in a population of doubled haploid lines of *Brassica oleracea* by means of RFLP and AFLP markers. *Theoretical and Applied Genetics*, **94**: 75-82.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van-de-Lee. T., Hornes, M., *et al*. (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research*, **23**: 4407-4414.

Vuylsteke, M., Mank, R., Antonise, R., Bastiaans, E., Senior, M.L., Stuber, C.W., *et al*. (1999). Two high-density AFLP® linkage maps of *Zea mays* L., Analysis of distribution of AFLP markers. *Theoretical and Applied Genetics*, **99**: 921-935.

Waines, J.G. and Bernhart, D. (1992). Biosystematic research in *Aegilops* and *Triticum. Hereditas*, **116**: 207-212.

Wang, C.M., Li, L.H., Zhang, X.T., Gao, Q., Wang, R.F. and An, D.G. (2009). Development and application of EST-STS markers specific to chromosome 1RS of Secale cereal. *Cereal Research Communication*, **37**: 13-21.

Wang, H. (2012). Dynamic QTL Analysis of Kernel Weight in Wheat at Different Developmental Stages. *Journal of Plant Genetic Resources*, **13**(6): 1055-1060.

Wang, H., Wang, X., Chen, P. and Liu, D. (2007). Assessment of genetic diversity of Yunnan, Tibetan, and Xinjiang wheat using SSR markers. *Journal of Genetics and Genomics* (Formerly *Acta Genetica Sinica*), **34**(7): 623-633.

Wang, L., Cui, F., Wang, J.P., Li, J., Ding, A.M., Zhao, C.H., *et. al*. (2012). Conditional QTL mapping of protein content in wheat with respect to grain yield and its components. *Journal of Genetics*, **91**: 303-312.

Wang, R.R.C., Chen, J. and Joppa, L.O.I.C. (1995). Production and identification of chromosome specific RAPD markers for Langdon durum wheat disomic substitution lines. *Crop Sciences*, **35**: 886-888.

Wang, S., Wong, D., Forrest, K., Allen, A., Chao, S., Huang, B.E., *et al*. (2014). Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. *Plant Biotechnology Journal*, **12**: 787-796.

Wang, Z.Y. and Tanksley, S.D. (1989). Restriction fragment length polymorphism in *Oryza* sativa. L. *Genome*, **32**: 1113-1118.

Wang. Z., Weber, J.L., Zhang, G. and Tanksley, S.D. (1994). Survey of plant short tandem DNA repeats. *Theoretical and Applied Genetics*, **88**: 1-6.

Warburton, M. and Hoisington, D. (2001). Applications of molecular marker techniques to the use of international germplasm collections. In: Plant Genotyping–The DNA Fingerprinting of Plants. Henry, R. (ed.). CAB International, Oxon, UK.

Ward, D.J. (1962). Some evolutionary aspects of certain morphological characters in a world collection of barley. USDA *Technical Bulletin*, **1276**: 1-112.

Weber, D. and Helentjaris (1989). Mapping RFLP loci in Maize using B-A translocations. *Genetics*, **121**: 583-590.

Weber, J.L. (1990). Informativeness of human (dC-dA)n (dGdT)n polymorphisms. *Genomics*, **7**: 524-530.

Weber, J.L. and May, P.E. (1989). Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *American Journal of Human Genetics*, **44**: 388-396.

Weising, K., Nybom, H., Woff and Meyer, W. (1995). DNA fingerprinting in plants and fungi. A Arbor (eds) CRC Press. Inc. Boca Raton, Florida, USA. **pp**. 1-3.

Weising, K., Nybon, H., Wolff, K. and Kahl, G. (2005). Applications of DNA fingerprinting in plant sciences. In: DNA Fingerprinting in Plants. Principles, Methods and Applications (2nd edition). Taylor & Francis (eds.) CRC Press Boca Raton, London, New York, Singapore. Chapter 6, **pp**. 235-276.

Welsh, J. and McCelland, M. (1990). Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Research*, **18**: 7213-7218.

Welsh, J., Hoercutt, R.J., McCelland, M. and Sobral, B.W.S. (1991). Parentage determination to maize hybrids using the arbitrary primed polymerase chain reaction (AP-PCR). *Theoretical and Applied Genetics*, **82**: 473-476.

Wemer, J.E., Endo, T.R. and Gill, B.S. (1992). Towards a cytogenetically based physical map of the wheat genome. *Proceedings of National Academy of Sciences (USA)*, **89**: 11307-1311.

Wenzl, P., Li, H., Carling, J., Zhou, M., Raman, H., Paul, E., *et al*. (2006). A high-density consensus map of barley linking DArT markers to SSR, RFLP and STS loci and agricultural traits. *BMC Genomics*, **7**: 206.

Westman, A.L. and Kresovich, S. (1997). Use of molecular marker techniques for description of plant genetic variation. In: Biotechnology and Plant Genetic Resources. Callow, J.A., Ford-Lloyd, B.V. and Newbury, H.J. (eds.). CAB International, UK. **pp**. 9-48.

Wicker, T., Krattinger, S.G., Lagudah, E.S., Komatsuda, T., Pourkheirandish, M., Matsumoto, T., *et al*. (2009). Analysis of intraspecies diversity in wheat and barley genomes identifies breakpoints of ancient haplotypes and provides insight into the structure of diploid and hexaploid Triticeae gene pools. *Plant Physiology*, **149**: 258-270.

Wiese, M.W. (1987). Wheat and Wheat Improvement. In: Compendium of wheat diseases (2nd ed.). Heyne, E.J. (ed.). APS Press, St Paul, Minn 55121, USA. **pp**. 3.

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

Winkler, H. (1920). Verebung und Ursache der Parthenogenesis im Pflanzen und Tierreich Fischer, Jena Street, USA.

Winter, P. and Kahl, G. (1995). Molecular marker technologies for plant improvement. *World Journal of Microbiology and Biotechnology*, **11**: 438-448.

Wolff, R.K., Plaetke, R., Jeffreys, A.J. and White, R. (1991). Unequal crossing over between homologous chromosomes is not the major mechanism involved in the generation of new alleles at VNTR loci. *Genomics*, **5**: 382-384.

Wu, J., Maehara, T., Shimokawa, T., Yamamoto, S., Harada, C., Takazaki, Y., *et al*. (2002). A comprehensive rice transcript map containing 6591 expressed sequence tag sites. *Plant Cell,* **14**: 525-535.

Xiao, Y., Yu, C., Yabes, J.M., Lei, J., Wang, X., Wang, Z., *et al*. (2013). Linkage disequilibrium in a diversity and stress adaptation rice panel for association mapping. *Molecular Plant Breeding*, **4**: 297-303.

Xie, D., Devos, K.M., Moore, G. and Gale, M.D. (1993). RFLP based genetic maps of the homeologous group 5 chromosomes of bread wheat (*Triticum aestivum*). *Theoretical and Applied Genetics*, **87**: 70-74.

Xu, C., Gao, J., Du, Z., Li, D., Wang, Z., Li, Y., *et al*. (2016). Identifying the genetic diversity, genetic structure and a core collection of *Ziziphus jujuba* Mill. var. jujuba accessions using microsatellite markers. *Nature Scientific Reports*, **6**: 31503. doi: 10.1038/srep31503.

Yadav, A.K., Singh, D. and Arya, R.K. (2013). Morphological characterization of indian mustard (*Brassica juncea*) genotypes and their application for dus testing. *Indian Journal of Agricultural Sciences*, **83**(12): 1305-1316.

Yan, W. and Fregeau-Reid, J. (2008). Breeding line selection based on multiple traits. *Crop Sciences*, **48**: 417-423.

Yan, W. and Wallace, D.H. (1995). Breeding for negatively associated traits. *Plant Breeding Review*, **13**: 141-177.

Yang, D., Liu, Y., Cheng, H., Chang, L., Chen, J., Chai, S., *et al*. (2016). Genetic dissection of flag leaf morphology in wheat (*Triticum aestivum* L.) under diverse water regimes. *BMC Genetics*, **17**(1): 94. doi: 10.1186/s12863-016-0399-9 (accessed on 12.8.2016).

Yen, T., Baenrige, P.S. and Moris, R. (1996). Genome constitution of bread wheat: Current status. In: Methods of genome analysis in plants. Jauhar, P.P. (ed.). CRC Press, Boca Raton, Florida, USA, **pp**. 359-373.

Yong, Q., Jérome, S.P., Devin, C.S., Adeline, D., Crossman, C.H., Gerard, R., *et al*. (2006). Types and rates of sequence evolution at *HMW-glutenin* locus in hexaploid wheat and its ancestral genomes. *Genetics*, **106**: 060-756.

Zabeau, M. and Vos, P. (1993). Selective restriction fragment amplification: a general method for DNA fingerprinting. European Patent, 0534858 A1.

Zadoks, J.C., Chang, T.T. and Konzak, C.F. (1974). A decimal code for the growth stages of cereals. *Weed Research*, **14**: 415-421.

Zahariev, M., Dimov, A., Stankova, P., David, J. and Monneveux, P. (2003). Morphological diversity and potential interest for wheat improvement of three *Aegilops* L. species from Bulgaria. *Genetic Resources and Crop Evolution*, **50**(5): 507-517.

Zane, L., Bargelloni, L. and Patarnello, T. (2002). Strategies for microsatellite isolation; a review. *Molecular Ecology*, **11**: 1-16.

Zarkti, H., Quabbou, H., Udupa, S.M., Gaboun, F. and Hilali, A. (2012). Agro-morphological variability in durum wheat landraces of Morocco. *Australian Journal of Crop Sciences*, **6**(7): 1172-1178.

Zeng, Y., Yang, S., Cui, H., Yang, X., Xu, L., Du, J., *et al*. (2009). QTLs of cold-related traits at the booting stage for NIL-RILsin rice revealed by SSR. *Genes and Genomics*, **31**: 143-154.

Zhang, D., Bowden, R.L., Yu, J., Carver, B.F. and Bai G. (2014). Association analysis of stem rust resistance in U.S. winter wheat**.** *PLOS*, **9**(7): e103747. doi: 10.1371/journal.pone.0103747. eCollection.

Zhang, H., Chen, J., Li, R., Deng, Z., Zhang, K., Liu, B., *et al*. (2016). Conditional QTL mapping of three yield components in common wheat (*Triticum aestivum* L.). *The Crop Journal*, **4**(3): 220-228.

Zhang, H.Q., Jia, J.Z., Yang, H. and Zhang, B.S. (2008). SSR mapping of stripe rust resistance gene from *Ae. tauschii*. *Yi Chuan*, **30**(4): 491-4.

Zhang, K., Wang, J., Zhang, L., Rong, C., Zhao, F., Peng, T., *et al*. (2013). Association Analysis of Genomic Loci Important for Grain Weight Control in Elite Common Wheat Varieties Cultivated with Variable Water and Fertiliser Supply. *PLoS ONE*, **8**(3): e57853.

Zhang, L. (2006). Study of the transferability of microsatellite markers derived from bread wheat (*T. aestivum*) or rice (*O. sativa*) ESTs (EST-SSRs) to their close and wild relative and evaluation of their potential for the organisation of genetic resources in the grass family. Ph.D. thesis (Published in English. online) submitted to Universit´e Blaise Pascal - Clermont-Ferrand II; Universit´e d'Auvergne - Clermont-Ferrand I, France.

Zhang, P., Dreisigacker, S., Buerkert, A., Alkhanjari, S., Melchinger, A.E. and Warburton, M.L. (2006). Genetic diversity and relationships of wheat landraces from Oman investigated with SSR markers. *Genetics Resource and Crop Evolution*, **53**: 1351–1360.

Zhao, Y., Wang, H., Chen, W. and Li, Y. (2014). Genetic structure, linkage disequilibrium and association mapping of verticillium wilt resistance in elite cotton (*Gossypium hirsutum* L.) germplasm population. *PLoS ONE*, **9**: e86308.

Zhou, W., Kolb, F.L., Domier, L.L. and Wang, S. (2005). SSR markers associated with fertility restoration genes against *Triticum timopheevii* cytoplasm in *Triticum aestivum*. *Euphytica*, **141**(1-2): 33-40.

Zhu, J., Gale, M.D., Quarrie, S., Jackson, M.T. and Bryan, G.J. (1998). AFLP markers for the study of rice biodiversity. *Theoretical and Applied Genetics*, **96**: 602-611.

Zhu, Y., Hu, j., Han, R., Wang, Y. and Zhu, S. (2011). Fingerprinting and identification of closely related wheat (*Triticum aestivum* L.) cultivars using ISSR and fluorescence-labeled TP-M13-SSR markers. *Australian Journal of Crop Sciences,* **5**(7): 846-850.

Zhu, Y., Queller, D.C. and Strassmann, J.E. (2000). A phylogenetic perspective on sequence evolution in microsatellite loci. *Journal of Molecular Evolution*, **50**: 324-338.

Zietkiewicz, E., Rafalski, A. and Labuda, D. (1994). Genome fingerprinting by simple sequence repeat (SSR). anchored polymerase chain reaction amplification. *Genomics*, **20**: 176-183.

Zwart, R.S., Bansal, U.K., Thompson, J.P., Williamson, P.M. and Bariana, H.S. (2008). QTL mapping of multiple disease resistance traits in a synthetic hexaploid x bread wheat population. In: proceedings of the 11th international wheat genetics symposium, August 24-29. Brisbane, Australia.

APPENDIX

APPENDIX

A1. Dye Information for genotyping on the 3730xl Genetic analyzer.

Dyes supported on the 3730: 6-FAM, VIC, NED, PET

Ladder: LIZ 500 bp size standard

6-FAM = Blue

VIC = Green

NED = Yellow (Black)

PET = Red

LIZ = Orange

A2. Microsatellite analysis workflow used in GeneMapper software

Set Up the Microsatellite Analysis.

1. Create a kit, panel, and markers for the project.
2. Create a new project and add sample files.
3. Set the analysis parameters and table settings for the project.
4. Perform an initial analysis.
5. Create a bin set and generate bins (using Auto Bin)

Analyze and Examine Results

1. Edit the analysis method to specify a bin set.
2. Analyze the samples in the project.
3. Examine the results

Print and Export the Results (Optional)

- Print results.
- Export results

A3. Analysis parameters used in GeneMapper software.

A collection of user-defined settings (including an analysis method, size standard, and panel) that determine the sizing and genotyping algorithms used by the GeneMapper® Software to analyze all sample files in a project.

Bin

A fragment size (bp) and dye color that define an allele within a marker. You create a bin for each possible allele associated with a marker.

Bin Set

A collection of bins (allele definitions), typically specific to a set of experimental conditions.

Marker

A microsatellite marker is defined by a name, fragment size range (bp), dye color, and repeat length.

Panel

A group of markers. In the GeneMapper Software, you associate a panel with a bin set to provide bin definitions for the markers.

Kit

A group of panels.

A4. Protocol used in GeneMapper software.

Create a New Kit

Create a Custom Panel

Create a Markers

Create a Project

Add Samples

Apply Analysis Settings

Start Analysis

Create a New Bin Set

Using the Auto Bin Function

A5. Quality analysis parameters used in GeneMapper software.

Quality Flags – This tab includes: – Settings that determine the importance of individual flagged Process Quality Values (PQVs) to the overall Genotype Quality (GQ). You can weight each PQV from 0 to 1, with 0 being of no importance and 1 meaning very important.

– Threshold settings that determine when the SQ and GQ are flagged as Pass, Check, or Low Quality.

– The SQ and GQ are given initial scores of 1. The value of any flagged PQVs are then subtracted from 1 to give the final SQ and GQ scores.

– Assume Linearity Range, where the size calling algorithm assumes the fragment migration is linear for a given size range when calculating the Sizing Quality (SQ).

Leave the default values for all settings.

A6. Detailed ANOVA table obtained from PowerMarker software

Marker Name	Major Allele frequency	No. of Genotype	Sample Size	No. of obs.	Allele No.	Availability	Gene Diversity	Hetero-zygosity	PIC	F
xgwm136	0.16	26	618	618	26	1.0	0.913	0.0	0.906	1.00
xgwm497	0.15	25	618	618	25	1.0	0.915	0.0	0.909	1.00
xgwm448	0.17	23	618	618	23	1.0	0.922	0.0	0.917	1.00
xgwm445	0.19	22	618	618	22	1.0	0.921	0.0	0.916	1.00
xgwm2	0.17	24	618	618	24	1.0	0.915	0.0	0.909	1.00
xgwm494	0.11	23	618	618	23	1.0	0.930	0.0	0.926	1.00
xgwm165	0.14	20	618	618	20	1.0	0.917	0.0	0.911	1.00
Xgwm637	0.26	22	618	618	22	1.0	0.877	0.0	0.866	1.00
xgwm443	0.21	23	618	618	23	1.0	0.896	0.0	0.888	1.00
Xgwm666	0.19	22	618	618	22	1.0	0.908	0.0	0.902	1.00
xgwm459	0.10	23	618	618	23	1.0	0.932	0.0	0.928	1.00
xgwm169	0.12	22	618	618	22	1.0	0.915	0.0	0.909	1.00
xgwm635	0.26	21	618	618	21	1.0	0.884	0.0	0.876	1.00
xgwm282	0.17	24	618	618	24	1.0	0.917	0.0	0.911	1.00
xgwm550	0.17	15	618	618	15	1.0	0.903	0.0	0.895	1.00
xgwm268	0.19	20	618	618	20	1.0	0.896	0.0	0.887	1.00
xgwm257	0.29	16	618	618	16	1.0	0.861	0.0	0.849	1.00
xgwm47	0.20	17	618	618	17	1.0	0.888	0.0	0.878	1.00
xgwm533	0.18	18	618	618	18	1.0	0.912	0.0	0.906	1.00
xgwm181	0.17	18	618	618	18	1.0	0.895	0.0	0.886	1.00
xgwm368	0.17	17	618	618	17	1.0	0.889	0.0	0.879	1.00
Xgwm513	0.18	18	618	618	18	1.0	0.896	0.0	0.887	1.00

Contd…

A6. Contd…

Marker Name	Major Allele frequency	No. of Genotype	Sample Size	No. of obs.	Allele No.	Availability	Gene Diversity	Hetero-zygosity	PIC	F
xgwm544	0.27	17	618	618	17	1.0	0.854	0.0	0.840	1.00
Xgwm335	0.18	18	618	618	18	1.0	0.901	0.0	0.893	1.00
Xgwm361	0.14	18	618	618	18	1.0	0.910	0.0	0.903	1.00
Xgwm219	0.19	20	618	618	20	1.0	0.906	0.0	0.899	1.00
xgwm400	0.14	18	618	618	18	1.0	0.907	0.0	0.899	1.00
xgwm131	0.14	20	618	618	20	1.0	0.917	0.0	0.911	1.00
Xgwm458	0.17	15	618	618	15	1.0	0.896	0.0	0.887	1.00
xgwm642	0.14	16	618	618	16	1.0	0.914	0.0	0.908	1.00
xgwm484	0.25	14	618	618	14	1.0	0.841	0.0	0.823	1.00
xgwm349	0.25	13	618	618	13	1.0	0.865	0.0	0.853	1.00
xgwm183	0.11	16	618	618	16	1.0	0.924	0.0	0.919	1.00
xgwm3	0.14	15	618	618	15	1.0	0.910	0.0	0.903	1.00
xgwm194	0.17	16	618	618	16	1.0	0.913	0.0	0.907	1.00
xgwm624	0.29	15	618	618	15	1.0	0.870	0.0	0.860	1.00
xgwm190	0.15	15	618	618	15	1.0	0.908	0.0	0.901	1.00
Xgwm182	0.22	13	618	618	13	1.0	0.889	0.0	0.880	1.00
Xgdm141	0.15	12	618	618	12	1.0	0.891	0.0	0.881	1.00
BARC1121	0.19	13	618	618	13	1.0	0.903	0.0	0.895	1.00
Xgwm44	0.18	13	618	618	13	1.0	0.890	0.0	0.880	1.00
Xgwm37	0.16	15	618	618	15	1.0	0.910	0.0	0.903	1.00
Mean	**0.18**	**18**	**618**	**618**	**18**	**1.0**	**0.900**	**0.0**	**0.892**	**1.00**

A7. Summary AMOVA table for A genome.

Source	Among Pops	Within Pops	Total
df	8	509	617
SS	718.132	15059.473	15777.605
MS	89.767	24.728	
Est. Var.	2.708	24.728	27.436
%	10%	90%	100%

A8. Summary AMOVA table for B genome.

Source	Among Pops	Within Pops	Total
df	8	509	617
SS	712.549	14777.535	15490.084
MS	89.069	24.265	
Est. Var.	2.699	24.265	26.964
%	10%	90%	100%

A9. Summary AMOVA table for D genome.

Source	Among Pops	Within Pops	Total
df	8	509	617
SS	697.507	14782.376	15479.883
MS	87.188	24.273	
Est. Var.	2.620	24.273	26.893
%	10%	90%	100%

A.10. Summary AMOVA table for ABD genome.

Source	Among Pops	Within Pops	Total
df	8	509	617
SS	2128.189	44619.384	46747.573
MS	266.024	73.267	
Est. Var.	8.027	73.267	81.293
%	10%	90%	100%

A.11 Ancestry detaill of 618 Wheat landraces derived from 42 SSR markers.

Label	accessions	FST1	FST2	FST3	FST4
1	IC_104537	0.99	0.007	0.001	0.002
2	IC_104540	0.001	0.004	0.993	0.001
3	IC_104543	0.001	0.003	0.995	0.001
4	IC_104551	0.002	0.104	0.001	0.893
5	IC_104564	0.002	0.001	0.995	0.003
6	IC_104566	0.005	0.001	0.002	0.992
7	IC_104568	0.001	0.001	0.995	0.003
8	IC_104573	0.001	0.677	0.002	0.319
9	IC_104576	0.002	0.002	0.991	0.006
10	IC_118722	0.006	0.001	0.003	0.99
11	IC_118723	0.341	0.209	0.216	0.234
12	IC_118727	0.305	0.252	0.227	0.216
13	IC_118729	0.933	0.02	0.02	0.027
14	IC_118730	0.33	0.205	0.222	0.243
15	IC_118731	0.002	0.001	0.995	0.001
16	IC_118732	0.002	0.001	0.994	0.002
17	IC_118736	0.522	0.468	0.008	0.002
18	IC_118737	0.483	0.504	0.005	0.008
19	IC_118739	0.001	0.001	0.975	0.023
20	IC_118740	0.001	0.651	0.004	0.344
21	IC_118741	0.937	0.025	0.034	0.004
22	IC_118758	0.022	0.658	0.319	0.001
23	IC_118762	0.001	0.001	0.997	0.002
24	IC_138845	0.001	0.001	0.355	0.643
25	IC_138849	0.001	0.987	0.009	0.002
26	IC_138895	0.003	0.001	0.532	0.463

Contd…

A.11 Contd…

Label	accessions	FST1	FST2	FST3	FST4
27	IC_138896	0.996	0.001	0.001	0.001
28	IC_138893	0.003	0.01	0.976	0.011
29	IC_138897	0.997	0.001	0.001	0.001
30	IC_138898	0.553	0.007	0.163	0.276
31	IC_210548	0.001	0.004	0.995	0.001
32	IC_212142	0.001	0.004	0.995	0.001
33	IC_212145	0.142	0.856	0.001	0.001
34	IC_212179	0.001	0.003	0.995	0.001
35	IC_212185	0.258	0.523	0.141	0.079
36	IC_28664	0.002	0.001	0.996	0.001
37	IC_28926	0.001	0.001	0.997	0.002
38	IC_28958	0.001	0.995	0.002	0.001
39	IC_28969	0.008	0.99	0.002	0.001
40	IC_28974	0.002	0.996	0.001	0.001
41	IC_29035	0.005	0.977	0.013	0.006
42	IC_30284	0.003	0.983	0.009	0.004
43	IC_30276	0.377	0.358	0.263	0.001
44	IC_31496	0.001	0.994	0.002	0.003
45	IC_32520	0.001	0.998	0.001	0.001
46	IC_35069	0.996	0.002	0.001	0.001
47	IC_35070	0.001	0.998	0.001	0.001
48	IC_35071	0.002	0.001	0.006	0.992
49	IC_47478	0.005	0.001	0.99	0.003
50	IC_532055	0.003	0.001	0.994	0.002
51	IC_532057	0.008	0.001	0.267	0.723
52	IC_532058	0.002	0.001	0.991	0.006
53	IC_532059	0.997	0.001	0.002	0.001
54	IC_532061	0.995	0.002	0.001	0.002
55	IC_532062	0.001	0.211	0.784	0.005
56	IC_532064	0.007	0.001	0.85	0.143
57	IC_532066	0.001	0.001	0.995	0.002
58	IC_532067	0.001	0.001	0.996	0.002
59	IC_532068	0.366	0.621	0.01	0.002
60	IC_532071	0.002	0.001	0.074	0.923
61	IC_532072	0.002	0.001	0.9	0.097
62	IC_532073	0.988	0.002	0.009	0.001
63	IC_532074	0.002	0.001	0.992	0.004
64	IC_532077	0.001	0.992	0.005	0.001

Contd…

A.11 Contd…

Label	accessions	FST1	FST2	FST3	FST4
65	IC_532078	0.001	0.003	0.995	0.001
66	IC_532081	0.001	0.001	0.001	0.997
67	IC_532083	0.002	0.99	0.006	0.002
68	IC_532084	0.003	0.001	0.977	0.02
69	IC_532085	0.002	0.001	0.993	0.005
70	IC_532087	0.001	0.002	0.006	0.991
71	IC_532088	0.002	0.989	0.007	0.002
72	IC_532089	0.002	0.001	0.992	0.005
73	IC_532090	0.003	0.002	0.993	0.002
74	IC_532091	0.004	0.001	0.994	0.001
75	IC_532092	0.991	0.002	0.005	0.002
76	IC_532093	0.001	0.996	0.002	0.001
77	IC_532094	0.002	0.001	0.996	0.001
78	IC_532095	0.001	0.995	0.002	0.001
79	IC_532096	0.995	0.001	0.002	0.002
80	IC_532097	0.001	0.006	0.011	0.982
81	IC_532098	0.037	0.001	0.957	0.006
82	IC_532100	0.002	0.001	0.06	0.937
83	IC_532101	0.009	0.001	0.049	0.942
84	IC_532102	0.01	0.001	0.985	0.003
85	IC_532103	0.993	0.001	0.005	0.002
86	IC_532106	0.001	0.002	0.996	0.001
87	IC_532110	0.002	0.001	0.982	0.015
88	IC_532116	0.001	0.001	0.698	0.3
89	IC_532118	0.001	0.001	0.001	0.997
90	IC_532119	0.002	0.001	0.991	0.007
91	IC_532120	0.002	0.001	0.994	0.003
92	IC_532121	0.001	0.001	0.001	0.997
93	IC_532122	0.001	0.001	0.001	0.997
94	IC_532124	0.003	0.001	0.013	0.982
95	IC_532126	0.002	0.001	0.011	0.986
96	IC_532129	0.001	0.001	0.014	0.984
97	IC_532130	0.003	0.007	0.112	0.879
98	IC_532131	0.982	0.001	0.009	0.009
99	IC_532132	0.002	0.001	0.013	0.984
100	IC_532133	0.002	0.001	0.881	0.116
101	IC_532134	0.986	0.001	0.012	0.001
102	IC_532136	0.001	0.001	0.001	0.997
103	IC_532137	0.713	0.002	0.283	0.003

Contd…

A.11 Contd…

Label	accessions	FST1	FST2	FST3	FST4
104	IC_532138	0.966	0.001	0.014	0.019
105	IC_532139	0.001	0.001	0.002	0.996
106	IC_532140	0.986	0.002	0.009	0.003
107	IC_532141	0.003	0.002	0.992	0.003
108	IC_532142	0.991	0.001	0.003	0.005
109	IC_532143	0.001	0.002	0.995	0.002
110	IC_532144	0.991	0.001	0.006	0.001
111	IC_532145	0.001	0.001	0.001	0.997
112	IC_532146	0.988	0.001	0.001	0.01
113	IC_532147	0.988	0.003	0.002	0.007
114	IC_532148	0.001	0.001	0.001	0.997
115	IC_532149	0.004	0.015	0.002	0.979
116	IC_532150	0.002	0.001	0.135	0.862
117	IC_532151	0.003	0.001	0.01	0.986
118	IC_532153	0.001	0.001	0.001	0.997
119	IC_532155	0.989	0.008	0.001	0.002
120	IC_532156	0.002	0.001	0.001	0.996
121	IC_532175	0.995	0.002	0.001	0.002
122	IC_532180	0.006	0.003	0.987	0.004
123	IC_532181	0.002	0.001	0.003	0.994
124	IC_532182	0.001	0.001	0.002	0.996
125	IC_532183	0.998	0.001	0.001	0.001
126	IC_532184	0.993	0.001	0.002	0.003
127	IC_532185	0.001	0.001	0.997	0.001
128	IC_532186	0.003	0.001	0.002	0.994
129	IC_532187	0.996	0.001	0.002	0.001
130	IC_532188	0.996	0.002	0.002	0.001
131	IC_532201	0.001	0.001	0.995	0.003
132	IC_532204	0.001	0.001	0.997	0.001
133	IC_532205	0.016	0.001	0.002	0.981
134	IC_532206	0.124	0.001	0.02	0.855
135	IC_532209	0.001	0.001	0.997	0.001
136	IC_532210	0.473	0.522	0.003	0.003
137	IC_532213	0.99	0.001	0.001	0.008
138	IC_532214	0.996	0.001	0.001	0.001
139	IC_532216	0.001	0.006	0.992	0.001
140	IC_532217	0.996	0.001	0.001	0.002
141	IC_532218	0.001	0.001	0.997	0.001
142	IC_532219	0.003	0.981	0.011	0.005

Contd…

A.11 Contd…

Label	accessions	FST1	FST2	FST3	FST4
143	IC_532221	0.002	0.001	0.002	0.994
144	IC_532223	0.996	0.001	0.001	0.003
145	IC_532224	0.001	0.001	0.002	0.996
146	IC_532225	0.002	0.001	0.996	0.001
147	IC_532227	0.002	0.001	0.001	0.997
148	IC_532228	0.001	0.003	0.994	0.002
149	IC_532229	0.001	0.02	0.978	0.001
150	IC_532230	0.122	0.002	0.003	0.874
151	IC_532231	0.413	0.583	0.001	0.003
152	IC_532232	0.002	0.001	0.006	0.991
153	IC_532237	0.357	0.614	0.013	0.016
154	IC_532238	0.001	0.001	0.996	0.002
155	IC_532239	0.997	0.001	0.001	0.001
156	IC_532240	0.435	0.564	0.001	0.001
157	IC_532241	0.011	0.987	0.002	0.001
158	IC_532242	0.037	0.002	0.108	0.853
159	IC_532243	0.001	0.001	0.993	0.005
160	IC_532244	0.001	0.001	0.994	0.004
161	IC_532245	0.001	0.001	0.997	0.001
162	IC_532246	0.002	0.001	0.002	0.995
163	IC_532247	0.005	0.001	0.002	0.992
164	IC_532248	0.004	0.001	0.081	0.914
165	IC_532249	0.001	0.001	0.001	0.997
166	IC_532250	0.008	0.99	0.001	0.001
167	IC_532251	0.033	0.963	0.002	0.002
168	IC_532253	0.001	0.001	0.001	0.998
169	IC_532256	0.009	0.986	0.004	0.001
170	IC_532257	0.001	0.997	0.001	0.001
171	IC_532258	0.02	0.902	0.042	0.036
172	IC_532259	0.001	0.001	0.001	0.997
173	IC_532261	0.996	0.001	0.001	0.002
174	IC_532262	0.007	0.001	0.09	0.902
175	IC_532263	0.996	0.001	0.001	0.002
176	IC_532264	0.008	0.001	0.009	0.982
177	IC_532265	0.001	0.001	0.001	0.998
178	IC_532267	0.996	0.001	0.001	0.002
179	IC_532268	0.001	0.001	0.995	0.003
180	IC_532271	0.001	0.997	0.001	0.001
181	IC_532272	0.003	0.001	0.005	0.992

Contd…

A.11 Contd…

Label	accessions	FST1	FST2	FST3	FST4
182	IC_532273	0.001	0.995	0.001	0.003
183	IC_532274	0.001	0.997	0.001	0.001
184	IC_532276	0.001	0.001	0.997	0.001
185	IC_532277	0.001	0.001	0.002	0.996
186	IC_532279	0.001	0.205	0.77	0.025
187	IC_532281	0.001	0.003	0.996	0.001
188	IC_532282	0.001	0.002	0.997	0.001
189	IC_532284	0.001	0.997	0.001	0.001
190	IC_532285	0.001	0.998	0.001	0.001
191	IC_532286	0.001	0.001	0.001	0.997
192	IC_532289	0.801	0.01	0.016	0.173
193	IC_532290	0.003	0.991	0.003	0.003
194	IC_532292	0.002	0.996	0.001	0.001
195	IC_532297	0.002	0.996	0.001	0.001
196	IC_532298	0.002	0.001	0.002	0.995
197	IC_532300	0.001	0.001	0.001	0.997
198	IC_532309	0.001	0.998	0.001	0.001
199	IC_532310	0.002	0.001	0.002	0.995
200	IC_532318	0.001	0.985	0.003	0.012
201	IC_532406	0.001	0.996	0.002	0.002
202	IC_532473	0.001	0.008	0.99	0.002
203	IC_532474	0.001	0.996	0.001	0.002
204	IC_532475	0.001	0.002	0.997	0.001
205	IC_532476	0.001	0.995	0.002	0.002
206	IC_532478	0.996	0.001	0.001	0.002
207	IC_532480	0.001	0.997	0.001	0.001
208	IC_532481	0.001	0.948	0.001	0.05
209	IC_532482	0.001	0.995	0.002	0.001
210	IC_532483	0.001	0.001	0.994	0.004
211	IC_532485	0.001	0.001	0.001	0.997
212	IC_532486	0.018	0.98	0.001	0.001
213	IC_532487	0.001	0.001	0.001	0.997
214	IC_532489	0.002	0.996	0.001	0.001
215	IC_532490	0.001	0.001	0.994	0.004
216	IC_532492	0.005	0.001	0.001	0.993
217	IC_532497	0.001	0.995	0.002	0.003
218	IC_532502	0.001	0.995	0.001	0.003
219	IC_532503	0.997	0.001	0.001	0.001
220	IC_532504	0.001	0.002	0.001	0.996

Contd…

A.11 Contd…

Label	accessions	FST1	FST2	FST3	FST4
221	IC_532697	0.001	0.994	0.001	0.004
222	IC_532698	0.002	0.995	0.001	0.001
223	IC_532699	0.001	0.996	0.002	0.001
224	IC_532700	0.002	0.99	0.002	0.005
225	IC_532701	0.997	0.001	0.001	0.001
226	IC_532704	0.002	0.996	0.001	0.001
227	IC_532705	0.001	0.001	0.001	0.997
228	IC_532707	0.01	0.001	0.002	0.987
229	IC_532708	0.997	0.001	0.001	0.001
230	IC_532719	0.002	0.003	0.001	0.993
231	IC_532723	0.001	0.997	0.001	0.001
232	IC_532725	0.001	0.997	0.001	0.001
233	IC_532726	0.001	0.001	0.001	0.997
234	IC_532727	0.997	0.001	0.001	0.002
235	IC_532728	0.002	0.001	0.001	0.996
236	IC_532729	0.996	0.001	0.001	0.002
237	IC_532736	0.001	0.036	0.957	0.005
238	IC_532737	0.001	0.998	0.001	0.001
239	IC_532738	0.002	0.003	0.001	0.995
240	IC_532768	0.997	0.001	0.001	0.001
241	IC_532773	0.001	0.001	0.997	0.002
242	IC_532775	0.001	0.994	0.004	0.002
243	IC_532779	0.001	0.001	0.998	0.001
244	IC_532784	0.997	0.001	0.001	0.001
245	IC_532787	0.002	0.001	0.996	0.001
246	IC_532790	0.997	0.001	0.001	0.001
247	IC_532794	0.001	0.997	0.001	0.001
248	IC_532805	0.001	0.003	0.001	0.995
249	IC_532807	0.001	0.989	0.009	0.001
250	IC_532811	0.001	0.003	0.001	0.994
251	IC_532813	0.997	0.001	0.001	0.001
252	IC_532815	0.001	0.969	0.029	0.001
253	IC_532821	0.997	0.001	0.001	0.002
254	IC_532832	0.001	0.969	0.029	0.001
255	IC_532833	0.001	0.987	0.011	0.001
256	IC_532835	0.707	0.29	0.001	0.002
257	IC_532837	0.753	0.244	0.001	0.002
258	IC_532839	0.998	0.001	0.001	0.001
259	IC_532841	0.996	0.001	0.001	0.001

Contd…

A.11 Contd…

Label	accessions	FST1	FST2	FST3	FST4
260	IC_532842	0.996	0.001	0.001	0.001
261	IC_532847	0.996	0.001	0.001	0.002
262	IC_532849	0.994	0.001	0.001	0.003
263	IC_532851	0.002	0.964	0.033	0.001
264	IC_532853	0.996	0.002	0.001	0.001
265	IC_532855	0.997	0.001	0.001	0.001
266	IC_532857	0.996	0.001	0.001	0.002
267	IC_532863	0.001	0.001	0.997	0.001
268	IC_532868	0.997	0.001	0.001	0.001
269	IC_532872	0.001	0.997	0.001	0.001
270	IC_532880	0.05	0.004	0.002	0.944
271	IC_532886	0.996	0.001	0.001	0.002
272	IC_532887	0.979	0.007	0.006	0.008
273	IC_532891	0.005	0.003	0.001	0.99
274	IC_532905	0.996	0.001	0.002	0.002
275	IC_532910	0.997	0.001	0.001	0.001
276	IC_532923	0.996	0.001	0.002	0.001
277	IC_532930	0.001	0.001	0.996	0.001
278	IC_532934	0.997	0.001	0.001	0.001
279	IC_532935	0.938	0.059	0.001	0.002
280	IC_532936	0.07	0.002	0.002	0.926
281	IC_532937	0.108	0.002	0.002	0.887
282	IC_533953	0.028	0.002	0.002	0.969
283	IC_533954	0.002	0.001	0.002	0.996
284	IC_533962	0.996	0.002	0.001	0.002
285	IC_533963	0.001	0.001	0.001	0.997
286	IC_533964	0.002	0.001	0.001	0.996
287	IC_533965	0.997	0.001	0.001	0.001
288	IC_533966	0.64	0.356	0.001	0.003
289	IC_533970	0.996	0.002	0.001	0.001
290	IC_533971	0.006	0.002	0.002	0.99
291	IC_534072	0.005	0.001	0.001	0.993
292	IC_534123	0.007	0.001	0.002	0.99
293	IC_534189	0.001	0.002	0.001	0.996
294	IC_534190	0.953	0.043	0.002	0.002
295	IC_534192	0.996	0.001	0.001	0.002
296	IC_534193	0.993	0.002	0.001	0.004
297	IC_534197	0.998	0.001	0.001	0.001
298	IC_534223	0.996	0.001	0.001	0.002

Contd…

A.11 Contd…

Label	accessions	FST1	FST2	FST3	FST4
299	IC_534235	0.001	0.001	0.998	0.001
300	IC_534322	0.001	0.001	0.998	0.001
301	IC_534323	0.997	0.001	0.001	0.001
302	IC_534334	0.001	0.001	0.997	0.001
303	IC_534363	0.996	0.002	0.001	0.001
304	IC_534419	0.995	0.001	0.002	0.001
305	IC_534430	0.001	0.001	0.995	0.003
306	IC_534432	0.001	0.001	0.996	0.001
307	IC_534435	0.997	0.001	0.001	0.001
308	IC_534451	0.001	0.001	0.998	0.001
309	IC_534455	0.001	0.001	0.997	0.001
310	IC_534480	0.001	0.001	0.994	0.004
311	IC_534481	0.001	0.001	0.998	0.001
312	IC_534509	0.997	0.001	0.001	0.001
313	IC_534524	0.996	0.001	0.001	0.002
314	IC_534534	0.996	0.001	0.001	0.002
315	IC_534543	0.997	0.001	0.001	0.001
316	IC_534549	0.008	0.004	0.986	0.002
317	IC_534553	0.995	0.003	0.001	0.001
318	IC_534554	0.001	0.001	0.997	0.001
319	IC_534555	0.997	0.001	0.001	0.001
320	IC_534556	0.002	0.001	0.995	0.002
321	IC_534557	0.004	0.006	0.001	0.989
322	IC_534560	0.997	0.002	0.001	0.001
323	IC_534561	0.002	0.001	0.996	0.001
324	IC_534562	0.001	0.997	0.001	0.001
325	IC_534564	0.997	0.001	0.001	0.001
326	IC_534565	0.995	0.003	0.001	0.001
327	IC_534566	0.002	0.001	0.993	0.004
328	IC_534568	0.996	0.001	0.001	0.001
329	IC_534572	0.997	0.001	0.001	0.001
330	IC_534582	0.997	0.001	0.001	0.001
331	IC_534605	0.001	0.995	0.003	0.001
332	IC_534609	0.003	0.001	0.993	0.004
333	IC_534688	0.998	0.001	0.001	0.001
334	IC_534690	0.997	0.001	0.001	0.001
335	IC_534697	0.001	0.996	0.001	0.001
336	IC_534720	0.001	0.998	0.001	0.001
337	IC_534721	0.997	0.001	0.001	0.001

Contd…

A.11 Contd…

Label	accessions	FST1	FST2	FST3	FST4
338	IC_534723	0.997	0.001	0.001	0.001
339	IC_534736	0.001	0.997	0.001	0.001
340	IC_534737	0.001	0.997	0.001	0.001
341	IC_534739	0.001	0.998	0.001	0.001
342	IC_534741	0.002	0.996	0.001	0.001
343	IC_534742	0.001	0.801	0.002	0.196
344	IC_534743	0.998	0.001	0.001	0.001
345	IC_534744	0.001	0.997	0.001	0.001
346	IC_534745	0.001	0.753	0.001	0.245
347	IC_534747	0.997	0.001	0.001	0.001
348	IC_534748	0.001	0.793	0.002	0.204
349	IC_534749	0.995	0.001	0.001	0.002
350	IC_534750	0.001	0.997	0.001	0.001
351	IC_534751	0.001	0.752	0.051	0.196
352	IC_534752	0.001	0.997	0.001	0.001
353	IC_534754	0.995	0.001	0.001	0.002
354	IC_534755	0.001	0.009	0.001	0.989
355	IC_534756	0.997	0.001	0.001	0.001
356	IC_534757	0.995	0.001	0.001	0.003
357	IC_534758	0.002	0.001	0.996	0.001
358	IC_534759	0.997	0.001	0.001	0.001
359	IC_534760	0.996	0.001	0.001	0.002
360	IC_534761	0.997	0.001	0.001	0.001
361	IC_534762	0.001	0.004	0.003	0.991
362	IC_534763	0.12	0.004	0.868	0.008
363	IC_534764	0.997	0.001	0.001	0.001
364	IC_534765	0.998	0.001	0.001	0.001
365	IC_534766	0.001	0.731	0.002	0.266
366	IC_534767	0.005	0.992	0.002	0.001
367	IC_534768	0.001	0.997	0.001	0.001
368	IC_534769	0.997	0.001	0.001	0.001
369	IC_534770	0.002	0.995	0.001	0.002
370	IC_534771	0.001	0.001	0.001	0.997
371	IC_534772	0.996	0.001	0.001	0.001
372	IC_534773	0.001	0.997	0.001	0.001
373	IC_534774	0.001	0.001	0.001	0.997
374	IC_534775	0.998	0.001	0.001	0.001
375	IC_534776	0.001	0.997	0.001	0.001
376	IC_534777	0.998	0.001	0.001	0.001

Contd…

A.11 Contd…

Label	accessions	FST1	FST2	FST3	FST4
377	IC_534778	0.001	0.001	0.997	0.001
378	IC_534786	0.001	0.997	0.001	0.001
379	IC_534787	0.998	0.001	0.001	0.001
380	IC_534792	0.002	0.996	0.001	0.001
381	IC_534794	0.001	0.001	0.001	0.998
382	IC_534798	0.211	0.001	0.787	0.001
383	IC_534802	0.002	0.002	0.001	0.995
384	IC_534805	0.001	0.001	0.997	0.001
385	IC_534806	0.001	0.001	0.001	0.997
386	IC_534808	0.001	0.001	0.001	0.997
387	IC_534811	0.001	0.001	0.001	0.997
388	IC_534814	0.001	0.001	0.997	0.001
389	IC_534819	0.001	0.002	0.001	0.996
390	IC_534820	0.997	0.001	0.001	0.001
391	IC_534822	0.001	0.001	0.996	0.001
392	IC_534823	0.997	0.001	0.001	0.001
393	IC_534854	0.001	0.001	0.001	0.997
394	IC_534855	0.997	0.001	0.001	0.001
395	IC_534857	0.997	0.001	0.001	0.001
396	IC_534858	0.002	0.983	0.001	0.013
397	IC_534859	0.001	0.001	0.001	0.997
398	IC_534864	0.001	0.994	0.001	0.003
399	IC_534870	0.002	0.004	0.001	0.993
400	IC_534871	0.001	0.994	0.002	0.003
401	IC_534873	0.002	0.004	0.001	0.993
402	IC_534883	0.001	0.001	0.997	0.001
403	IC_534884	0.69	0.307	0.001	0.002
404	IC_534885	0.003	0.003	0.001	0.993
405	IC_534886	0.981	0.016	0.001	0.002
406	IC_534887	0.003	0.001	0.001	0.994
407	IC_55507	0.006	0.001	0.003	0.99
408	IC_55578	0.002	0.001	0.002	0.995
409	IC_55593	0.997	0.001	0.001	0.001
410	IC_55617	0.001	0.001	0.001	0.996
411	IC_55636	0.001	0.001	0.001	0.997
412	IC_55652	0.004	0.002	0.001	0.994
413	IC_55656	0.944	0.052	0.001	0.003
414	IC_55657	0.074	0.915	0.003	0.009
415	IC_55659	0.995	0.002	0.001	0.002

Contd…

A.11 Contd…

Label	accessions	FST1	FST2	FST3	FST4
416	IC_55664	0.001	0.001	0.001	0.997
417	IC_55685	0.001	0.001	0.001	0.997
418	IC_57844	0.997	0.001	0.001	0.001
419	IC_57983	0.002	0.96	0.035	0.003
420	IC_57998	0.02	0.001	0.002	0.976
421	IC_59179	0.002	0.001	0.001	0.996
422	IC_59191	0.026	0.002	0.003	0.969
423	IC_59511	0.997	0.001	0.001	0.002
424	IC_59544	0.995	0.003	0.001	0.001
425	IC_59612	0.997	0.001	0.001	0.001
426	IC_64238	0.001	0.002	0.997	0.001
427	IC_73198	0.001	0.001	0.997	0.001
428	IC_73201	0.001	0.001	0.997	0.001
429	IC_73205	0.001	0.001	0.997	0.001
430	IC_73207	0.238	0.196	0.42	0.146
431	IC_73210	0.243	0.217	0.392	0.148
432	IC_73215	0.724	0.275	0.001	0.001
433	IC_73232	0.002	0.996	0.001	0.001
434	IC_73493	0.001	0.997	0.001	0.001
435	IC_75327	0.001	0.997	0.001	0.001
436	IC_75328	0.001	0.996	0.001	0.002
437	IC_75333	0.085	0.052	0.001	0.862
438	IC_75339	0.001	0.997	0.001	0.001
439	IC_75351	0.001	0.001	0.001	0.998
440	IC_78096	0.001	0.985	0.001	0.013
441	IC_78707	0.763	0.235	0.001	0.001
442	IC_78728	0.254	0.744	0.001	0.001
443	IC_78715	0.001	0.001	0.001	0.998
444	IC_78824	0.001	0.001	0.001	0.998
445	IC_78828	0.001	0.001	0.001	0.998
446	IC_78832	0.291	0.693	0.001	0.015
447	IC_78836	0.001	0.001	0.001	0.998
448	IC_78837	0.001	0.001	0.001	0.998
449	IC_78839	0.001	0.001	0.001	0.998
450	IC_78843	0.001	0.001	0.001	0.998
451	IC_78854	0.001	0.001	0.001	0.998
452	IC_78859	0.001	0.001	0.001	0.998
453	IC_78860	0.001	0.966	0.001	0.033
454	IC_78862	0.001	0.001	0.001	0.997

Contd…

A.11 Contd…

Label	accessions	FST1	FST2	FST3	FST4
455	IC_78865	0.001	0.001	0.001	0.998
456	IC_78869	0.001	0.001	0.001	0.998
457	IC_78872	0.991	0.006	0.001	0.002
458	IC_78877	0.001	0.993	0.001	0.005
459	IC_78888	0.001	0.001	0.001	0.998
460	IC_78889	0.003	0.001	0.003	0.992
461	IC_78890	0.004	0.002	0.001	0.993
462	IC_78891	0.01	0.002	0.002	0.986
463	IC_78895	0.001	0.001	0.001	0.997
464	IC_78897	0.001	0.001	0.001	0.997
465	IC_78899	0.001	0.001	0.001	0.998
466	IC_78901	0.001	0.001	0.001	0.998
467	IC_78905	0.001	0	0.001	0.998
468	IC_78908	0.001	0	0.001	0.998
469	IC_78911	0.001	0	0.001	0.998
470	IC_78912	0.001	0.001	0.001	0.998
471	IC_78914	0.001	0	0.001	0.998
472	IC_78915	0.001	0	0.001	0.998
473	IC_78918	0.001	0	0.001	0.998
474	IC_78919	0.001	0.001	0.001	0.998
475	IC_78920	0.001	0.001	0.001	0.997
476	IC_78923	0.001	0.001	0.997	0.001
477	IC_78924	0.001	0.001	0.997	0.001
478	IC_78925	0.002	0.001	0.995	0.002
479	IC_78927	0.996	0.001	0.001	0.002
480	IC_78928	0.991	0.004	0.002	0.003
481	IC_78930	0.013	0.002	0.002	0.983
482	IC_78931	0.001	0.996	0.003	0.001
483	IC_78933	0.001	0.987	0.001	0.012
484	IC_78935	0.001	0.001	0.001	0.998
485	IC_78936	0.001	0	0.001	0.998
486	IC_78937	0.001	0.002	0.994	0.003
487	IC_78940	0.001	0.001	0.996	0.002
488	IC_78941	0.001	0	0.001	0.998
489	IC_78945	0.001	0.001	0.001	0.998
490	IC_78948	0.001	0	0.001	0.998
491	IC_78960	0.001	0.001	0.001	0.998
492	IC_78968	0.998	0.001	0.001	0.001
493	IC_78981	0.998	0.001	0.001	0.001

Contd…

A.11 Contd…

Label	accessions	FST1	FST2	FST3	FST4
494	IC_78982	0.998	0.001	0.001	0.001
495	IC_78987	0.998	0	0.001	0.001
496	IC_78990	0.997	0	0.002	0.001
497	IC_78991	0.567	0.003	0.172	0.258
498	IC_78996	0.994	0	0.005	0.001
499	IC_78999	0.001	0.996	0.002	0.001
500	IC_79022	0.001	0.001	0.001	0.997
501	IC_79023	0.001	0.001	0.001	0.998
502	IC_79026	0.001	0.001	0.997	0.001
503	IC_79028	0.987	0.01	0.001	0.001
504	IC_79038	0.002	0.002	0.995	0.001
505	IC_79039	0.001	0.001	0.997	0.001
506	IC_79040	0.001	0.002	0.996	0.001
507	IC_79041	0.002	0.996	0.001	0.001
508	IC_79042	0.002	0.001	0.997	0.001
509	IC_79043	0.001	0.998	0.001	0.001
510	IC_79046	0.995	0.002	0.001	0.001
511	IC_79047	0.002	0.996	0.001	0.001
512	IC_79050	0.994	0.002	0.002	0.002
513	IC_79052	0.314	0.683	0.001	0.001
514	IC_79053	0.001	0.001	0.997	0.001
515	IC_79055	0.001	0.004	0.987	0.008
516	IC_79056	0.998	0	0.001	0.001
517	IC_79062	0.002	0.001	0.997	0.001
518	IC_79063	0.696	0.302	0.001	0.001
519	IC_79065	0.998	0.001	0	0.001
520	IC_79066	0.001	0.001	0.001	0.998
521	IC_79067	0.001	0.001	0.998	0.001
522	IC_79068	0.998	0.001	0.001	0.001
523	IC_79077	0.998	0	0	0.001
524	IC_79079	0.001	0.001	0.001	0.998
525	IC_79080	0.001	0.001	0.001	0.998
526	IC_79083	0.001	0.001	0.001	0.998
527	IC_79085	0.998	0	0	0.001
528	IC_79090	0.998	0	0	0.001
529	IC_79091	0.998	0	0.001	0.001
530	IC_79095	0.996	0.002	0.001	0.001
531	IC_79100	0.001	0.002	0.997	0.001
532	IC_79102	0.001	0.001	0.001	0.997

Contd…

A.11 Contd…

Label	accessions	FST1	FST2	FST3	FST4
533	IC_79106	0.002	0.001	0.996	0.001
534	IC_79107	0.997	0.001	0.001	0.001
535	IC_79108	0.996	0.001	0.001	0.002
536	IC_79110	0.001	0.029	0.968	0.001
537	IC_79940	0.001	0.004	0.994	0.001
538	IC_82116	0.001	0.994	0.004	0.001
539	IC_82118	0.003	0.002	0.994	0.001
540	IC_82119	0.002	0.032	0.001	0.966
541	IC_82125	0.001	0.997	0.001	0.001
542	IC_82126	0.001	0.997	0.001	0.001
543	IC_82127	0.014	0.982	0.001	0.003
544	IC_82130	0.001	0.001	0.001	0.997
545	IC_82132	0.027	0.001	0.971	0.001
546	IC_82133	0.001	0.997	0.001	0.001
547	IC_82134	0.001	0.997	0.001	0.001
548	IC_82135	0.001	0.997	0.002	0.001
549	IC_82136	0.001	0.997	0.001	0.001
550	IC_82145	0.001	0.997	0.001	0.001
551	IC_82146	0.001	0.001	0.998	0
552	IC_82147	0.001	0.007	0.991	0
553	IC_82156	0.001	0.001	0.997	0.001
554	IC_82157	0.001	0.001	0.998	0
555	IC_82158	0.001	0.001	0.997	0.001
556	IC_82159	0.002	0.002	0.995	0.001
557	IC_82160	0.002	0.002	0.996	0.001
558	IC_82161	0.567	0.374	0.001	0.058
559	IC_82163	0.001	0.997	0.001	0.001
560	IC_82164	0.001	0.996	0.001	0.001
561	IC_82165	0.002	0.002	0.994	0.002
562	IC_82167	0.003	0.001	0.995	0.001
563	IC_82169	0.016	0.411	0.008	0.565
564	IC_82177	0.997	0.001	0.001	0.001
565	IC_82179	0.001	0.001	0.998	0.001
566	IC_82180	0.995	0.002	0.001	0.003
567	IC_82181	0.001	0.998	0.001	0.001
568	IC_82185	0.595	0.398	0.001	0.005
569	IC_82187	0.725	0.27	0.001	0.004
570	IC_82189	0.001	0.998	0.001	0.001
571	IC_82190	0.996	0.002	0.001	0.001

Contd…

A.11 Contd…

Label	accessions	FST1	FST2	FST3	FST4
572	IC_82192	0.002	0.019	0.001	0.978
573	IC_82193	0.001	0.997	0.001	0.001
574	IC_82195	0.978	0.014	0.002	0.006
575	IC_82197	0.979	0.018	0.001	0.001
576	IC_82198	0.002	0.096	0.001	0.901
577	IC_82199	0.002	0.001	0.001	0.997
578	IC_82200	0.409	0.588	0.001	0.002
579	IC_82204	0.002	0.002	0.001	0.996
580	IC_82206	0.993	0.003	0.002	0.001
581	IC_82210	0.394	0.602	0.001	0.002
582	IC_82236	0.001	0.998	0.001	0.001
583	IC_82247	0.001	0.998	0.001	0.001
584	IC_82256	0.997	0.001	0.001	0.001
585	IC_82257	0.996	0.003	0.001	0.001
586	IC_82259	0.008	0.975	0.001	0.016
587	IC_82263	0.001	0.997	0.001	0.001
588	IC_82285	0.001	0.997	0.001	0.001
589	IC_82286	0.502	0.003	0.271	0.224
590	IC_82303	0.548	0.004	0.207	0.241
591	IC_82385	0.001	0.002	0.995	0.002
592	IC_82338	0.514	0.008	0.221	0.257
593	IC_82342	0.35	0.648	0.001	0.001
594	IC_82344	0.023	0.964	0.001	0.012
595	IC_82367	0.997	0.001	0.001	0.001
596	IC_82369	0.511	0.006	0.478	0.005
597	IC_82369	0.001	0.004	0.993	0.002
598	IC_82371	0.003	0.993	0.001	0.003
599	IC_82372	0.005	0.993	0.001	0.001
600	IC_82375	0.255	0.743	0.001	0.001
601	IC_82377	0.001	0.006	0.001	0.991
602	IC_82381	0.19	0.803	0.001	0.006
603	IC_82385	0.008	0.975	0.002	0.015
604	IC_82386	0.001	0.001	0.997	0.001
605	IC_82387	0.001	0.001	0.997	0.001
606	IC_82388	0.002	0.012	0.003	0.983
607	IC_82393	0.001	0.001	0.997	0.001
608	IC_82394	0.005	0.056	0.938	0.001
609	IC_82398	0.001	0.001	0.997	0.001
610	IC_82400	0.028	0.101	0.002	0.868

Contd…

A.11 Contd…

Label	accessions	FST1	FST2	FST3	FST4
611	IC_82410	0.008	0.002	0.001	0.988
612	IC_82421	0.001	0.988	0.001	0.01
613	IC_82425	0.001	0.12	0.007	0.872
614	IC_82426	0.002	0.986	0.001	0.011
615	IC_82432	0.002	0.829	0.001	0.168
616	IC_82440	0.001	0.003	0.001	0.995
617	IC_82221	0.005	0.988	0.001	0.006
618	IC_82555	0.001	0.001	0.997	0.001

PUBLICATION

Indian J. Plant Genet. Resour. 30(1): 80-83 (2017)
DOI 10.5958/0976-1926.2017.00012.2

SHORT COMMUNICATION

High Kernel Weight Genotypes of Indian Bread Wheat (*Triticum aestivum* L.)

HH Kumaraswamy[1], GP Singh[2] and NK Singh[1]
[1]*ICAR-National Research Centre on Plant Biotechnology, New Delhi-110012, India*
[2]*ICAR-Indian Institute of Wheat and Barley Research, Gahoon Vihar, Karnal, Haryana-132001 India*

(Received: 20 February 2015; Revised: 3 August 2016; Accepted: 4 August 2016)

In an experiments on assessment and identification of elite Indian wheat genotypes, IC138896, IC138893, IC28664, IC35069, IC532841, IC534435, IC534777, IC534787, IC82197, and IC82198, showed high kernel weight compared to 'Moti (HD-1949)'. These can be gainfully used as parental lines in breeding for increasing grain weight, both for hybridization and selection.

Key Words: Indian wheat, High kernel weight, Parental lines, Phenological characteristics, *Triticum aestivum*

Wheat (*Triticum* spp. L.) is one of the most important cereal crops that was crucial in the rise of famous civilizations and has been the chief vehicle of green revolution across the world. It has been playing the key role in India's food security and self-sufficiency.

India's rate of gain in wheat production has declined during the last two years due to adverse weather conditions (Singh, 2016). On the other hand, strong selection pressure for yield under high input conditions during and after green revolution has eliminated considerable genetic diversity in the breeding pools of major crops, including wheat, causing erosion of genetic potential for adaptation to emerging challenges of climate change (Voss-Fels *et al*., 2015). In addition, the bottlenecks of hexaploidisation followed by breeding have considerably narrowed down the bread wheat (*Triticum aestivum* L.) diversity to the extent that yields in many regions appear to be unexpectedly stagnating (Voss-Fels *et al*., 2015). Genetic diversity analyses and identification of elite genotypes can address this dilemma by providing detailed knowledge to characterize and utilize the available genetic diversity in the germplasm collections.

As compared to modern high yielding varieties, local cultivars and landraces have higher genetic variability and adaptation to the different local stress conditions and, therefore, are valuable tools for identifying genes for achieving high grain yield in arid and semi-arid areas (Lammerts-van-Bueren *et al*., 2005). Worldwide efforts are on to assess the phenotypic diversity of local landraces in order to mine useful genes and to assess the exploitable diversity, in wheat and other crops to utilize this untapped diversity: for instance, Dikshit *et al*. (2004) and Routray *et al*. (2007) reported on rice landraces and Uttaranchal wheat landraces, respectively. While genetic diversity can be estimated based on various criteria, agro-morphological characterization is a primary step for conservation and utilization of plant genetic resources (Zarkti *et al*., 2012).

In order to generate information, useful for wheat crop improvement through breeding and/or biotechnology, six hundred and eighteen Indian wheat collections, supplied with thanks by Late Dr SK Mishra, ICAR-National Bureau of Plant Genetic Resources (ICAR-NBPGR), New Delhi, were assessed for morphological and phenological characteristics. These were grown at the experimental farm of ICAR-Indian Agricultural Research Institute, New Delhi (latitude: 80°38'23"N, longitude:77°09'27"E, and with an elevation of 228.61m above msl) for two consecutive cropping seasons (2011-12 and 2012-13), in two replicates using randomized block (RBD) design, following standard agronomic practices. Data were recorded on five randomly tagged plants in each plot. Recorded value of each trait was averaged over 5 individuals, two replications, and two cropping seasons. Data were analyzed using statistical software SAS, version 9.3 (SAS Institute Inc., USA, 2011).

**Author for Correspondence: Email- hhkswamy@gmail.com, hhk.swamy@icar.gov.in*

Statistical parameters such as standard error mean (SEm), critical difference (CD) and coefficient of variation (CV, %) were estimated for the entire set of 618 genotypes and were used for the comparison. In the present study, as presented in Table 1, performances of these ten genotypes were compared with that of commercial cultivar HD-1949 (IC73210). Selection of superior wheat genotypes regarding high grain yield would be as effective as selection for its components, including thousand seed weight (Peltonen-Saino *et al*., 2007). The relation of these traits with grain yield and the interrelationships have special importance as the basis for selecting high yielding genotypes (Sabaghnia *et al*., 2014). In the set of 618 Indian wheat accessions, IC73210 was a commercial cultivar Moti (HD-1949), which is a triple dwarf wheat with grain color amber and medium bold, good for chapati making (www.dacnet.nic.in) and, therefore, ten high kernel weight genotypes identified in the present study were compared with HD-1949.

The genotype IC5344777, as shown in Table 1, registered the highest thousand kernel weight (65.5 g in 2011-12, and 69.99 g in 2012-13) amounting to more than the double of the commercial cultivar HD-1949 (32.5 g in 2011-12, and 30.5 g in 2012-13) during both the years followed by genotypes IC534787 (58.6 g in 2011-12, and 56.37 g in 2012-13), IC82197 (48.2 g in 2011-12, and 46.54 g in 2012-13), IC35069 (45.8 g in 2011-12, and 44.6 g in 2012-13), IC138893 (45.05 g in 2011-12, and 44.52 g in 2012-13), IC532841 (43.3 g in 2011-12, and 41.84 g in 2012-13), IC82198 (41.5 g in 2011-12, and 41.88 g in 2012-13), IC138896 (41.23 g in 2011-12, and 42.84 g in 2012-13), IC534435 (40.25 g in 2011-12, and 42.05 g in 2012-13), and IC28664 (40 g in 2011-12, and 41.08 g in 2012-13). Our result showed higher values than those of previous reports for bread wheat (Sabaghnia *et al*., 2014). In terms of days to head emergence, days to 50% flowering, and days to physiological maturity; most of these genotypes were comparable to the HD-1949, except IC534787 that showed the longest duration for physiological maturity (155 days in 2011-12 and 150 days in 2012-13) followed by IC534777 (140 days in 2011-12, and 145 days in 2012-13), and IC82198 (100 days in both the years). IC28664, IC534435, IC82197, and IC82198 produced 50% flowers in 5 days after head emergence, while the other genotypes took 10 to 15 days. Our results were congruent with the report of Zarkti *et al.* (2012). In case of average spike length, as depicted in Fig. 1,

Table 1. Comparative performance of twelve high-kernel-weight genotypes of Indian bread wheat (*Triticum aestivum* L.)

Genotypes	Thousand kernel weight (g)		Days to physiological maturity		Days to 50% flowering		Days to head emergence		Plant height (cm)		Spike length (cm)		Number of seeds per spike	
	2011-12	2012-13	2011-12	2012-13	2011-12	2012-13	2011-12	2012-13	2011-12	2012-13	2011-12	2012-13	2011-12	2012-13
IC_73210 (Moti/ HD-1949)	32.5	30.5	90	90	70	70	55	60	53.25	55.6	8.3	8.5	35	33
IC_138896	41.23	42.84	90	90	70	70	60	65	92.5	90.3	10.7	9.5	39	37
IC_138893	45.05	44.52	100	95	75	70	60	60	95.25	95.55	9.8	9..5	44	43
IC_28664	40	41.08	90	90	70	70	65	65	80	82.1	12	12.5	49	47
IC_35069	45.8	44.6	95	90	75	70	60	65	110.4	109.1	12.5	12	45	43
IC_532841	43.5	41.84	95	90	70	70	60	60	91.5	89.05	8	8.5	25	27
IC_534435	40.25	42.05	95	95	70	70	65	65	106	106.05	10.5	10	56	57
IC_534777	65.5	69.99	140	145	110	120	110	110	102.6	103.6	10.5	10.5	42	38
IC_534787	58.6	56.37	155	150	115	120	110	110	98.07	99.85	12.8	12.5	30	42
IC_82197	48.2	46.54	100	95	75	70	65	65	91.5	91.85	11.2	11.5	55	53
IC_82198	41.5	41.88	100	100	75	70	65	65	80.04	78.95	10.3	9.5	58	61
SE(m)	0.13	0.12	0.38	0.35	0.06	0.05	0.04	0.04	0.03	0.03	0.04	0.05	0.9	0.87
CD (0.05%)	0.36	0.33	1.05	0.97	0.17	0.14	0.11	0.11	0.08	0.08	0.11	0.14	2.49	2.41
CD (0.01%)	0.47	0.44	1.38	1.27	0.22	0.18	0.15	0.15	0.11	0.11	0.15	0.18	3.28	3.17
CV (%)	25.92	28.37	19.21	19.36	23.01	24.67	25.61	27.28	8.56	9.02	16.95	18.58	37.94	39.27

Fig. 1. Spike characteristics of high kernel weight genotypes of Indian bread wheat: A, IC82198; B, IC534787; C, IC82197; D, IC28664; E, IC534777; F, IC534435; G, IC138893; H, IC138896; I, IC532841; and J, IC35069.

and presented in Table 1, longer spikes ranging from 12 to 12.8 cm were recorded in genotypes IC28664, IC35069, and IC534787; whereas, the shorter spikes were observed in genotype IC532841 (8 cm in 2011-12, and 8.5 cm in 2012-13) similar to those of commercial cultivar HD-1949 (8.3 cm in 2011-12 and 8.5 cm in 2012-13). Number of seeds per spike is another important component of grain yield. As shown in Table 1, the highest number of seeds per spike was registered in IC82198 (58 in 2011-12 and 61 in 2012-13) and the lowest (25 in 2011-12 and 27 in 2012-13) values that were below those of commercial cultivar HD-1949 (35 in 2011-12 and 33 in 2012-13) were registered in IC532841. According to Subedi *et al*. (2000), when wheat is stressed during reproductive development, the grain number in fertile florets is reduced. This explains the reason for variation in number of seeds per spike within each genotype between the two cropping seasons (2011-12 and 2012-13).

Figure 1 illustrates, further, visual/qualitative characteristics of spikes of these high kernel weight genotypes. Five genotypes were awnless and the other five were awned; spike density was compact in most of the genotypes except IC534787 which exhibited lax

type; most of them were yellow in glume color except IC534787 which was brown; and, five (IC82198, IC534787, IC82197, IC534777, IC532847, and 35069) were pubescent and the other five lacked glume hairiness. Glume hairiness is another important trait of breeding value as it influences reaction of plant, particularly spike portion, to fungal infection such as powdery mildew.

In wheat breeding, thousand seed weight is one of the important parameters in that it increases seed germination percent, seedling emergence, tillering capacity, spike and yield (Sabaghnia *et al*., 2014). The results of this study will support efforts of conservation and utilization of Indian wheat landraces in wheat genetic improvement programmes as these high kernel weight genotypes of Indian bread wheat, which are already adapted to local condition due to continuous cultivation and selection for long time by farmers, can be gainfully used by breeders and biotechnologists as parental lines in breeding for increasing grain weight, both for hybridization and selection, as well as in developing mapping population to tag genomic regions for high seed weight.

Acknowledgements

We acknowledge, with thanks, the facility and financial support from ICAR-NPTC project, Post Graduate School, ICAR-Indian Agricultural Research Institute, New Delhi and ICAR-National Research Centre for Plant Biotechnology, New Delhi; the kind supply of plant materials used in this research by late Dr SK Mishra, ICAR-National Bureau of Plant Genetic Resources; land facility for field experiments provided by Division of Genetics, ICAR-Indian Agricultural Research Institute; and the study leave for the first author granted by Director, ICAR-IIOR, Hyderabad, and Indian Council of Agricultural Research, New Delhi.

References

Dikshit N, N Sivaraj and AB Das (2014) Genotypic variations in agro-morphological and physico-chemical traits in ethnic rice (*Oryza sativa*) landracs of orissa. *Indian J. Agric. Sci.* **84(8):** 987-992.

Lammerts-van-Bueren ET, LJM Van-Soest, EC De-Groot, IW Boukema and AM Osman (2005) Broadening the genetic base of onion to develop better-adapted varities for organic farming systems. *Euphytica* **146:** 125-132.

Peltonen-Sainioa P, A Kangas, Y Salo and L Jauhiainen (2007) Grain number dominates grain weight in temperate cereal yield determination: Evidence based on 30 years of multilocation trials. *Field Crop Res.* **100:** 179-188.

Routray P, O Basha, M Garg, NK Singh and HS Dhalikal (2007) Genetic diversity of landraces of wheat (*Triticum aestivum* L.) from hilly areas of Uttaranchal, India. *Genet. Resour. Crop Evol.* **54:** 1315-1326.

Sabaghnia N, M Janmohammadi and AB Segherloo (2014) Evaluation of some agro-morphological traits diversity in Iranian bread wheat genotypes. *Ann. Univ. Mariae Curie-sklodovska* **69(1):** 79-92.

Singh, SK (2016) India: Grain and feed annual (Grin report numer, IN6033). www.gain.fas.usda.gov. (Accessed on May 13, 2016, Please see the URL, http:// gain.fas.usda.gov/Recent%20GAIN%20Publications/Grain%20and%20Feed%20Annual_New%20 Delhi_India_2-26-2016.pdf).

Subedi, KD., PJ Gregory, RJ Summerfield and MJ Goodig (2000) Pattern of grain set in boron-deficient and cold-stressed wheat (*Triticum aestivum* L.). *J. Agric.Sci., Cambridge*, **134**: 25-31.

Voss-Fels K, M Frisch, L Qian, S Kontowski, W Friedt, S Gottwald and RJ Snowdon (2015) Subgenomic diversity patterns caused by directional selection in bread wheat gene pools. *Pl. Gen.* **8(2):** 1-13.

Zarkti H, H Quabbou, SM Udupa, F Gaboun and A Hilali (2012) Agro-morphological variability in durum wheat landraces of Morocco. *Aust. J. Crop Sci.***6(7):** 1172-1178.