Chapter 1
Wheat Evolution, Domestication, and Improvement

Perry Gustafson, Olga Raskina, XueFeng Ma, and Eviatar Nevo

SUMMARY

(1) Wheat is the world’s largest and most important food crop for direct human consumption; therefore, continued wheat improvement is paramount for feeding an ever-increasing human population.

(2) Wheat improvement is tightly associated with the characterization and understanding of wheat evolution and the genetic diversification of various wheat species and relatives. The evolution of the genus *Triticum* mainly resulted from inter- and intraspecific hybridization, polyploidization, and recurrent formation of wheat and its wild relatives.

(3) An understanding of the process of species domestication and genome evolution that has occurred and is still occurring within the primary, secondary, and tertiary gene pools is critical for further exploring the improvement of wheat production.

(4) Also critical is to evaluate the relative importance of the evolutionary processes, such as the pivotal genome concept and cell cycle differences, driving and shaping the structural rearrangements and genomic changes occurring within the genomes of polyploid wheat.

INTRODUCTION

Cereals, including wheat (*Triticum* spp.), rice (*Oryza sativa* L.), and maize (*Zea mays* L.), are the major food crops for all humans and are the principal resources that have led to the emergence of human civilization as we know it today. Domestication of cereals during the past 10 millennia is one of the most dramatic demonstrations involving humans’ manipulation of the evolutionary processes of speciation, natural selection, and adaptation. Plant domestication revolutionized human cultural evolution and is primarily responsible for the advances that have occurred in our civilization. A post-Pleistocene global temperature increase following the ice age may have induced the expansion of economically important thermophilous plants, which in turn promoted the change from a world of hunter–gatherer societies to complex foraging and plant-cultivating societies. The shift from foraging to steady production definitely led to the occurrence of an incipient agriculture in many parts of the world and to a decline in genetic diversity of the world’s crops, which has been accelerated by domestication and the recent breeding of modern cultivars.

The evolution of domestication should be considered in two main contexts: (i) the evolution of new species of crop plants by humans through strong artificial selection (see Darwin 1859); (ii) the evolution of human civilizations, and the current consequences of population explosion and increasing world hunger in developing countries. Both contexts provide fascinating insight into the evolutionary process.
WHEAT DOMESTICATION AND HUMAN CIVILIZATION

The earliest signs of crop domestication appeared 10,000–12,000 years ago in the Fertile Crescent of the Near East, in Central America, and in southern China, involving different crops and independent cradles of domestication. Cereal domestication was founded, in the Fertile Crescent of the Near East, on crop reliability, yield, and suitability for storage. Recent botanical, genetic, and archaeological evidence has pointed to a small core area within the Fertile Crescent—near the upper reaches of the Tigris and Euphrates rivers, in present-day southeastern Turkey—northern Syria—as the cradle of cereal agriculture (Lev-Yadun et al., 2000). Further evidence is needed to clarify when and where wheat domestication and agriculture, as driving forces of modern civilization, originated. Was it spread in time or space, or both, in the Fertile Crescent?

The genetic changes required for wheat domestication to occur were relatively straightforward and rapid, including selection for nonshattering, free-threshing, nonbrittle rachis and hull-less spike characteristics and for higher yield. In all cereals, the gene complexes for ease of harvest, yield, and suitability for short- and long-term storage have been critical for domestication. The domestication of cereals was essential for human populations’ change to an agriculture-based society. Evolution of any crop species from their wild progenitors to full domestication, and the emergence of agricultural ecologies from preagricultural ones, clearly established human movement from hunter–gatherer societies to sedentariness, urbanization, culture, and an unprecedented population explosion (Harlan 1975, p. 295).

Considerable progress has been achieved in characterizing the wild ancestry of Old World crops, including cereals. The wild progenitors of most of our cultivated plants have been satisfactorily identified by comparative morphology and genetic analysis. The distribution and ecological ranges of wild relatives have also been established. Furthermore, comparisons between wild types and their cultivated counterparts have revealed many of the evolutionary changes that resulted in domestication. Research has enabled us to assess the relative importance of the evolutionary forces driving wheat evolution—hybridization, migration, drift, and natural selection—interacting in generating the contemporary wheat genotype. Studies suggested that, besides polyploid hybridization, natural selection played a large role and oriented wheat evolution primarily through the mechanisms of diversifying and balancing selection regimes (Nevo et al., 2002).

Wheat has become the world’s largest and most important food crop for direct human consumption, with an annual harvest of more than 620 million tonnes produced in over 40 countries for more than 35% of the global population (Williams 1993). The US produces approximately 55 to 60 million tonnes per year and supplies about 40% of the world’s exports. Wheat is currently grown from 67°N, in Norway, Finland, and Russia, to 45°S, in Argentina and Chile. The world’s main wheat-producing regions are in temperate and southern Russia, the central plains of the US, southern Canada, the Mediterranean Basin, northern China, India, Argentina, and Australia. Wheat makes up 29%–30% of the world’s total cereal production and is humans’ most important source of protein. As a crop for direct human consumption, only rice comes close to matching wheat production. As a food grain, wheat is the major dietary component throughout the world; in 1996 it served as the source of over 55% of the world’s carbohydrates (http://www.fao.org).

Wheat cultivars are superior to most other cereals in their nutritive value (Zohary and Hopf 2000). Besides the grain containing from 60% to 80% starch, it also contains from 7% to 22% storage protein, which in elite wild genotypes can reach as much as 17% to 28% (Avivi 1978, 1979; Avivi et al., 1983; Grama et al., 1983; Nevo et al., 1986; Levy and Feldman 1987). The gluten proteins in the seed endosperm impart unique bread-baking qualities to wheat dough, which has made wheat the staple food in the ancient and modern world for billions of people. Only minor amounts of wheat are occasionally used as animal feed, with the amount being highly dependent
on wheat prices compared with other feed grains. A very small portion of the world’s wheat and wheat flour is used for industrial purposes such as starch and gluten production (Morrison 1988).

Global efforts to increase wheat production and to keep up with population growth and rising demand have been relatively successful in maintaining a steady increase in wheat yield, representing roughly a threefold increase over production levels of the 1960s. It should be noted that, despite dramatic increases in global wheat production, in 2003 more than 800 million people in the world suffered daily from severe undernourishment and hunger (http://www.fao.org). Protein deficiency is one of the most serious problems and threatens to become a real nutritional disaster in the near future, primarily in Asia and Africa, where about 80% of the human diet is protein supplied by plants. The urgent need to increase high-quality protein sources is exacerbated by major problems affecting cultivated crops, including the cereals, with respect to the reduction in genetic diversity (Plucknett et al., 1983, 1987).

In the future, the nutritional composition of the world wheat supply will become even more critical as world demand for wheat continues to grow and world wheat stocks continue to decrease. World population, which currently stands at well over 6 billion, was projected in 2001 to reach 8.3 billion by the year 2030 and 9.3 billion by 2050 (http://www.fao.org). Income growth and urbanization, which are shifting consumer preference away from rice, coarse grains, and tubers to more wheat-based food products and meat, are also expected to continue to increase in many developed and developing countries. World demand is projected to require approximately a 66% increase in agricultural production by 2040. In addition, our ability to bring more land into wheat cultivation is rapidly diminishing due to population growth, environmental pressures, and the increasingly limited availability of arable land (Young 1999). The need for future improvement in wheat production will clearly coincide with a loss of flexibility and availability of traditional resources.

The success of wheat improvement programs to meet future demands will require complementing the traditional breeding approaches with innovative, nontraditional methodologies that will enhance genetic variation in wheat. One of many approaches to improving wheat production will be the manipulation of secondary and tertiary gene pools for new sources of biotic and abiotic stress tolerance. A key to the successful manipulation of the primary, secondary, and tertiary gene pools is to fully understand the evolution of the cultivated wheat species. Unfortunately, some loss of genetic diversity involving most of the world’s crops, including wheat, has accelerated in recent decades. The dynamic conservation of wheat germplasm and wild wheat relatives offers one of the best hopes for sustained wheat improvement (Nevo 1998). However, it is clear that conservation of germplasm is not the only answer. To achieve a more efficient and comprehensive utilization of the gene pools of wheat and wild wheat relatives, it is critical that we learn how to predict, screen, manipulate, maintain, and properly evaluate genetic diversity and resources (Nevo 2001). After all, plant breeding is basically an accelerated manipulation of natural evolution. Once we understand the evolutionary processes involved in the formation and stabilization of wheat, we can better design wheat improvement programs that will enable a more efficient restructuring of gene complexes within and between wheat, wheat-related species, and genera to capitalize on the value-added traits that may be economically important for wheat improvement.

WHEAT CULTIVATION

Until the late 19th century, all cultivated wheat existed as highly heterogeneous landraces. Wheat cultivars were morphologically uniform mixtures of inbred lines and hybrid segregates, the products of low levels of random crossing within a landrace. Any artificial selection was primarily for increased yield, larger seed size, better flour quality, and adaptation to a wider range of climatic and farming regimes (Feldman et al., 1995). Many landraces still exist today, in fields in many
regions around the world and in germplasm collections. Over the past century of modern breeding, attempts to produce cultivars that meet the advanced agriculture demands of an ever-increasing population has resulted in the landraces being almost wholly displaced by genetically uniform cultivars. The result of modern agriculture has been a marked narrowing of the genetic base in probably all advanced agricultures (Harlan 1975, 1976, 1992). While wheat yields have kept up with population demands in advanced agricultures (e.g., Avery 1985), genetic homogeneity has also dramatically increased due to modern agricultural practices (Frankel and Bennett 1970; Frankel and Hawkes 1975; Harlan 1975, 1976, 1992; Frankel and Soulé 1981; Nevo et al., 1982; Nevo 1983, 1986, 1989, 1995, 1998, 2001; Plucknett et al., 1983, 1987; Lupton 1987; Nevo and Beiles 1989). Consequently, the genetic base of many cultivated crops, including wheat, has been narrowed and placed under serious risk (Frankel and Soulé 1981; Plucknett et al., 1987). A global network of gene banks has been established to provide plant breeders with the genetic resources for maintaining germplasm collections and for developing more resistant and tolerant crops that will improve production (Lupton 1987; Plucknett et al., 1987; Brown et al., 1989, 1990).

Dynamic in situ conservation of landraces and wild relatives, the best hope for improving cultivated plants (Feldman and Sears 1981), is being actively discussed as an optimal conservation strategy (Hawkes 1991; Heyn and Waldman 1992; Valdes et al., 1997; Maxted et al., 1997; Nevo 1998). Just as important as the conservation of diverse germplasm is the achievement of a more efficient and comprehensive utilization of conserved wild gene pools. It is essential to be able to efficiently predict, screen, and evaluate promising genetic diversity and resources, thereby optimizing crop improvement (Nevo 1983, 1989, 1992, 1995, 2001; Peng et al., 2000a,b,c). The analysis of genetic diversity across the geographic range, at both the macro- and microscale, will unravel patterns and forces driving wheat genome evolution and lay open the full potential of its genetic resources for utilization.

**ORIGIN, DOMESTICATION, AND EVOLUTION OF WHEAT**

Modern wheat cultivars belong primarily to two polyploid species: hexaploid bread wheat \([T. aestivum\ (2n = 6x = 42 chromosomes)]\) and tetraploid hard or durum-type wheat \([T. turgidum\ L. \ (Thell.)\ (2n = 4x = 28)]\) used for macaroni and low-rising bread. The cultivated diploid species \(T. monococcum\ L.\) einkorn wheat \((2n = 2x = 14)\) is a relic and is only found in some mountainous Mediterranean regions. Wheat is predominantly self-pollinated; hence, genetic diversity is represented in the wild by numerous clones, in vast national and international germplasm collections, and in current cultivation by some 25,000 different cultivars. Wild and primitive wheat forms have hulled grains and brittle ears that disarticulate at maturity into individual spikelets, with each spikelet having a wedge-shaped rachis internode at its base, and an arrowlike device that inserts the seed into the ground (Zohary 1969). By contrast, all cultivated wheat forms have nonbrittle ears that stay intact after maturation, thus depending on humans for reaping, threshing, and sowing. The nonbrittleness and nakedness of cultivated wheat is controlled by the \(Q\) locus (Luo et al., 2000), located on chromosome 5 of genome A, which may have arisen from the \(q\) gene of the hulled varieties by a series of mutations (Feldman et al., 1995).

**Polyploidy, a form of plant evolution**

The evolution of the genus *Triticum* serves as one of the best models of polyploidy, one of the most common forms of plant evolution (Elder and Turner 1995; Soltis and Soltis 1999). The gradual shift to a steady-production-based agriculture has been the main driving force behind the domestication of wheat. The evolution of domestication can be considered as the evolution of new crop species by natural and artificial selection, and the evolution of human civilization as we know it. Unfortunately, this has resulted in a massive population explosion and greatly increased world hunger in many regions of the world.
From a practical perspective, a large number of simply inherited dominant or recessive genes conferring different types of resistances are still available in wheat germplasm and wild relatives of wheat. A solid knowledge of the mechanisms of polyploidization will help scientists in manipulating gene pools to improve cultivated wheat. Scientists and historians have long been searching for an explanation of the evolution and domestication of the various forms of cultivated wheat (diploid, tetraploid, and hexaploid; *T. monococcum*, *T. turdium*, and *T. aestivum*, respectively). The origin of polyploid wheat is complex because its evolution, since the various grass species diverged, has involved a long-established massive intervention of human and environmental selection pressures. The evolution within the entire Triticeae tribe included early widespread intra- and intergenome hybridization followed by introgression, gene flow, gene fixation, and rapid diversification within and among the ancestral diploid and polyploid species (Kellogg et al., 1996). Unequal rates of evolution, parallel evolution, DNA sequence deletion and/or amplification, and silencing during the evolution of present-day wheat species has been postulated to explain the complexity in phylogenetic relationships (McIntyre 1988; Appels et al., 1989; Feldman 2001).

Evolutionary studies involving various plant taxa have demonstrated that not only wheat, but also many polyploids, evolved from different progenitor populations. Independently formed polyploids most likely came in contact to hybridize with each other, thus resulting in ever-expanding primary and secondary germplasm pools (reviewed by Soltis and Soltis 1999). The formation of many allopolyploids was also accompanied by considerable genomic changes and structural reorganization within all or some of the parental genomes, including rapid nonrandom coded and noncoded sequence elimination, genic silencing, intergenic colonization by repeats and transposable elements, intergenic homogenization of divergent DNA sequences, DNA methylation changes, and other genomic modifications (Ozkan et al., 2001; Liu and Wendel 2002; Ma and Gustafson 2005). These genomic changes have been well demonstrated in the polyploids of the Triticeae tribe (Feldman et al., 1997; Kashkush et al., 2002; Han et al., 2003; Ma et al., 2004; Ma and Gustafson 2005, 2006). Such genomic changes, coupled with the likely repeated occurrence of polyploid formation, also contribute to the conflicting determinations of phylogenetic relationships and origins of many species, including wheat.

The origin, evolution, and domestication of cereals were among the major events shaping the development and expansion of human culture and will continue to shape the world in which we live. The domestication of cereals, which occurred approximately 10,000 years ago, was critical in laying the groundwork for the Neolithic revolution that transformed humanity to more centralized, sedentary farming societies (for a complete discussion see Kimber and Feldman 1987; also see especially Feldman 2001). There is no question that the various grass species (approximately 10,000 species), growing in every habitat in the world, and our understanding of the evolution of grasses are critical to developing the potential for grasses to feed the world’s ever-increasing population.

Polyploidy has been defined as the presence of more than one genome per cell and is probably the most common mode of speciation in plants (Stebbins 1950; Wendel 2000). Polyploids are classified into autopolyploids, which are formed from intraspecific chromosome doubling, and allopolyploids, which are the result of the interspecific or intergeneric hybridization of two or more genomes from differentiated species (Stebbins 1947). Polyploidy is one of the most important evolutionary events leading to a massive increase of genetic diversity, thus allowing species to adapt to varying environments. The most important and best-characterized group of allopolyploids, from an agricultural point of view, is the wheat genus (Kimber and Sears 1987; Feldman 2001). The evolutionary development of the various cultivated wheat species comprises several converging and diverging polyploid events involving several *Triticum* and *Aegilops* species from the Triticeae tribe.

It has been estimated that the Triticeae tribe began diverging from its progenitor approxi-
Section I Making of a Wheat Plant

approximately 35 million years ago (MYA) and that the *Triticum* group separated out about 11 MYA. The formation of the various polyploid wheat species within the *Triticum* genus began approximately 10,000 years ago. Since the early 1900s it has been known that the wheat species and the entire Triticeae tribe have a basic chromosome number of $n = 1x = 7$. Cultivated wheat consists of diploid (einkorn; $2n = 2x = 14$, AA), tetraploid (emmer, durum, rivet, Polish, and Persian; $2n = 4x = 28$, BBAA), and hexaploid (spelt, bread, club, and Indian shot; $2n = 6x = 42$, BBAADD) species. The various diploid genomes of the Triticeae tribe appear to be highly conserved in gene order along the seven pairs of chromosomes (Gale and Devos 1998). The chromosomes (1 through 7) in the various diploid genomes (B, A, and D) are considered to be evolutionarily related, that is, homoeologous in nature. When combined in the same nucleus, homoeologues can be induced to pair with each other. The importance of this fact in controlling our ability to make interspecific crosses and manipulate genes from one species to another will be considered elsewhere in this and other chapters.

Origin of the A genome

It is apparent that the key to understanding the evolution of wheat involves an elucidation of the evolution of the tetraploid wheat species. Early cytogenetic studies led to the conclusion that the A genome of the tetraploid species, *T. timopheevi* and *T. turgidum*, was contributed by *T. monococcum* (Sax 1922; Kihara 1924; Lilienfeld and Kihara 1934). However, it became apparent that diploid einkorn wheat actually comprised two biological species, *T. monococcum* and *T. urartu* Tum. Ex Gand. Chapman et al. (1976) determined that the A genome originated from *T. urartu*. Konarev et al. (1979) concluded, from studies of the immunological properties of seed-storage proteins, that the A genome in *T. turgidum* was contributed by *T. urartu*, and the A genome of *T. timopheevi* (Zhuk.) Zhuk. (GGAA) was contributed by *T. monococcum*. However, Nishikawa et al. (1994) suggested that the A genomes in both diploid species were contributed by *T. urartu*.

Clearly, the diploid component of the *Triticum* genus is composed of two defined species, *T. urartu* and *T. monococcum*. *Triticum monococcum* is the only cultivated diploid wheat species and was first found in Greece by Boissier (1884). Both *T. urartu* and *T. monococcum* have been identified in natural habitats ranging from southwestern Iran, northern Iraq, Transcaucasia, eastern Lebanon, southeastern Turkey, western Syria, and beyond into neighboring Mediterranean areas (Kimber and Feldman 1987). The sterility of their hybrids (Johnson and Dhaliwal 1976, 1978) confirms that they are valid biological species. It has been established that *T. urartu* contains approximately 4.93 pg DNA (http://data.kew.org/cvalues/introduction.html) and is the donor of the A genome to all polyploid wheat species. Dvořák et al. (1988) showed that variation in A-genome repeated nucleotide sequences, present in both tetraploid wheat species, was more related to the A genome of *T. urartu* than to the A genome of *T. monococcum*.

Origin of the B genome

Both the B and G genomes of tetraploid wheat have undergone massive changes following ancestor divergence and polyploidization, and they are widely considered to be modified S genomes having evolved from a common ancestor. Gu et al. (2004) indicated that the B genome diverged before the separation of the A and D genomes. There has been considerable controversy over the donor of the S-genome progenitor, but it was correctly identified as an ancestor of *Ae. speltoides* Tausch ($2n = 2x = 14$) in 1956 (Sarkar and Stebbins 1956; Riley et al., 1958; Shands and Kimber 1973; Dvořák and Zhang 1990; Daud and Gustafson 1996) and contains 5.15 pg DNA (http://data.kew.org/cvalues/introduction.html). Cytoplasmic analyses have shown that *Ae. speltoides* was the maternal donor of not only tetraploid but also hexaploid wheat (Wang et al., 1997). It is clear that the B genome has undergone significant intergenomic noncoded and coded DNA changes (in both the diploid and the tetraploid wheats) since the formation of tetraploid wheat. The B-genome component of polyploid wheat is the
largest of the wheat genomes and, because of the large degree of change at the DNA level, the true donor of the B genome since polyploid formation has been very difficult to establish. It would be useful to test representatives of A- and B-genome donors from across their geographic ranges to settle some of the past and present controversies over their origins. See the discussions based on genomic in situ hybridization (GISH) molecular cytogenetics by Belyayev et al. (2000) and Yen and Baenziger (1996). Regardless, since the cytoplasm donor is the female in the original cross creating the polyploid and is always listed first in any pedigree, the tetraploid genome designations should technically be BBAA or GGAA.

Emmer and durum wheat

Origin of Triticum urartu

Triticum urartu exists only in its wild form, contains 4.93 pg DNA (http://data.kew.org/cvalues/introduction.html), and supplied the male parent of tetraploid wheat (Feldman and Sears 1981), including several cultivated species. The most important are T. turgidum (BBAA), containing 12.28 pg DNA (http://data.kew.org/cvalues/introduction.html), and the sometimes cultivated, non–free-threshing T. timopheevi, which contains 11.30 pg DNA (http://data.kew.org/cvalues/introduction.html) and includes wild subspecies T. timopheevi araraticum (Jakubz.) Mac Key and cultivated subspecies timopheevi = T. turgidum ssp. timopheevi (Zhuk.). Triticum turgidum is further divided into several species, including T. turgidum ssp. dicoccoides (Korn. Ex Asch. and Graebn.), which is well known as the progenitor of all modern cultivated polyploid wheat species—that includes T. turgidum ssp. durum (Desf.) Husn., which is widely cultivated and is commonly called durum or macaroni wheat.

Outside the Fertile Crescent area, where T. dicoccoides wheat (Color Plate 1) reached the range of Ae. tauschii, the two species hybridized (Van Zeist 1976; Van Zeist and Bakker–Heeres 1985) and formed the hexaploid wheat group. This key hybridization event most likely occurred in the Caspian Sea region approximately 10,000 years ago. Notably, T. dicoccoides wheat is more adapted to Mediterranean environments, whereas noncultivated hexaploid wheat grows in cooler and more continental parts of Europe and western Asia (Fig. 1.1). Thus, T. dicoccoides deserves a particular in-depth study as suggested by Aaronsohn (1910, 1913), since it is the main genetic resource for improving both tetraploid and hexaploid wheat. Here we review theoretical and applied studies on T. dicoccoides that are important for all polyploid wheat improvement, including genetic structure across its range and its genetic resources. These topics are critically important to overcome the dangerous process of homogeneity occurring in all cultivated wheat gene pools. In addition, we will discuss the genome organization and evolution of T. dicoccoides.

Origin of Triticum dicoccoides (wild emmer)

Triticum dicoccoides is an annual, is predominantly self-pollinated, and has large and brittle ears with large elongated grains (Nevo et al., 2002), similar to cultivated emmer and durum wheats. It is the only wild ancestor in the genus Triticum that is cross-compatible and fully interfertile with cultivated emmer wheat (Zohary 1969; Chapman et al., 1976; Miller 1987, 1992; Harlan 1992; Zohary and Hopf 1993; Feldman et al., 1995). Triticum dicoccoides (Fig. 1.1) is a tetraploid containing the A (male) and B (female) genomes and is the female progenitor of all hexaploid wheat species. Triticum aestivum is the most important of the hexaploid wheats, followed by several primitive hulled types (spelta wheat) and numerous modern free-threshing forms (Zohary and Hopf 1993).
wild subspecies of the *T. turgidum* complex. Because of its central place in the evolution of cultivated wheat, wild emmer is among the best sources for obtaining insights into wheat evolution and improvement (Xie and Nevo 2008).

*Triticum dicoccoides* is a valid biological species (Miller 1992) that has a unique ecological niche in nature, where the seed dispersal mechanism involves “wild type” rachis disarticulation (brittle rachis), and spikelet morphology reflects adaptive-specialized traits that ensure survival in nature (Zohary 1969). Under the human selection system of reaping, threshing, and sowing, the selection and maintenance of the nonbrittle phenotype was highly advantageous and resulted in accelerated domestication (Miller 1992). Wild and domesticated forms also differ in kernel morphology (Van Zeist 1976); in cultivated tetraploid species the grain is wider, thicker, and rounder than in *T. dicoccoides*. Unique chromosomal translocations (Kawahara et al., 1993; Nishikawa et al., 1994; Joppa et al., 1995; Kawahara and Nevo 1996) and genetic polymorphisms (Nevo et al., 1982; Nevo and Beiles 1989; Fahima et al., 1998, 1999; Nevo 1998, 2001) also characterize *T. dicoccoides*. This combined evidence justifies its traditional classification as a separate species, as implied in the name *T. turgidum* ssp. *dicoccoides*, and it is clearly the progenitor of cultivated tetraploid and hexaploid wheat.

*Triticum dicoccoides* is found in Israel and Syria (which are its centers of distribution based on genetic diversity), Jordan, Lebanon, southeast Turkey, northern Iraq, and western Iran (Nevo and Beiles 1989; Nevo 1998). It was discovered in 1906, in eastern Galilee on the slopes of Mt. Hermon by Aaronsohn, who recognized its potential importance for all wheat improvement.
Chapter 1  Wheat evolution, domestication, and improvement

(Aaronsohn and Schweinfurth 1906; Aaronsohn 1910, 1913; Schiemann 1956; Feldman 1977; Nevo 1983, 1989, 1994, 2001). The genetic resource value of *T. dicoccoides* for wheat improvement far exceeded Aaronsohn’s vision (Nevo 2001; Peng et al., 2000a,b,c). In the northeastern distribution area of *T. dicoccoides*, where the sympatric area of *T. araraticum* is located, the two species are separated by strong sterility barriers (Maan 1973). Even though they are similar morphologically, they are practically indistinguishable without cytogenetic analysis.

*Triticum dicoccoides*, like *T. boeoticum*, was collected for human consumption long before its domestication (Kislev et al., 1992; Zohary and Hopf 1993; Lev–Yadun et al., 2000; Nesbitt 2001). Brittle *T. dicoccoides*–like plants with relatively narrow grains appeared in early Neolithic and Natufian Near Eastern settlements. However, 9–10 millennia ago, they also coexisted with nonbrittle seeds in Turkey (Jarmo, Iraq, Cayonu) (Hillman and Colledge 1998), in northern Syria (Tel Aswad and Tel Abu Hureira) (Zohary and Hopf 1993; Nesbitt 1998; Nesbitt and Samuel 1998; Lev–Yadun et al., 2000), and in Syria (Tell Mureybet I and II; 9000–8000 BC). *Triticum dicoccoides* was also discovered in Neolithic sites in Syria (Jerf el–Ahmar, Mureybet III, and Djade) and Turkey (Cayonu) (8000–7500 BC) and in sites near pre–Neolithic Turkey (Hallan Cemi Tepesi) and Iraq (Neolithic Qermez Dere and M'lefaat) (Nesbitt 1998; Lev–Yadun et al., 2000). Domesticated forms appeared in core–area Neolithic sites in Syria (Tell Abu Hureira 2A) and Turkey (Cafer Huyuk) about 7500 BC, and soon thereafter in Turkey (Cayonu and Nevali Cori) (Kislev et al., 1992; Nesbitt and Samuel 1998). From the early beginnings of agriculture in the Near East, 10,000 years ago and throughout the Chalcolithic and Bronze times, emmer was the principal wheat of newly established farming settlements; approximately 7000 years ago it spread from there to Egypt, the Indian Subcontinent, and Europe.

Patterns of allozyme diversity in wild *T. dicoccoides* suggest the following: (i) during the evolutionary history of wild *T. dicoccoides*, diversifying and balancing natural selections, through climatic, edaphic, and biotic factors, were major agents of creating genetic structure and maintaining differentiation; (ii) wild *T. dicoccoides* harbors large amounts of genetic diversity that can be utilized to improve both tetraploid and hexaploid wheat.

Wild *T. dicoccoides* grows extensively in the catchment areas of the Upper Jordan Valley (in northern Israel, in the eastern Upper Galilee Mountains, and the Golan Heights). Elsewhere in the Fertile Crescent (Fig. 1.1), populations of wild *T. dicoccoides* are semi-isolated and isolated and display a patchy structure. The highly subdivided, archipelago-type ecological population structure of wild *T. dicoccoides* is matched by its genetic population structure. Substantially more gene differentiation has been found within and between populations that were sometimes geographically very close within Israel, than between wild *T. dicoccoides* populations in Israel and Turkey (Nevo and Beiles 1989), where 40% of the *T. dicoccoides* genetic diversity existed within populations and 60% existed between populations. Only 5% of the genetic diversity was found between the Israel and Turkey metapopulations. This conclusion was reinforced based on edaphic, topographic, and temporal differentiation, on local microclimatic differentiation, on the extreme case of local isozyme differentiation in the Golan Heights (Nevo et al., 1982; Golenberg and Nevo 1987; Nevo et al., 1988a,b), and on recent DNA analyses (Fahima et al., 1999; Li et al., 1999, 2000a,b,c,d). The DNA results suggested that at least part of the noncoding regions were also subjected to natural selection. Genetic diversity was eroding across coding and noncoding regions of the *T. dicoccoides* genomes during and following domestication (Fahima et al., 2001). The *T. dicoccoides* genomes have been molded, in part, by diversifying natural selection from various ecological stresses.

The genetic differentiation within and between populations of *T. dicoccoides* was also reflected by an analysis of allele distribution (Nevo and Beiles 1989), which showed that 70% of all variant alleles were not widespread but revealed a definite localized somewhat sporadic distribution. Likewise, the analysis of genetic distances between populations supported the conclusion that sharp
local differentiation over short geographic distances was the rule, and the frequency of some common alleles (>10%) was localized and high. The population genetic structure of wild *T. dicoccoides* is obviously a mosaic and reflects the underlying ecological heterogeneity, which has been derived from local and regional geological, edaphic, climatic, and biotic differentiations. The genetic landscape is definitely not random between loci, populations, and habitats, and it most likely displays adapted patterns predictable on the basis of environmental factors. Could these polymorphisms represent adaptation to fluctuating environments? It has been possible to decide if selection is responsible for the occurrence of many DNA variants across the coding and non-coding regions of a genome, and it is clear that major DNA changes can and do occur within and between *T. dicoccoides* populations over a relatively short time frame, paralleling that of allozymes.

Nevo and Beiles (1989) predicted that neither migration nor genetic drift could have generated the patterns observed between loci and alleles of wild *T. dicoccoides* and that selection remained a vital explanatory model. Environmental selection also partly affected loci differentially, but differently from migration. This was supported by data from Nevo and Beiles (1989) for three reasons: (i) variation was found among loci; (ii) in an autocorrelation analysis, positive correlations were found in different distant *T. dicoccoides* groups, and not necessarily in the first one as would be expected if migration determined the interpopulation genetic structure; and (iii) the predominance of negative correlations in the larger distant groups was found to be due to decreasing ecological similarity often observed with increasing distance.

The maintenance of polymorphisms in wild *T. dicoccoides* may be explicable by both spatial and temporal variation in selection. Theory indicates that selection, acting differentially in space, coupled with limited migration, which is typical of wild *T. dicoccoides*, will maintain a substantial amount of polymorphism (Karlin and McGregor 1972; Hedrick 1986; Nevo et al., 2000). Thus, different polymorphisms will be favored in different climatic and edaphic niches, from regional to local, and at miniscule levels within a locality. Microniche ecological selection (e.g., climatic factors related to temperature, available water, and biotic and abiotic stresses) could be a major cause of genetic differentiation rather than stochastic processes.

### Origin of hexaploid wheat

There are two main forms of hexaploid wheat, including *T. zhukovskyi* Men. & Er., which was the result of a recent hybridization involving *T. timopheevi* and *T. monococcum*, the only example of hexaploid wheat to have the GGAAAmAm constitution (Upadhya and Swaminathan 1963). The most important hexaploid wheat group comprises *T. aestivum* (BBAADD) and its several subspecies containing 21 pairs of chromosomes with seven pairs belonging to each of the A, B, and D genomes (Sears 1954; Okamoto 1962) and containing 17.33 pg DNA (http://data.kew.org/cvalues/introduction.html).

*Triticum aestivum* originated approximately 10,000 years ago after the domestication of tetraploid wheat and was derived from the hybridization of a primitive tetraploid (BBAA), as the female, and *T. tauschii* ssp. *strangulata* [* Ae. tauschii* (Coss.) Schmal, also known as *Ae. squarrosa*, DD, 2n = 2x = 14, 5.10 pg DNA], as the male (Kihara 1944; McFadden and Sears 1944, 1946a,b; Kimber and Sears 1987; Kimber and Feldman 1987; Dvořák et al., 1998). The first primitive hexaploid wheat was probably a hulled-type like *T. aestivum* var. *spelta*, *macha*, or *vavilovii*. The current free-threshing types, *T. aestivum* var. *aestivum*, *sphaerococcum*, or *compactum*, were the result of a mutation at the *Q* gene locus (Mura-matsu 1986) followed by selection. All polyploid wheat species are disomic in inheritance due to complete diploidlike chromosome pairing, which is controlled by two main homoeologous pairing genes *Ph1* (Riley and Chapman 1967) and *Ph2* (Mello-Sampayo 1971) and several minor genes (for a complete review, see Sears 1977). As previously stated, since the cytoplasm donor of hexaploid wheat was the female in the original cross creating the polyploid, it should be listed first in any pedigree or genome designation; therefore,
hexaploid genome designations should be stated as BBAADD as noted by Feldman (2001).

Research has shown that hexaploid wheat is less variable than its diploid progenitors, suggesting the possibility of a genetic bottleneck caused by a very limited number of initial hybridizations (Appels and Lagudah 1990). However, given the obviously large distribution area of primitive tetraploid wheat and *Ae. tauschii* populations within the cradle of agriculture (for an excellent review of cultivation regions, see Feldman 2001), the natural occurrence of multiple tetraploid wheat and *Ae. tauschii* hybrids could be a more common occurrence than originally thought. The presence of several sets of alleles and microsatellites established that hexaploid wheat resulted from several hybridizations (Dvorák et al., 1998; Talbert et al., 1998; Lelley et al., 2000; Caldwell et al., 2004; Zhang et al., 2008). Zohary and Hopf (1993) suggested that these hybridizations are still occurring today.

Clearly, under certain environmental conditions, some degree of outcrossing in wheat does occur (Grifﬁn 1987; Martin 1990) and hybrids of various ploidy levels can be formed (McFadden and Sears 1947; Ohtsuka 1998), both of which would indicate less of an evolutionary bottleneck in the development of hexaploid wheat than previously suggested. In addition, hexaploid wheat originated and still originates in a region where all of the progenitors reside, thus allowing for a continuous intercrossing and backcrossing with diploid progenitors. Even with the presence of ploidy and genome differences, the various types of primitive wheat species are capable of widespread intercrossing, culminating in intraspeciﬁc hybrid swarms which would signiﬁcantly increase the potential for gene ﬂow over time. This is possible because all polyploid wheat progenitors share at least one common genome, which can serve as a buffer or a pivot around which unpaired homoeologous chromosomes can pair. Any homoeologous chromosome pairing, within the genomes of the Triticeae tribe, can and does allow for the occurrence of additional gene recombination and exchange. Since tetraploid and hexaploid wheat are predominantly self-pollinated, homozygosity for any gene exchanges favored by natural selection would be rapidly achieved and available for artiﬁcial selection.

This presence of modiﬁed genomes along with unmodiﬁed or pivotal genomes, widespread throughout the Triticeae tribe, was originally suggested by Zohary and Feldman (1962) and later in wheat–rye hybrids by Gustafson (1976), and it has been shown to occur more often than expected. The presence of the pivotal (buffering) genomes in primitive polyploid wheat crosses made possible the rapid and very successful expansion of wheat in a very short time. This manner of polyploid speciation allowed for a greater degree of gene ﬂow and genome modiﬁcation than that which has been observed in any diploid system of speciation. This process of wheat polyploid speciation needs to be kept clearly in mind when attempting to make wide crosses (interspeciﬁc and intergeneric) to introduce genes from other species and genera into wheat. The presence of pivotal genomes in polyploid wheat complexes makes it easier for breeders to understand the processes involved in manipulating gene complexes from related grass species into wheat. It also makes very clear the importance of maintaining and expanding all existing diploid and polyploid germplasm collections of wheat and wheat relatives as vast primary, secondary, and tertiary gene pools for future use in wheat improvement.

The polyploid wheat species represent a converging form of evolution where several genomes have been combined. This form of species hybridization coupled with inbreeding has resulted in a very successful polyploid that is highly adaptable to a wide range of environmental growing conditions. The evolution of polyploid wheat and its intimate connection with the transition of human societies from hunting–and–gathering to an agricultural culture occurred over a long time and involved vast mixtures of wild and increasingly domesticated populations, and of hulled and free-threshing forms, ultimately resulting in a diverse and dynamic wheat gene pool.

The genetic composition of polyploid wheat species fully accounts for their successful establishment. The evolutionary development of a genetic system conferring diploidization (*Ph*
mutant; for a complete review, see Sears 1977), thus preventing multivalent chromosome formation with deleterious intergenomic exchange, was critical for the stabilization of all polyploid wheat species. Mutations, within the various wheat genomes, also played a major role in allowing wheat to increase in variability, stabilize as a species, and become the major food crop. The two and three genomes present within tetraploid and hexaploid wheat, respectively, and the self-pollinating character of all species, resulted in the accumulation of mutations that became available for selection. This allows for individuals within populations to become a main driving force upon which natural selection operates. This form of gene formation, modification, and stabilization is one of the most powerful processes in plant evolution.

GENOME EVOLUTION AND MODIFICATION

We now have a voluminous amount of information concerning the ancestors and evolutionary processes that created polyploid wheat. To fully understand the genomic evolution of polyploid wheat, it is important to ask why each of the diploid genomes comprising polyploid wheat is so massive relative to other grass species such as rice. The B genome is 5.15 pg DNA (Furuta et al., 1986); the A genome is 4.93 pg DNA (Bennett and Smith 1976); and the D genome is 5.10 pg DNA (Rees and Walters 1965). On the other hand, rice contains only 0.6–1.0 pg DNA in japonica and indica types, respectively (Bennet and Leitch 1997; http://data.kew.org/cvalues/introduction.html). However, the various wheat genomes and the rice genome appear to have similar genetic composition with a good macrolevel syntenic relationship (Gale and Devos 1998; Sorrells et al., 2003; Tang et al., 2008). Flavell et al. (1974) and Gu et al. (2004) established that over 80% of the hexaploid wheat genome comprised noncoded highly repeated DNA sequences and highly active and nonactive retrotransposons. The intergenic regions (Bennetzen 2000; Feuillet and Keller 2002; Wicker et al., 2003) and genic regions (Gu et al., 2004) of many species are mainly composed of retrotransposons, and the vast numbers of these retrotransposons are correlated with genome size (Kidwell et al., 2002). Most of the retroelements in the three wheat genomes are not colinear, which suggests that their present location was the result of genome divergence after the individual A, B, and D genome parental species were combined (Gu et al., 2004). When analyzing the glutenin genes of wheat, Gu et al. (2004) found that more genes from the glutenin region of the A genome contained retrotransposons than occurred in orthologous regions of either the B or D genome.

Is all or most of the noncoded DNA present in hexaploid wheat really “junk” DNA? From an evolutionary view, it is highly unlikely that any genome would expend a vast percentage of its energy production maintaining DNA that was of little or no value. The reason behind the presence and function of vast amounts of noncoded DNA in the wheat genomes remains largely unknown. To fully understand and be able to manipulate wheat genome evolution, the function and purpose of this noncoded DNA needs to be investigated.

There is an abundance of data supporting the ability of a genome to increase and/or decrease in DNA amount over time, compared with that observed in its original progenitor. Such genomic changes (deletions and additions, gene conversions, transposon activation and silencing, chromosomal rearrangements, epigenetic events, etc.) are known to occur widely in grass genomes (Feldman et al., 1997; Liu et al., 1998; Ozkan et al., 2001; Shaked et al., 2001; Kashkush et al., 2002; Han et al., 2003; Ma et al., 2004; Ma and Gustafson 2005, 2006) and other polyploid plant genomes, including, for example, *Brassica napus* polyploids (for an excellent article, see Gaeta et al., 2007). The frequency of such events is not uniform across individual chromosomes or within complete genomes. The selection pressures acting on DNA deletion or insertion in either a plant or animal genome can be different, depending on whether or not changes are located in repeated DNA, heterochromatin regions, or gene-rich regions. Diaz-Castillo and Golic (2007) noted in
Drosophila that gene structure and expression were influenced by the location of genes proximal to heterochromatin and were evolving at a rate in response to their chromosomal location.

No fully satisfactory explanation has been suggested for why these major evolutionary genome modification, deletion, and addition processes take place mainly in the noncoded portion of the genome. Wicker et al. (2003) and Gaut et al. (2007) have made a strong case for illegitimate recombination having a major influence on genome evolution. Illegitimate recombination is capable of generating deletions, inversions, gene conversions, and duplications within any chromosome of any genome. However, it is difficult to envision illegitimate recombination as the main cause for such a sizeable DNA deletion, of up to about 10% or more in the genomes of many allopolyploid cereals. It is likely that no single explanation will answer the question of why the cereal genomes vary so much in size. It will most certainly require a number of working hypotheses and a large body of new evidence and knowledge bearing on the problems associated with the evolution of genome size in grasses to resolve this question. See a recent review on synteny and colinearity in plant genomes by Tang et al. (2008).

We can propose one possible cause for many of the observed vast changes in grass genome composition. Clearly every grass genome goes through its cell cycle at a specific rate, which varies with each genome. Van’t Hof and Sparrow (1963) first proposed the existence of a relationship between DNA content, nuclear volume, and mitotic cell cycle, and suggested that any mitotic cell cycle is greatly influenced by the amount of DNA present in the genome. They made it clear that the amount of DNA present in a genome does affect cell cycle, and ultimately plant growth, regardless of whether or not it was coded. Recently, Francis et al. (2008) concluded that the speed of DNA replication was identified as the limiting factor in the cell cycle. Therefore, it follows that individual genome cell cycle differences cause problems of maintaining their synchrony when two or more genomes, with different volumes of DNA, are placed together in a cell.

For example, in the wheat–rye hybrid triticale (×Triticosecale Wittmack), Bennett and Kaltsikes (1973) showed that the meiotic duration of wheat and rye differed from that observed in the hybrid, and the hybrid had a meiotic cell cycle closer to the wheat parent. Their observations made it clear that if one genome of a hybrid has not completed its cell cycle by the time cell wall formation has initiated, the possibility of breakage-fusion-bridges occurring in the genome with the lagging cell cycle will be greatly increased, most likely resulting in DNA elimination or addition. This is what happens in a wheat–rye hybrid and can be readily seen in the formation of large aberrant nuclei that are readily visible in the early cenocytic stages of endosperm development before cellularization takes place (Fig. 1.2). The formation of cell walls at the first division of the embryo would definitely cause breakage-fusion-bridges to occur immediately and lead to the decrease—or even increase—of DNA present in the genome with the lagging cell cycle.

![Fig. 1.2](a) A wheat–rye hybrid (triticale) seed only 48 hours after pollination with a cenocytic endosperm and a cellular embryo (arrow); (b) a nuclear division (24 hours after pollination) showing bridges that have formed during anaphase; (c) nuclear divisions (48 hours after pollination) showing rye telomeres that have formed bridges during anaphase; and (d) nuclear divisions (72 hours after pollination) showing rye telomeres that have formed bridges during anaphase.
As has been observed in the early endosperm development of wheat–rye hybrids, deletions within the rye genome were clearly detected (Gustafson and Bennett 1982; Bennett and Gustafson 1982). Deletions and increases in DNA have even been detected within heterochromatic regions of the genus Secale (Gustafson et al., 1983). Variations in DNA content have been observed in polyploids of the Tribiceae tribe (Feldman et al., 1997; Ozkan et al., 2001; Kashkush et al., 2002; Liu and Wendel 2002; Han et al., 2003; Ma et al., 2004; Ma and Gustafson 2005, 2006), in rice (Wang et al., 2005), in maize (Messing et al., 2004), in a few Hordeum species (Jakob et al., 2004), and in synthetic polyploids of Brassica (Song et al., 1995). In some induced dihaploids of Nicotiana (Dhillon et al., 1983; Leitch et al., 2008) and Gossypium species (Grover et al., 2007), an increase in DNA has actually been detected. For an excellent review of polyploids showing a decrease in DNA over time, see Leitch and Bennett (2004). Any genome cell cycle differences could easily be the major cause of genome variation in DNA content between an allopolyploid and its parental species.

MECHANISMS FOR CHROMOSOME EVOLUTION

As stated previously, the evolution of the genus Triticum serves as a good model of polyploidy, one of the most common forms of plant evolution (Elder and Turner 1995; Soltis and Soltis 1999). From a practical perspective, large stores of simply inherited genes that confer different types of resistance are available in wheat and its wild relatives via germplasm collections. Knowledge of the mechanisms of polyploidization will help plant breeders to enrich the gene pool of cultivated wheat. The origin and co-evolution of A and B genomes of tetraploid wheat has long been controversial (Feldman and Sears 1981). Unknown are the details of the co-evolution of A and B genome repetitive sequence arrays in allotetraploid wheat. There is no reason to regard the process of allopolyploidization as a mechanical combination of sequences from two genomes, but still less is known about the interaction between sequences from different arrays in chromatin fractions (Wendel 2000). In this section we draw attention to several critical points of speciation-related chromosomal changes.

Chromosomal rearrangements and repetitive DNA

Major structural chromosome rearrangements including deletions, duplications, translocations, and inversions are often associated with cytogenetically detectable heterochromatic regions composed of repetitive DNA, and they frequently appear in heterochromatin–euchromatin borders (Badaeva et al., 2007). Chiasmata in meiosis appear very close to the terminal and intercalary C-bands and mark the point of exchange (Loidl 1979). Well-studied intraspecific C-banding polymorphisms can be regarded as a manifestation of this interdependence. The diploid–polyploid Aegilops–Triticum complex exemplifies abundant C-banding polymorphism based on chromosomal rearrangements (Badaeva et al., 1996, 1998, 2002, 2004, 2007; Friebe and Gill 1996; Rodriguez et al., 2000a,b; Maestra and Naranjo 1999, 2000). A good example of this is where the combination of C-banding techniques and fluorescence in situ hybridization (FISH) with ribosomal RNA genes, 5S and 18S–5.8S–26S rDNA (45S rDNA), and with a D-genome-specific repetitive DNA sequence pAs1 revealed species-specific patterns of heterochromatin, rDNA, and pAs1 clusters for six D-genome-containing allopolyploid Aegilops species: Ae. cylindrica, Ae. ventricosa, Ae. uniaristata, Ae. crassa, Ae. vavilovii, and Ae. juvenalis (Badaeva et al., 2002). A wide spectrum of chromosomal rearrangements, particularly species-specific, and genome-specific redistribution of repetitive DNA clusters led to hypothesizing the phylogenetic relationships in this group of polyploid Aegilops species.

Heterochromatin

An inherent feature of heterochromatin is the complex composition of tandem repeats of various
types (Sharma and Raina 2005) and transposable elements (TEs), predominantly retrotransposons (elements of Class I that transpose via RNA intermediates) (Lipman et al., 2004). Three groups of retroelements—Ty1-copia, Ty3-gypsy, and LINE—were found in large quantities in heterochromatin of the diploid B/G-genome progenitor Ae. speltoides (Belyayev et al., 2001). Clusters of tandem repeats and TEs form species-specific and chromosome-specific heterochromatin patterns. There is a certain correlation between distribution-clustering of retroelements and chromosome location of tribe-specific and species-specific sequences within cereal genomes (Color Plate 2). Thus, the independently discovered tribe-specific tandem repeat Spelt 52 (Anamthawat-Jonsson and Heslop-Harrison 1993; Friebe et al., 2000; Giorgi et al., 2003) and Ae. speltoides—specific tandem repeat Spelt 1 (Salina et al., 1997; Pestsova et al., 1998) cluster together with retroelements at the same chromosomal locations corresponding to AT-enriched heterochromatin. Moreover, this complex of distal-terminal chromosomal regions enriched by TEs of different types and tribe- and species-specific tandem repeats could be classified as a faster-evolving part of the genome (Belyayev and Raskina 1998) (Color Plate 2, note the green signal on the distal regions of chromosome 4 after GISH).

Repetitive DNA

The repetitive DNA fraction plays an important role during polyploidization and post-polyploidization changes (Dvořák and Zhang 1990, 1992; Dvořák and Dubcovsky 1996; Feldman and Levy 2005; Ma and Gustafson 2005, 2006). In the genomes of allopolyploid wheat, T. dicoccoides (B and A genomes) and T. aestivum (B, A, and D genomes), the distribution pattern of highly repetitive DNA clusters and Ty1-copia retroelements differs from those of their diploid progenitors, T. urartu and Ae. speltoides (Color Plate 2a–d) (Raskina et al., 2002). Significant intercalary repositioning and decay of a majority of distal-terminal clusters of AT-positive heterochromatin were observed in the B genome of allopolyploid wheat in contrast to the S genome of Ae. speltoides (Color Plate 2c). Similar differences have been observed between populations of Ae. speltoides involving a series of distal–terminal chromosomal rearrangements (Raskina et al., 2004a,b). Multiple translocations and deletions occurred, which led to the current heterochromatin pattern (Color Plate 2e) and a majority loss of Ae. speltoides–specific tandem repeat Spelt 1 clusters (A. Belyayev, pers. comm.). Since new allopolyploids continue to occur in the periphery distribution areas of existing species (Grant 1981), we will continue to observe allopolyploidization involving the Ae. speltoides genome containing numerous chromosomal rearrangements and mobile elements (Raskina et al., 2004a,b). The present-day B genome of wild and cultivated wheat carries from zero to two Spelt 1 clusters per haploid genome in contrast to the G genome of existing allopolyploids, which contains up to six Spelt 1 clusters (Salina et al., 2006). These data are in accordance with the purported independent origin of the B and G genomes of allopolyploid wheat (Jiang and Gill 1994b; Rodriguez et al., 2000a).

Ongoing permanent intragenomic mutagenesis in plant populations is a generator of heterozygosity leading to intraspecific genetic variability and creates the basis for natural selection under changing environments. Significant inter-B/A-genomic interactions, in allotetraploid wild emmer wheat, revealed major substitutions of part of the A-genome heterochromatin clusters by satellite DNA from the B genome (Color Plate 2c) (Belyayev et al., 2000). Enrichment of these clusters with mobile Ty1-copia retroelements suggests an important role of TEs in rebuilding and homogenizing the allopolyploid genome, leading to stabilization of T. dicoccoides as a new species. Retroelements are known to play a large role in gene and genome evolution (Flavell et al., 1997; Kidwell and Lisch 2001; Bennetzen 2002). In T. aestivum, substitution of part of the heterochromatin from the “youngest” D genome by repetitive DNA from the A and B genomes was revealed to a far lesser degree (Color Plate 2d, 14 red D-genome chromosomes marked by asterisk).

Transposable elements can directly change molecular composition and/or DNA amount in the regions of insertions. They also can mediate
ectopic chromosomal exchanges when homologous and/or nonhomologous chromosome recombination moves sequences within and between genomes. Furthermore, insertions of TEs may create new crossing-over “hot” spots that provoke transposable element-mediated homologous or nonhomologous chromosome rearrangements. The latter include spontaneous translocations, inversions, and deletions and are potential mechanisms for rapid genome reorganization during speciation and stabilization of any allopolyploid. For example, in the wheat 4AL–7BS translocation (Naranjo 1990), a cluster of Ty1-copia retrotransposons was detected (Color Plate 2b), and chromosomes 4A and 7B were also involved in a 4AL–7BL translocation, which was detected in a natural population (Raskina et al., 2002).

The manipulation of repetitive DNA complexes plays an important role in evolutionary genome transformation. Changes in repetitive DNA may cause chromosomal rearrangements and, in turn, chromosomal rearrangements may cause repetitive DNA change through mechanisms of concerted evolution (Elder and Turner 1995). Therefore, these processes are interdependent.

**Repatterning of rDNA arrays in the wheat genome**

In addition to the direct detection of major chromosomal rearrangements, it is also possible to indirectly estimate the level of microevolutionary genomic change by evaluating the repatterning of well-determined chromosomal markers and by the mobility of rDNA clusters. It is obvious that speciation-related chromosome structure change establishes further increases or decreases in the number of rDNA sites or their repositioning, but the dynamics of rDNA clusters may be regarded as an indicator for significant intragenomic processes (Jiang and Gill 1994a; Dubcovsky and Dvořák 1995; Raskina et al., 2004b).

The location, number, and mobility of rDNA clusters have been described in many plant species and may involve major loci, small numbers of copies of the repeat unit, or fragments of a repeat unit, which are often known not to be transcribed (for review, see Heslop-Harrison 2000). There is evidence that rDNA repeat sites may alter chromosomal location without the involvement of translocations or other chromosomal rearrangements (Dubcovsky and Dvořák 1995). Schubert and Wobus (1985) examined the mobility of nucleolar organizing regions (NORs) in *Allium* and proposed TE activity as one of the possible sources for rDNA movement. Recent studies proposed that transposons (*En/Spm*-like; elements of Class II that move by extinction and reintegration) might be involved in rDNA repatterning (Raskina et al., 2004a). The ability of some classes of transposons (Pack-MULES, Helitrons) to capture entire genes and move them to different parts of the genome has been documented (Jiang et al., 2004; Lai et al., 2005).

Therefore, any interaction of ribosomal genes and TEs relative to evolutionarily significant chromosomal repatterning appears to be of tremendous interest yet remains largely unexplored. Certain remodeling of chromosome-specific repetitive DNA patterns may lead to meiotic abnormalities. In extremes, these abnormalities are capable of causing reproductive (postzygotic) isolation (Grant 1981).

A series of *in situ* hybridization (ISH) experiments revealed permanent clustering of different TEs in the NOR (which contains 45S rDNA loci) as well as near or within clusters of 5S rDNA (Belyayev et al., 2001, 2005; Raskina et al., 2004a,b). Therefore, we can suggest that the possible association of TEs and rDNA loci arise due to, first, the insertion preference of the TEs in the rDNA arrays. Indeed, rDNA arrays are common targets for several LINE retrotransposons (Eickbush and Eickbush 2003; Averbeck and Eickbush 2005) and also for some Class II transposons (Penton and Crease 2004). Second, these two components may accumulate preferentially within the same genomic context, perhaps driven over time by selection against insertions elsewhere in the genome (e.g., heterochromatin in the case of retroelements). Third, a possible functional relationship exists between the dispersion of TEs and rDNA genes. Additional molecular–bioinformatic studies may further explain TE–rDNA gene interactions.
Repetitive DNA and mobile elements as perpetual generators of diversity and evolution

Speciation in wild diploid and polyploid wheat is tightly connected with significant repatterning of rDNA sites (Dubcovsky and Dvořák 1995). Considering rDNA in terms of temporary genome changes, we face a certain paradox. On one hand, rDNA is the most conservative fraction in the eukaryotic genome, and ribosomal RNA genes undergo minimal changes over hundreds of millions of years. On the other hand, this conservatism appears to be a source of genome instability. Due to the similarity of rDNAs, any chromosome that carries extended rDNA arrays has the potential for involvement in heterologous synapses and recombination (Raskina et al., 2004b). Thus, any rDNA cluster could consist of several layers of different origins, especially in a polyploid species, with a high level of interhaplome invasion (Belyayev et al., 2000). This is clearly seen in wheat (Color Plate 2f), where differently labeled nontranscribed spacers (NTS) of 5S rRNA genes of different origins (short A1 and long G1) show slightly different positions inside the rDNA cluster on chromosome 1A of *T. dicoccoides* (Baum et al., 2004).

Due to the known capability of mobile elements to provoke ectopic exchanges, the consequences of TE–rDNA interaction make it possible to propose that the collocation of different TEs within recombinogenic hot spots could intensify homologous and heterologous recombination. Moreover, TE-mediated intragenomic transfer of rDNA fragments and the inheritance of such mutations may cause significant evolutionary changes in chromosomal distribution of rDNA clusters (Raskina et al., 2004a,b).

Another possible consequence of the physical association of rDNA–TE within the 45S rDNA region (NOR) could be the loss of chromosome satellites with all their genetic content. McClintock (1946) suggested that TEs could cause chromosomal breakages. High concentrations of TEs around 45S rDNA increase the fragility of this site (Color Plate 2c, chromosomes 1 and 6). In the case of a satellite loss, the remainder of the 45S rDNA block will be in the telomeric position as has been detected in many wheat and *Aegilops* species (Mukai et al., 1990; Dubcovsky and Dvořák 1995; Badaeva et al., 1996; Baum et al., 2004). When chromosomes are broken, the breakpoints become highly unstable and acquire the ability to fuse with other broken ends (McClintock 1941). However, the breakpoints are eventually stabilized, and the reconstructed chromosomes are transmitted to the daughter cells. This phenomenon, known as healing of breakpoints, involves the addition of repetitive telomere sequences at the breakpoints by telomerase, the enzyme that normally synthesizes the telomere sequence at normal chromosome terminals (Tsujimoto et al., 1997). According to Tsujimoto et al. (1999), rDNA sequences provide insight into the properties of telomerase activity at the breakpoints. The telomere sequences initiate two- to four-nucleotide motifs in the original rDNA sequence. These motifs are also found in the repeat unit of telomere sequences. Thus, it has been documented in many plant species that rDNA in terminal positions could stimulate de novo rapid synthesis of telomeres.

Therefore, we can emphasize that single Class I or II TEs constantly form distinct clusters in or around regular and irregular rDNA sites, and that the presence of TEs in or around rDNA sites increases the possibility of recombination and satellite loss. Apparently this event is common in plant karyotype evolution, since in many plant species rDNA clusters in terminal positions have been detected.

THE POTENTIAL OF WILD EMMER IN WHEAT IMPROVEMENT

Studies on wild emmer (*T. dicoccoides*), the progenitor of most tetraploid and hexaploid wheats, have revealed rich genetic resources applicable to wheat improvement, given its diverse single- and multilocus adaptations to stressful abiotic and biotic environments (Xie and Nevo 2008). The available resources have been described (Zohary 1970; Feldman 1979; Lange and Jochemsen 1992; Grama et al., 1983; Nevo 1995, 2001; Nevo et al.,...
22

Section I Making of a Wheat Plant

2002) and they include genotypic variation for traits such as: (i) germination, biomass, earliness, nitrogen content, and yield; (ii) amino acid composition; (iii) grain protein content and storage protein genes (HMW glutenins); (iv) disease resistances, including resistance to powdery mildew \([Blumeria graminis\) (DC) E.O. Speer f. sp. \(tritici\] \(Blumeria graminis\) (DC) E.O. Speer f. sp. \(tritici\] f. sp. \(tritici\] Eriks.] \(Puccinia triticina\) Eriks.), and \(Soilborne wheat mosaic virus\ (WSBMV); (v) high photosynthetic yield; (vi) salt tolerance; (vii) herbicide resistance; (viii) amylases and \(\alpha\)-amylase inhibitors; and (ix) micronutrients such as \(Zn\) and \(Fe\). This is only a preliminary list of the vast potential genetic resources existing in wild emmer that remain to be exploited for wheat improvement.

Quantitative trait loci (QTLs) and beneficial cryptic, agronomically important alleles have now been extensively described. The current genetic map of \(T. dicoccoides\), with 549 molecular markers and 48 significant QTLs for 11 traits of agronomic importance (Peng et al., 2000b), will permit the unraveling of beneficial alleles of candidate genes that are otherwise hidden. These beneficial alleles could be introduced into cultivated wheat by marker-assisted selection.

The Near East, in general, and Israel, in particular (Nevo 1986), are the centers of origin and diversity of wild emmer, where it developed wide genetic adaptations against multiple pathogens and diverse ecological stresses. Genetic diversity is transferable from the wild to the cultivated gene pool, so genes of wild emmer are directly accessible for future wheat improvement. Consequently, exploration of \(in situ\) and \(ex situ\) collections (with optimized sampling strategies) along with utilization programs should maximize the contribution of wild emmer to wheat improvement. Among the potential donors for improving wheat, wild emmer occupies a very important and unique position due to its direct ancestry of bread wheat and its rich and largely adaptive genetic diversity. This was first suggested by Aaronsohn (1913) and later elaborated on by many authors (see Feldman 1977; Nevo 1983, 1995, 2001, 2006; Xie and Nevo 2008).

There are many ongoing programs around the world utilizing genes of wild emmer for wheat improvement, primarily involving genes coding for resistance to powdery mildew and the rusts, for high protein content, and for improved baking quality. Cultivars based on introgression of \(T. dicoccoides\) genes have appeared and will continue to appear in the near future. With \(T. dicoccoides\) at least three to four backcrosses with bread wheat are a necessity in breeding programs to minimize linkage drag (Groenewegen and van Silfhout 1988; Reader and Miller 1991). Wheat improvement programs will continue utilizing \(T. dicoccoides\) and other wheat relatives (Xie and Nevo 2008). Extensive work on transferring genes for high protein content from \(T. dicoccoides\) to cultivated wheat is currently underway in several laboratories (e.g., Weizmann Institute; US Department of Agriculture, Fargo, North Dakota; and University of California, Davis).

CONCLUDING REMARKS ON THE PROCESS OF WHEAT EVOLUTION

The molecular diversity and divergence of wheat species displays parallel ecological–genetic patterning and demonstrates the following: (i) significant genetic diversity and divergence exists at single- and multilocus structures of allozymes, random amplified polymorphic DNA, simple-sequence repeats, and single-nucleotide polymorphisms over very short distances of several to a few dozen meters; (ii) genetic patterns across coding and largely noncoding genomic regions are correlated with, and predictable by, environmental stress (climatic, edaphic, biotic) and heterogeneity (the niche–width variation hypothesis), displaying significant niche-specific and niche–unique alleles and genotypes; and (iii) genomic organization of wheat, including the noncoding genome, is nonrandom, heavily structured, and at least partly, if not largely, adaptive. The process of wheat evolution defies explanation by genetic drift, neutrality, or near neutrality alone as the primary driving forces. The main viable model to explain wheat genomic organization seems to be natural selection, primarily diversifying and bal-
ancing, and cyclical selection over space and time according to the two- or multiple-niche ecological models. Natural selection interacts with mutation, migration, and stochastic factors, but it overrides them in orienting the evolutionary processes of wheat.

FUTURE PERSPECTIVES

What will be the next step in wheat improvement in the current genomic and postgenomic eras? Conceptually, in-depth probing of comparative genome structure and function are the major challenges, including analyses of the intimate relationship between the coding and noncoding regions of the wheat genomes. Such studies will unravel genome evolution and highlight the rich genetic potentials for wheat improvement residing in wheat and various wheat relatives, including *Triticum* and *Aegilops* species as well as other Triticeae.

We believe that the theoretical and applied perspectives for future wheat improvement will encompass the following. First is characterization of the genome structure, function, regulation, and evolution at macro- and microgeographic scales of wheat and wheat-related species. Second is to combine multilocus markers and fitness-related traits to produce direct estimates of adaptive fitness differentiations within and between populations. Third, a critical activity will be to analyze the genetic system determining the enormous genetic flexibility of the various wheat and wheat-related species in diverse ecological contexts, mutation rates in different elements of the genome, recombination properties of the genome with their genetic and ecological control, and genomic distribution and function of structural genes (primarily abiotic and biotic stress genes). Fourth, it would be prudent to characterize the interface between ecological and genomic spatio-temporal dynamics and adaptive systems, to characterize genome evolution and the polyploidization processes, and to conduct colinearity studies between the grasses, including model species with small genome size such as rice and *Brachypodium distachyon*.

REFERENCES


Bennetzen, J.L. 2002. Mechanisms and rates of genome
Bennetzen, J.L. 2000. Comparative sequence analysis of plant
Bennett, M.D., and J.B. Smith. 1976. Nuclear DNA amounts
Bennett, M.D., and P.J. Kaltsikes. 1973. The duration of
Brown, A.H.D., G.J. Lawrence, M. Jenkin, J. Douglass, and
Brown, A.H.D., M.T. Clegg, A.L. Kahler, and B.S. Weir
Bennett, M.D., and J.P. Gustafson. 1982. The effect of telo-
distribution of reverse transcriptase-containing retroele-
ments in two Triticeae species. Chromosome Res. 9:129–
Ty3-gypsy retrotransposons chromosomal distribution in
topologies of two wild Triticeae species. Cytogenet.
Bennett, M.D., and J.P. Gustafson. 1982. The effect of telo-
meric heterochromatin from Secale cereale L. on triticale
(X Triticosecale Wittmack): II. The presence or absence of
blocks of heterochromatin in isogenic backgrounds. Can.
Bennett, M.D., and P.J. Kaltikes. 1973. The duration of
meiosis in a diploid rye, a tetraploid wheat and the hexa-
196.
Bennett, M.D., and J.B. Smith. 1976. Nuclear DNA amounts
274:227–274.
Bennetzen, J.L. 2000. Comparative sequence analysis of plant
nuclear genomes: Microcolinearity and its many exceptions.
Bennetzen, J.L. 2002. Mechanisms and rates of genome
expansion and contraction in flowering plants. Genetica
Lycon, Switzerland.
Brown, A.H.D., M.T. Clegg, A.L. Kahler, and B.S. Weir
(ed.) 1990. Plant population genetics, breeding and genetic
resources. Sinauer, Sunderland, MA.
Brown, A.H.D., G.J. Lawrence, M. Jenkin, J. Douglass, and
E. Gregory. 1989. Linkage drag in backcross breeding in
Caldwell, K.S., J. Dvořák, E.S. Lagudah, E. Akhunov, M.C.
Luo, P. Wolters, and W. Powell. 2004. Sequence polymor-
phism in tetraploid wheat and their D-genome diploid
Chapman, V., T.E. Miller, and R. Riley. 1976. Equivalence of
the A genome of bread wheat and that of Triticum urartu.
Darwin, C.D. 1859. On the origin of species by means of
natural selection, or preservation of favoured races in the
Daud, H.M., and J.P. Gustafson. 1996. Molecular determina-
tion of a B genome progenitor of wheat (Triticum aesti-
icum) using a genome-specific repetitive DNA sequence.
Evaluation of nuclear DNA content and heterochromatin
changes in anther-derived diploids of tobacco (Nicotina
173.
sequence in response to chromosomal location. Genetics
177:359–374.
Dubcovsky, J., and J. Dvořák. 1995. Ribosomal RNA multi-
gene loci: Nomads of the Triticeae genomes. Genetics
140:1367–1377.
Dvořák, J., and J. Dubcovsky. 1996. Genome analysis of
polyploid species employing variation in repeated nucleo-
of genome analysis in plants. CRC Press, Boca Raton,
FL.
The structure of Aegilops tauschii genepool and the evolution
sources of the A genomes of wheats inferred from the
polymorphism in abundance and restriction fragment
length of repeated nucleotide sequences. Genome
30:680–689.
nucleotide sequences sheds light on the origin of the
wheat B and G genomes. Proc. Natl. Acad. Sci. USA
87:9640–9644.
phylogeny of the genus Triticum from variation in repeated
of endogenous and exogenous R2 elements in the rRNA
23:3825–3836.
repetitive DNA sequences in eucaryotes. Q. Rev. Biol.
70:297–320.
Fahima, T., J.H. Peng, J.P. Cheng, M.S. Roder, Y.I. Ronin,
maps in tetraploid wild emmer wheat, Triticum dicoccoides,
based on microsatellite and AFLP markers. Abstract. Isr.
Microsatellite DNA polymorphism divergence in Triticum
dicoccoides accessions highly resistant to yellow rust. Theor.
Fahima, T., G.L. Sun, A. Beharav, T. Krugman, A. Beiles,
and E. Nevo. 1999. RAPD polymorphism of wild emmer
wheat population, Triticum dicoccoides, in Israel. Theor.
Chapter 1  Wheat evolution, domestication, and improvement

Harlan, J.R. 1975. Crops and man. ASA, Madison, WI.
Harlan, J.R. 1992. Crops and man. 2nd ed. ASA, CSSA, Madison, WI.


Chapter 1  Wheat evolution, domestication, and improvement


