CHARACTERISATION OF DNA FROM ARCHAEOLOGICAL WHEAT (*Triticum L.*) SEEDS FROM ANATOLIA

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
THE MIDDLE EAST TECHNICAL UNIVERSİTY

BY

MEHMET SOMEL

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
IN
THE DEPARTMENT OF BIOTECHNOLOGY

SEPTEMBER 2003

Approval of the Graduate School of Natural and Applied Sciences

_____

Prof. Dr. Canan Özgen

Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

_____

Prof. Dr. Ayhan Demir

Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

_____	_____
Prof. Dr. Mahinur S. Akkaya	Prof. Aykut Kence
Co-supervisor	Supervisor

Examining Committee Members

Prof. Dr. İnci Togan	_____
Prof. Dr. Musa Doğan	_____
Prof. Dr. Mahinur S. Akkaya	_____
Assoc. Prof. Dr. Can Bilgin	_____
Assoc. Prof. Dr. Emel Oybak Dönmez	_____

ABSTRACT

CHARACTERISATION OF DNA FROM ARCHAEOLOGICAL WHEAT (*Triticum L.*) SEEDS FROM ANATOLIA

Somel, Mehmet

M.Sc. Department of Biotechnology

Supervisor: Prof. Dr. Aykut Kence

Co-supervisor: Prof. Dr. Mahinur S. Akkaya

September 2003, 99 pages

Ancient DNA analysis of archaeological wheat remains may serve to clarify unknown or controversial points in the history of wheat. In the first part of this study, extraction and amplification of DNA from Anatolian charred ancient wheat seeds obtained from different locations and ages was attempted. None of the our extraction samples yielded any PCR amplification. The possible reasons for this result were investigated by constructing an artificial charring experiment. The results suggest that the chances of obtaining DNA from the charred archaeological samples used in this study by the methods used are very low. Moreover, strong PCR inhibition by these charred seed extracts was observed.

The second part of the study aimed to develop new DNA based markers for ancient wheat DNA analysis. Markers linked to the brittle rachis character exhibiting domestication status were sought, but no result was obtained. Primers targeting plasmon sequences were developed and tested. A primer pair amplifying a 400 bp portion of the chloroplast TrnL-TrnF intergenic region was focused upon. A short piece of this region was amplified using ancient wheat DNA extracted in another study. This short piece appeared non-polymorphic

upon sequencing. The sequence spanning a wider portion of this region contained a number of length polymorphisms. Phylogenetic reconstruction using maximum parsimony showed that these polymorphisms were able to distinguish wheat taxa at the maternal ancestor level.

Keywords: Wheat, charred seed, ancient DNA, brittle rachis, domestication, Anatolia, PCR inhibition, length polymorphism.

ÖZ

ANADOLU KÖKENLİ ARKEOLOJİK BUĞDAY (*Triticum L.*) TOHUMLARINDAN ELDE EDİLEN DNA'NIN KARAKTERİZASYONU

Somel, Mehmet
Yüksek Lisans, Biyoteknoloji Bölümü
Tez Yöneticisi: Prof. Dr. Aykut Kence
Ortak Tez Yöneticisi: Prof. Dr. Mahinur S. Akkaya

Eylül 2003, 99 sayfa

Arkeolojik buğday tohumlarından elde edilen "arkeolojik DNA"nın analizi, bir tarım ürünü olarak buğdayın tarihindeki bilinmeyen veya tartışmalı noktaları açıklığa kavuşturmak için kullanılabilir. Bu çalışmanın ilk bölümünde farklı bölge ve tarihsel dönemlere ait Anadolu kökenli yanık buğday tohumlarından verimli bir şekilde DNA özütlemesi ve çoğaltılmasını amaçlamıştır. Özüt örneklerinden hiçbiri Polimeraz Zincir Reaksiyonunu (PZR) ürünü vermemiştir. Bu durumun olası sebepleri bir yapay yanık tohum deneyi düzenlenmesi suretiyle araştırılmıştır. Sonuçlar, bu çalışmada kullanılan yanık arkeolojik örneklerden eldeki yöntemlerle DNA elde edilme ve çoğaltılma olasılığının çok düşük olduğuna işaret etmektedir. Dahası, yanık tohum özütlerinin PZR'yi ciddi biçimde engellediği gözlemlenmiştir.

Çalışmanın ikinci bölümününde arkeolojik buğday analizinde kullanılmak üzere yeni DNA bazlı belirteçler geliştirmek amaçlanmıştır. "Kırılgan rakilla" karakterine bağlantılı ve dolayısıyla evcillik durumunu gösteren belirteçler aranmış ancak sonuç

alınamamıştır. Ek olarak, kloroplast ve mitokondri bölgelerini hedefleyen primerler geliştirilmiş ve denenmiştir. Kloroplast TrnL ve TrnF genleri arasındaki 400 baz çiftlik bir bölgeye yönelik bir primer çifti üzerinde yoğunlaşılmış, bu bölgenin kısa bir parçası ise, bir diğer çalışmada özütlenmiş arkeolojik buğday DNA'sı kullanılarak çoğaltılmıştır. Dizilimi okunan bu bölgenin genetik farklılık içermediği ortaya çıkmıştır. Bu bölgenin daha geniş bir bölümünün ise bir dizi uzunluk polimorfizmi içerdiği gösterilmiştir. Gerçekleştirilen filogenetik analiz, bu polimorfizmlerin buğday gruplarını en azından ana-ataları düzeyinde ayrıştırmada kullanılabileceğini ortaya koymuştur.

Anahtar Kelimeler: Buğday, yanık tohum, arkeolojik DNA, moleküler belirteç, kırılgan rakilla, evcilleşme, PZR engelleme, uzunluk polimorfizmi.

ACKNOWLEDGEMENTS

I would firstly like to thank both Prof. Mahinur Akkaya and Prof. Aykut Kence for their vital supervision, inspiration and encouragement.

All labmates in lab B20, namely Adnan, Elif, Ertuğrul, Figen, Mine, Tolga, Turgay, Yavuz and Yasemin gave enourmous support. Also those in the Department of Biology, especially friends in lab 37 and Gülen Ulusoy, were always ready to help.

I am obliged to Dr. Hatice Bilgiç for generously sharing all her results, and Dr. Erdoğan E. Hakkı for his precious assistance. I must mention Dr. Sachihiro Omura and Dr. Andy Fairbairn for providing and identifying the Kaman seeds, and Dr. Emel Oybak Dönmez for explaining archaeobotanical identification. I must also thank Dr. Angela Schlumbaum, Dr. Terry Brown and Dr. Mark Nesbitt for their willingness to communicate and for their suggestions.

The TÜBİTAK Directorate Of Human Resources Development has been supporting this study since March 2002.

Lastly, I am grateful to my parents, grandparents, my brother and all my comrades, without whose precious advice and support I would neither end up being a biologist nor a researcher. And of courıse Nazlı Somel, without whom it would be impossible for me to finish this work.

And I have to finally mention that, the working people of Turkey, our labour class, are worthy of much more refined research, much more enlightment, much better lives, which stands as our sole duty today.

TABLE OF CONTENTS

LIST OF TABLES

TABLE

LIST OF FIGURES

FIGURE

LIST OF ABBREVIATIONS

[α ^{32}P]-dATP	: [α ^{32}P]-adenosine triphosphate
µg	: Microgram
µl	: Microliter
aDNA	: ancient DNA
AFLP	: Amplified Fragment Length Polymorphism
BC	: years before Christ
BP	: years before present
bp	: base pairs
BSA	: Bovine Serum Albumin
Cox1	: Cytochrome c oxidase I
CSE	: Charred Seed Extract
CTAB	: Cethyltrimethylammonium bromide
ÇH	: Çatalhöyük
ds	: double stranded
dd	: double distilled
Glu	: Glutenin
HMW	: high molecular weight
kb	: kilobase
KK	: Kaman Kalehöyük
min	: minute
mM	: Milimolar
MYA	: Million Years Ago
ng	: Nanogram
PAGE	: Polyacrylamide Gel Electrophoresis
PCR	: Polymerase Chain Reaction
pmol	: Pico mole
PTB	: *N*-phenacylthiazolium bromide
RAPD	: Randomly Amplified Polymorphic DNA

rDNA	: Ribosomal DNA
RFLP	: Restriction Fragment Length Polymorphism
Rpm	: Rotation per minute
Ta16S	: *Triticum aestivum* 16S RNA gene
TrnL-F	: Intergenic region between the *Trn* Leu and *Trn* Phe genes
UV	: Ultra Violet
v/v	: volume/volume
w/v	: weight/volume

CHAPTER I

INTRODUCTION

1.1. Wheat and the spread of agriculture

1.1.1. The beginnings of agriculture

The start of agriculture, as a practice of selection and controlled reproduction of natural resources, has been the most prominent step in the transformation of our tool-using ancestors into modern human beings. The process of cultivation which lead to the domestication of wild species is the primary one that has enabled the production of economic surplus and the subsequent establishment of human civilizations.

The origins of agriculture in the archaeological record can be traced back to the Neolithic period around the 8th millennium BC (Van Zeist, 1992; Zohary and Hopf, 1994); findings also suggest that crop domestication developed more or less simultaneously with animal breeding (Harris, 1996). Although a relatively recent development in the history of mankind, arising much later than the start of tool-making about 2 million years ago (Lewin, 1999), the initiation of agricultural practice was in fact a consequence of a complex process, having ecological, cultural and economic aspects (van Zeist, 1992; ed. Harris, 1996), and in turn, it had drastic effects in all these areas. Moreover, it should be noted that this accomplishment would be impossible without the previous accumulation of the tools and practices of the hunter-gatherer societies during the preceding ages (the rapid progress of tool-making practices from 40,000 to 10,000 years ago is described in Lewin, 1999).

The so-called “Fertile Crescent”, the arc stretching from Northern Syria through Southeast Turkey, Northern Iraq and Western Iran, is the first known area where domestication started, where the “founder crops” emmer and einkorn wheat and barley were present in the wild (Zohary and Hopf, 1994). Agriculture became established in the Near East until 7000 BC (van Zeist, 1992). The practice later spread into North Africa and Europe, and there also occurred independent initiations in South East Asia and Central America. In all, agriculture became a world-wide phenomenon in a few millennia. When compared to the history preceding it, this rise and spread seem to have occurred in a relatively short time.

Unfortunately, the actual time and location of the start of agriculture of different species and how these practices were passed on to other communities at the Neolithic period are questions that remain mostly obscure, mostly due to the scarcity of archaeological material (van Zeist, 1992; Nesbitt, 2001). Similarly, the developments regarding the domestication of new species, development of new farming practices and artificial selection, and their exchange throughout human history -until the 20th century- are generally a matter of debate, or even unknown.

Today, in addition to archaeological, historical and ethnographical investigations, molecular genetic studies based on phylogenetic reconstruction of modern species’ and archaeological genetic material have become both popular and informative. Using these tools, interdisciplinary research will be able to answer more and more questions regarding the history of agriculture, and in turn, these will provide substantial clues for the understanding human history itself (*e.g.* Nesbitt, 1993). Moreover, knowledge on both the history of agriculture and the patterns of speciation and variation of domesticated species may have important implications regarding current agricultural and ecological issues, such as conservation and utilization of natural variation, or genetic modification of crop species.

1.1.2. The importance of wheat in the history of agriculture

Wheat is today considered the most prominent crop species in the world, with a yearly production that has exceeded 600 million tonnes during the last decade

(http://www.fao.org/inpho/compend/text/ch06.htm). In Turkey, wheat can be considered the most important crop from both from economic and cultural aspects: During the 1990's the production was about 20 thousand tonnes per year, and it comprised about 50% of total field crops; most of this produce is consumed domestically, such that wheat consumption in Turkey is among the highest in the world with yearly 200 kg per capita (Bilgiç, 2002).

During the last century wheat production has been increasing as a general tendency. This phenomenon is rather a reflection of new and more efficient farming practices and novel resistant and high-yielding species, rather than aerial spread (http://www.fao.org/inpho/compend/text/-ch06.htm) which has been the dominant trend since the beginning of wheat domestication in the Fertile Crescent. On the other hand, those countries which cannot keep pace with the technological developments, Turkey included, face the danger of losing their competitive position (Braun, 2000).

Apart from its critical position as a modern crop, and the importance of the study of wheat genetics to improve breeding programs, the domestication and later history of wheat has also been an attractive issue for scientific investigation. This is natural, when wheat's position as mankind's first cereal crop to be cultivated and domesticated (alongside barley) is considered. Actually many other cereals, such as sorghum and rice, were domesticated much later (Zohary and Hopf, 1994).

As einkorn and emmer -the early forms- and later as bread wheat -since historical times (Harlan, 1987)- wheat must have had a very important social influence. It has a highly nutritious character, such that besides being rich in carbohydrate, it has a protein content that exceeds all other cereal crops (http://www.fao.org/inpho/-compend/text/ch06.htm); hence it must have efficiently fed large populations and allowed more population growth.

Furthermore, and in partial contrast to barley, its domestication and later cultivation appears as not a simple but quite a complex process: Domestication events may have occurred more than once and in different places; the process itself has involved multiple wild species living in different habitats; and hybridizations between these species have also played a role, resulting in polyploidy.

Moreover, the cultivation of wild (undomesticated) wheat species and even collection from wild stands has continued until recently in relatively remote locations, including Anatolia, (Hillman and Davies, 1990) which suggests that perhaps, local processes of domestication may have continued for a long time after the initial events.

Wheat's robust quality as a crop has also allowed its spread throughout the world, and in these new areas wheat has been established as highly divergent landraces, while mating continued in search of better crops. On the other hand, the accumulation of variation has been reversed during the last centuries, and many landraces have been lost by leaving their place to standardized, more resistant and/or high yielding cultivars (Jones *et al*., 1996; Gregova *et al*. 1997).

1.2. The evolution, domestication and speciation of wheat

1.2.1. Classification of wheat

Wheat is the general name used for member species of two genera, namely *Triticum L.* and *Aegilops L.*, the former including the cultivated and domestic forms, and the latter comprised solely of wild cereal species. Species of *Triticum* and *Aegilops* are annual, dominantly self-fertilizing monocot grasses belonging to the *Gramineae* (=Poaecea) family of plants including the cereal crops: wheat, barley, rice, maize. These have a rather recent ancestry, and wheat and maize have diverged only 50-70 million years ago (Graur and Li, 2000). Inside the grass family, wheat is placed in the Pooideae subfamily and the Triticeae tribe, together with barley (Bilgiç, 2002).

Both *Triticum L.* and *Aegilops L.* contain diploid (n=7, 2n=14), tetraploid (2n=28) and hexaploid (2n=42) species. Tetraploid and hexaploid species are allopolyploids that have resulted from natural hybridizations between species with chromosome sets (genomes) too similar to pair normally during meiosis (Sauer, 1993). The haploid genomes of diploid species in *Aegilops L.* are denoted by D, C, M, N, S, S^b, S^1 T or U (Morrison, 2001), and this genera also comprises polyploid species with certain combinations of these genomes. In

Triticum L. only type A, B, D and G genomes and their combinations are found. All, except types B and G, are readily detectable in genus *Aegilops L.*, while these latter two are considered to have initiated from an ancestral S genome (see section 1.2.7.1.), from a member of the *Sitopsis* section of *Aegilops.*

Genus *Triticum L.* comprises more than 10 species with a wild distribution in Europe, the Mediterranean and Western Asia; genus *Aegilops L.* comprises 22 species distributed in western Mediterranean to central Asia and China (Watson and Dallwitz, 1999; Nesbitt, 2001). The exact number of species is not definite since there is no consensus on wheat classification (see Table 1.2 as an example), such that the same taxa are classified by different taxonomists as species or subspecies.

Actually, wheat taxonomy has been subject to much controversy and confusion, similar to other recently radiated and incompletely diverged and isolated genera (Morrison, 2001). The complications are further enhanced by the polyploidic feature. The taxonomy of *Triticum L.* has also been controversial due to the high number of locally cultivated and adapted populations (*e.g.* the case of *T. parvicoccum* discussed in Hillman, 2001; Morrison, 2001).

Another complicating point regarding wheat taxonomy is that *Triticum L.* appears to be a paraphyletic group. There are essentially no reproductive genetic barriers between the two taxa. (Kimber and Feldman, 1987). When *T. aestivum* is considered as an example, two of its three genomes are originated from *Aegilops* L. species, including the non-nuclear genetic material. This suggests that the conventional nomenclature may be somewhat anthropocentric. In fact, many taxonomists have hitherto attempted to combine the two taxa (Morrison, 1993).

In Table 1.1 a list of the *Triticum* species indicating ploidy level, domestication status, character regarding grain type and common name in English is given. In this study, the nomenclature of Dorofeev and Migushova (1979) will be used for *Triticum* species due to its relative simplicity (Nesbitt, 2001).

Table 1.2 gives the alternative nomenclature for five *Aegilops L.* species that are among the progenitors of domestic wheat or have been used as outgroups in this study. The nomenclature of van Slageren, 1994, will be used hereafter.

Table 1.1. Comparative classification table for *Triticum L.* classification using nomenclature of Dorofeev and Migushova (1979). Brittle= Hulled, fully brittle rachis. Semi-tough= Hulled, fully brittle rachis. Tough= Free-threshing, fully tough rachis.

Genome	Domestic status	Hulled/free-threshing character	Common name	Classifications
AA	Wild	Brittle	Wild einkorn	*T. urartu*
AA	Wild	Brittle	Wild einkorn	*T. boeoticum*
AA	Domesticated	Semi-tough	Einkorn	*T. monococcum*
AA	Domesticated	Tough	-	*T. sinskajae*
AABB	Domesticated	Tough	-	*T. aethiopicum*
AABB	Domesticated	Tough	Persian wheat	*T. carthlicum*
AABB	Wild	Brittle	Wild emmer	*T. dicoccoides*
AABB	Domesticated	Brittle	Emmer	*T. dicoccum*
AABB	Domesticated	Tough	Macaroni wheat	*T. durum*
AABB	Domesticated	Semi-tough	-	*T. ispahanicum*
AABB	Domesticated	Tough	Polish wheat	*T. polonicum*
AABB	Domesticated	Tough	Khorasan wheat	*T. turanicum*
AABB	Domesticated	Tough	Rivet wheat	*T. turgidum*
AAGG	Wild	Brittle	-	*T. araraticum*
AAGG	Domesticated	Semi-tough	-	*T. timopheevii*
AABBDD	Domesticated	Tough	Bread wheat	*T. aestivum*
AABBDD	Domesticated	Tough	Compact wheat	*T. compactum*
AABBDD	Domesticated	Semi-tough	-	*T. macha*
AABBDD	Domesticated	Semi-tough	Spelt	*T. spelta*
AABBDD	Domesticated	Tough	-	*T. sphaerococcum*
AABBDD	Domesticated	Semi-tough	-	*T. vavilovii*
AAAAGG	Domesticated	Semi-tough	-	*T. zhukovskyi*

Table 1.2. Comparative classification table for five species of *Aegilops L.*

Genome	Classifications		
	Hammer, 1980	**van Slageren,1994**	**Kimber &Sears, 1987**
CD	*Ae. cylindrica*	*Ae. cylindrical*	*T. cylindricum*
D	*Ae. tauschii*	*Ae. tauschii*	*T. tauschii*
DM	*Ae. crassa*	*Ae. crassa*	*T. crassum*
S	*Ae. speltoides* subsp. *speltoides*	*Ae. speltoides* var. *speltoides*	*T. speltoides (speltoides*)
UM	*Ae. lorentii*	*Ae. biuncialis*	*T. macrochaetum*

1.2.2. Wild, primitive and modern wheat

As can be noticed from Table 1.1, the *Triticum L.* genus comprises both wild and domesticated wheat species. The distinction between the two will be discussed in depth, but here it will suffice to say that the main distinctive feature of domestic forms, in contrast to wild species, is their lack of capacity to distribute their seeds. Thus for their survival they are dependent on human action, *i.*e., threshing of the heads and sowing the seeds on tilled land. More specifically, their tough rachis does not spontaneously disarticulate, whereas the brittle rachis of wild species allows the seeds to disperse during the maturation of the head (Zohary and Hopf, 1994).

A further distinction among domesticated *Triticum* species leads to their classification as "primitive" or "modern". Modern wheat include tetraploid macaroni wheat (*T. durum*) and hexaploid common bread wheat (*T. aestivum*). These are also called "free-threshing" or "naked" wheat: Upon threshing, the rachis segments stay attached to each other, while the glumes and the other parts of chaff break apart, releasing freely the grain (Nesbitt and Samuel, 1996).

Primitive wheat, on the other hand, include einkorn (diploid), emmer (tetraploid) and spelt (hexaploid). Their grains, collected upon threshing, are surrounded by a hull, and hence their other name: Hulled wheat (wild wheat also contain this hull, and thus can be called the same name). When a spike of hulled wheat is threshed, it breaks up into its component spikelets, each consisting of tough glumes attached to a rachis segment and enclosing one or more seeds. The hulled character is the result of two differences in the structure of the spike: the semi-brittle joints between the rachis internodes, and the toughened glumes (Nesbitt and Samuel, 1996). These are relatively primitive features, remnants of the wild character, and have obviously been selected against in the lineages leading to the "modern wheat". On the other hand, genetic introgression from wild or primitive wheat to modern wheat has also played a role in hulled wheat evolution as will be discussed later (see section 1.2.7.4).

Cultivation of primitive wheat has continued together with modern wheat, and especially in relatively harsh -cold or wet- environments. This can be attributed to the robust

and resistant character of hulled wheat (Nesbitt and Samuel, 1996). Meanwhile, hulled wheat have been dramatically losing their position as crops in recent times (Karagöz, 1996).

1.2.3. Selfing and polyploidization

A number of characters of *Triticum* L. can be determined as important for their utilization in domestication. One is their predominantly self-pollinating character, which is actually shared by all "founder crops" and most of the contemporary crop species. It provides numerous advantages: The primary one is the isolation of species from the wild relatives, thus facilitating selection for domestication. Secondly, selfers provide the opportunity to establish true breeding lines from every individual, which allows easy and efficient breeding (Zohary and Hopf, 1994). Besides, Zohary (1971) notes that being a predominant selfer (with a less than one percent out-crossing frequency) in contrast to obligatory selfers, must have contributed to the accumulation of genetic variation and again enriched the material for breeding.

Another feature of wheat that must have facilitated its domestication and artificial selection is its being an annual species.

A third character of the *Triticum-Aegilops* genera has also greatly affected their evolutionary history. This is the tendency towards speciation by polyplodization, which has occurred naturally including both wild and domestic species (Briggle, 1967). In fact this is a common feature of many domesticated species, from potato to sugar cane (Gepts, 2002); perhaps most importantly, it leads to bigger cells and bigger plants. Although many polyploids being born spontaneously probably cannot survive due to problems in meiosis, there have also arisen quite a number of successful polyploids. Mechanisms of genome stabilization (Belyayev, 2000; Rieseberg, 2001) have certainly played a role; moreover selfing must also have helped in preserving the new chromosomal structure. On the other hand, being an allopolyploid was probably an important opportunity for a domesticated species, while alloployploid species carried multiple copies of each allele rather than two. This should have allowed for more thorough selection (Wendel, 2000) and robust characters.

We can discern two important polyploidization events in the history of cultivated wheat (see section 1.2.7.1. and Table 1.1): The first has been the hybridization that bore wild emmer, *T. dicoccoides*, most probably between *T. urartu* and an ancestral type of *Ae. speltoides*, which occurred long before cultivation (Bahrman *et al*., 1988; Miyashita *et al*., 1994). The second has occurred quite recently: the contribution of the D genome to a cultivated tetraploid by *Ae. tauschii* (Hillman, 1975; Dvorak *et al*., 1998) to yield the hexaploid bread wheat, and here, the tetraploid species has been the maternal donor. But this fusion has not been the last, and more recent hybridisations or introgressions have also been postulated or proved.

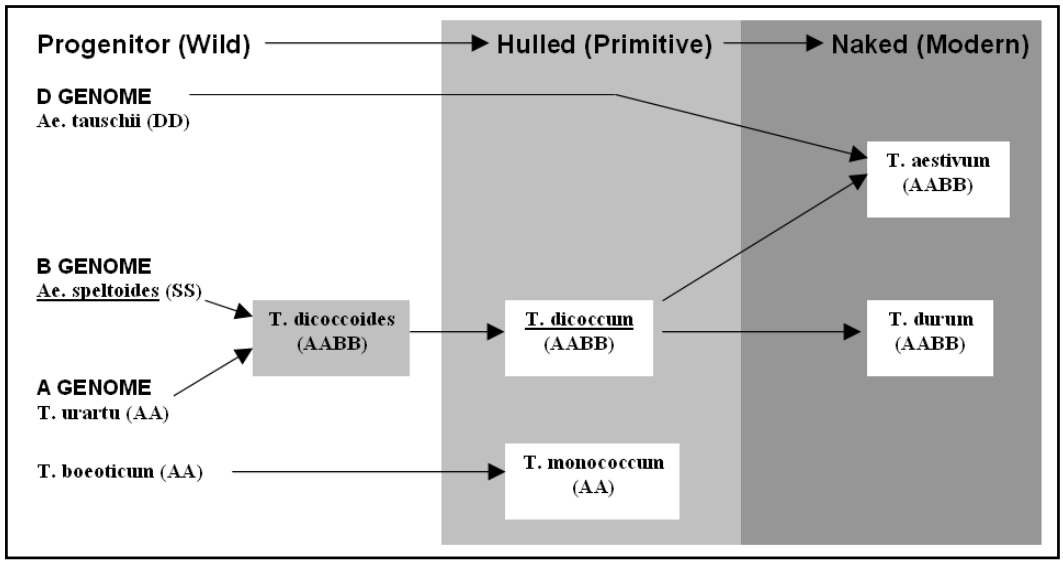

Figure 1.1. A summary of the polyploidy and speciation of major *Triticum* species. The genome compositions are given, and the maternal donors in hybridization events are underlined.

1.2.4. The initiation of wheat cultivation

As mentioned earlier, the events that led humans to start cultivation must have been complex. On one hand is the changing distribution of wild food species, in our case cereals, on the other, we have the changing gatherer practices of humans and their demographic dynamics. There must have been a series of special coincidences that led to humans to adopt cereals for cultivation between 9000-8000 BC in the Near East, but not earlier, not anywhere else, and not any other food species.

The pollen records would have been useful to follow the change in wild cereal distribution; if only wheat species were not selfers and withheld their pollen! As for plant and other remains unearthed at archaeological sites (see Table 1.3 and Figure 1.2), neither do these convey sufficient data, except the general location of the first events: The Fertile Crescent. In short, the evidence at hand is quite scarce.

One parameter that must have been decisive for both wild wheat and human history is the climatic change following the Glacial Maximum, *i.e.* a tendency of global warming and increasing climatic fluctuations. But there seems to be quite a bit of controversy regarding the exact dates and characteristics of these climatic changes (van Zeist, 1992; Hillman, 1996).

Nevertheless, Hillman (1996) provides a comprehensive description of a probable scenario: Following 13,000 BC, a climatic change that led to oak woodland expansion in northern Fertile Crescent appeared, in northwest to southeast direction (which can be traced from the pollen record). It was probably accompanied by a faster spread of certain cereal species including wild einkorn, rye and barley (this deduction is made by considering the floral succession patterns). This spread would have changed the food-resource base of hunter-gatherer communities, which had hitherto depended on a richer but scarcer diet based on roots, shrublet seeds, woodworm seeds, etc. The dense stands of einkorn, rye and barley and their large seeds (which also enhanced their competence against other grass species) provided a very calorie-rich source for humans. This, in return, resulted in an increase in population carrying capacity, a tendency towards sedentism and population growth. A second point is that the climatic change included climatic unpredictability and seasonal fluctuations, which might have led humans to elaborate their food storage practices.

It can be expected that such a population increase, although at first compensated by increased food gathering, would eventually end up in food stress. The response to stress could have taken different forms, but clearing stands where cereals did not naturally grow by burning weeds and sowing seeds would have been an alternative, which could have been developed by braking up the soil. Hence cultivation could have started in the period between 12,000-9,000 BC, and most probably initiating in the west of the northern Fertile Crescent, where the climatic change first showed effect (Hillman, 1996). Einkorn and emmer were among the initial crops to be cultivated.

1.2.5. The domestication of wheat

Domestication of a species can be defined as a genetic conversion of a wild species leading to the systematic control its reproduction and selection. The domestication of wheat has thus involved the fixation of two main genetic based changes: **i)** The loss of capacity of spike and grain dissemination, thus independent reproduction, **ii)** the loss of seed dormancy, which has allowed sowing of seeds at different times of the year. The latter development has been related to the thinning of the seed coat (Zohary and Hopf, 1994). In addition, plants with bigger and more numerous seeds have also been selected for.

But the former character has been paramount in the process of domestication, similar to many other species (*e.g.* in legumes, the other group of founder crops in the Fertile Crescent, this change has involved non-dehiscent pods (Zohary, 1999)). The specific modification in wheat has been the transformation of the a) wild type fully-brittle rachis, which allows the dissemination of spikelets during maturation, b) to a semi-brittle rachis, such as that found in emmer and spelt, and due to which the head can be easily collected while reaping with a sickle, c) and lastly to a fully tough rachis, that of naked bread wheat and durum wheat, which does not disarticulate whatsoever.

There are different views and finding as to how many and which genes are controlling this brittle-rachis character in wheat, but like many other species, it is believed that there are more than one (Sharma and Waines, 1980; also discussed in section 1.3.1). It is also unlikely for such a mutation to be frequent, while it is deleterious in the wild.

Nevertheless, it has been shown that once present in the field, a recessive mutation could be easily selected for under appropriate circumstances, thanks to the self-pollinating character of wheat (Gepts, 2002).

Hillman and Davies (1990) have performed a computer simulation and field experiments to test how much time it would take for cultivation to succeed in a fully non-brittle rachis population. Their most important and probably valid assumption is that the human communities engaged in cultivation were unaware of the non-brittle rachis type or at least its possible advantages; thus the process was one of unconscious selection. Secondly, pre-cultivation domestication (*i.e.*, while gathering activity continues) is considered unlikely in the case of wheat (in contrast to the contentious case of legumes).

In the mentioned study, Hillman and Davies have showed that, there are four possibilities to occur once a mutant form is present:

i) If humans used harvesting methods that would positively select for non-brittle rachis mutants, such as sickle reaping, and varying with such options as to how much harvested seed was sown, and whether to virgin plots or the same field, it could be expected for domestication to occur in 20-30 to 200-300 years, once the mutant was present in the field.

ii) If the same conditions were valid but harvesting was done before head ripening, there would not be a positive selective pressure for mutants, and thus this "pre-domestication" cultivation could extend for a much longer time.

iii) If the communities used a method which does not advantage but even disadvantages non-brittle rachis mutants, such as collecting grains by beating them in a basket, then this "non-domestication cultivation" could last indefinitely. It can be added that this last option seems to waste more grains that the former two.

iv) Recognition of the mutant character could have accelerated the process to finalize in a few years. The authors consider a mutant frequency of 1-5% as threshold for

such awareness. In this case, the duration of full domestication could be even as short as 10 years. But in my view, this threshold could even be lower, because people were probably familiar with a non-brittle rachis phenotype if the heads were reaped unripe.

So the problem seems to be mostly based on a cultural-economic one, which we have only meagre evidence from the archaeological record. Nevertheless, it is logical to suppose that domestication of a species was achieved in a limited number of communities, a point that is also supported by experimental genetic data (to be later discussed), and that other groups subsequently adopted these new mutant crop species from the former. This process is called “diffusion” (Harris, 1996).

Accepting this scenario, we face the problems of exactly where and when the cultivator communities first domesticated the so-called founder crops: Einkorn and emmer (and barley). Although it has been strongly argued that agriculture began in the Jordan Valley (Smith, 1995) based on some of the evidence shown in Table 1.3, the picture taken as a whole more strongly supports northern Syria and southeast Turkey (Nesbitt and Samuel, 1996).

Another question could be on how rapid and through what kind of mechanisms they diffused, the answers of which could have implications on the social relations at that time between communities.

Table 1.3 and Figure 1.2 give a list of archaeological sites where charred remains of wheat and barley have been found. As can be seen from the table, the meager amount of information does not make a direct suggestion. One can add that indirect evidence, such as crop processing tools, are also scanty and contentious (Hillman and Davies, 1990).

1.2.6. The subsequent history of wheat

As has been described, einkorn and emmer were the two wheat species that were first domesticated. Their domestication followed their spread as crops throughout the Fertile

Crescent, to nearby regions such as the south of the Caspian Sea or the Nile Valley, and from these areas to other parts of the world, especially central and east Asia and Europe (Zohary and Hopf, 1994). The sharing of crops between sedentary farmers must have occurred simultaneously with migratory farmers carrying their crops.

During this process, when farming reached the eastern habitat of *Ae. tauschii*, the abovementioned fusion of the D genome to a tetraploid cultivated wheat occurred, resulting in bread wheat. The spread thus continued, with increasing number of cultivated or hybrid species, and a much more rapidly increasing level of intraspecies diversity (Qualset, 2001).

This latter process would be driven by selection of varieties by farmers for better agronomic traits. One such example is the easily detectable increase in grain size, which helps to distinguish wild and domestic species of today (Nesbitt, 2001). Other traits such as bread making quality must have attained importance in later periods with cultural changes.

Local adaptation of wheat species also should have continued in habitats differing from the original one, producing an enormous number of landraces. In later periods, as a result of increasing population pressure and growing ties between agricultural producers, the need to develop better varieties allowed plant breeding to become an important practice, which involved both selection within and hybridisation among these landraces (Qualset, 2001). Today, there are over 25,000 varieties of wheat (Feldman *et al*., 1995). This process, together with the previously mentioned among-species hybridisation and introgression events, should be expected to have continuously reshaped the genetic makeup of domestic wheat. And the last global change that has been taking effect since the last two centuries in Europe (Jones *et al*., 1996) and most probably since the mid-20th century in Turkey, has been in the reverse direction: the rapid decay of genetic variability by the introduction of standard high-yielding or resistant varieties (Gregova, 1997).

Today, it is tempting for many to investigate this history more in depth. For example, by using the developed tools at hand and global information, it might be possible to construct a dendogram describing the domestication and later genetic changes that occurred in wheat, and to try to relate this with the change of human social structure, of climate, etc.

Table 1.3. Selected archaeological sites in the Fertile Crescent and wild and domestic wheat occurrences. EI = einkorn. EM = emmer. NAK = naked wheat, and any detected ploidy level denoted in parentheses. G = identification based on grain. C = identification based on chaff. ? = uncertain identification. From Nesbitt (2001).

Site (phase)	**Country**	**Date (uncal BC)**	**Economy**	**Wild**		**Domestic**		
				EI	EM	EI	EM	NAK
Ohalo	Palestine	17000	Foraging		C			
Abu Hureyra (I)	Syria	9500-8000	Foraging	G				
Mureybit	Syria	8500-7600	Foraging	G			C ?	
Qermez Dere	Iraq	8200-7700	Foraging	G				
Netiv Hagdud	Palestine	8000-7400	Foraging		C			
M'lefaat	Iraq	7900-7700	Foraging	G				
Jerf al Ahmar	Syria	7800-7700	Foraging	C				
Dja'de	Syria	7600-7000	Foraging	C ?				
Abu Hureyra (2A)	Syria	7500-6000	Farming	G		G	C	C
Beidha	Jordan	7200-6600	Farming			C	C	
Cafer Höyük (XIII-IX)	Turkey	7200-?7000	Farming	C	G	C	C	
Jericho	Palestine	7200-6800	Farming			G	G	
Nahal Hemar (3-4)	Palestine	7100-6000	Farming				C	
Nevalı Çorı	Turkey	7200	Farming			C	C	
Çayönü (g-c)	Turkey	7000-6700?	Farming	G	G ?	C	C	
Tell Aswad (II)	Syria	6900-6500	Farming			G	G	C
Aşıklı Höyük	Turkey	6900-6500	Farming	G		C	C	C
Wadi el-Jihat (7)	Jordan	6800-6400	Farming	G	G ?	G	G	
Ghoraife	Syria	6800-6200	Farming	G		G	G	C (4, 6)
Halula	Syria	6700-6600	Farming		C	?	C	**G**
Can Hasan III	Turkey	?6500-6200	Farming	G		G	C	C (4, 6)
Cafer Höyük (III-IV)	Turkey	6600-5800	Farming	C		G	C	C (6)
Abdul Hosein	Iran	6500	Farming				G	
Ras Shamra (Vc)	Syria	6500-6000	Farming				G	
Jarmo	Iran	6400?	Farming	G	C		C	
Ali Kosh (BM)	Iran	?6400-6000	Farming	G		C	C	
Tell Bouqras	Syria	6350-5850	Farming	G		G	G	C
Tell Ramad (I)	Syria	6200-6100	Farming	G		G	G	C
Wadi Fidan A	Jordan	6000	Farming			G	G	
Wadi El-Jihat 13	Jordan	5900-5800	Farming	G		G	G	
El Kowm II-Caracol	Syria	5800-5700	Farming			C	C	C (4, 6)

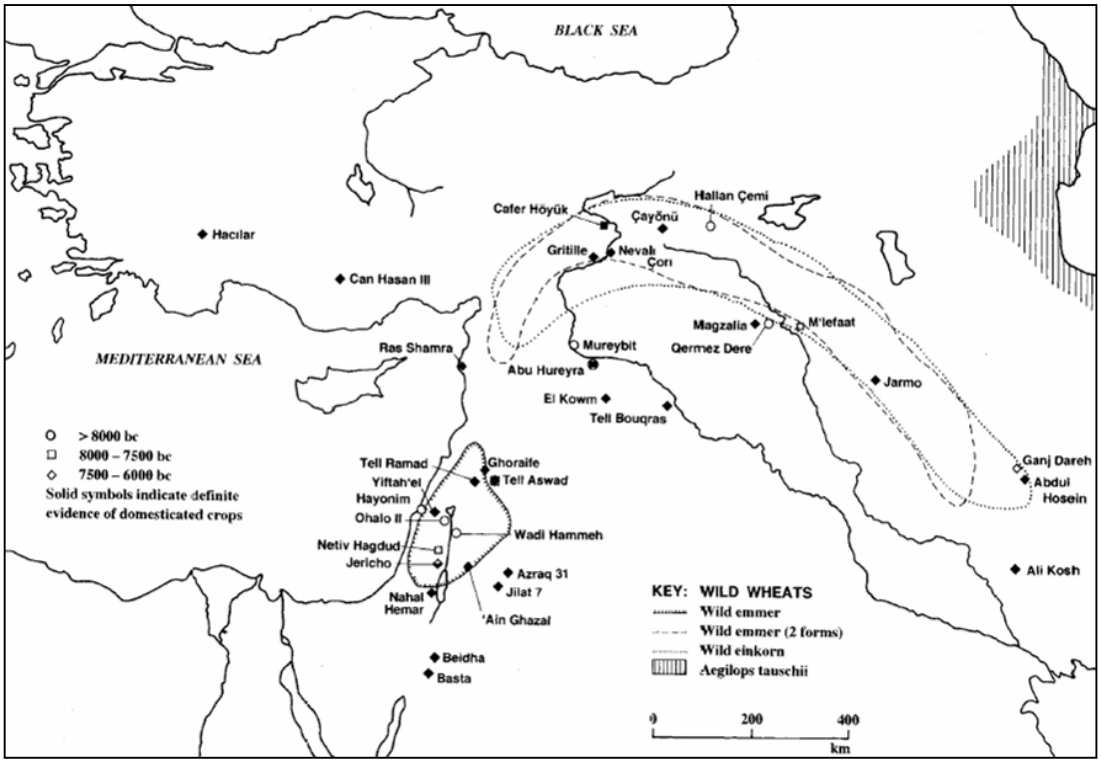

Figure 1.2. Map of Near Eastern archaeological sites and the wild distributions of wheat species *T. monococcum* (einkorn), *T. dicoccum* (emmer) and *Ae. tauschii*. From Nesbitt and Samuel (1996).

1.2.7. Points of obscurity in wheat domestication and evolution

Whereas the general picture described above and depicted in Figure 1.1 is today well understood, there are more than a few controversies and unknown facts regarding this complex process of wheat evolution, some of which are listed below:

1.2.7.1. The wild ancestry of domestic species

Since the last few centuries, studies on this issue have resolved most of the debates. *T. urartu* has been more or less confirmed as the A genome donor, three studies pointing to this conclusion are Dvorak *et al.* (1992), Miyashita *et al*. (1994) and Yıldırım (2001). But still some do not seem to be totally satisfied (*e.g.* Huang *et al*., 2002).

The majority of studies on the emmer ancestor issue have suggested *Ae. speltoides* as the B and G genome donor, and moreover the maternal parent of both *T. dicoccoides* and *T. araraticum*, respectively (Bahrman *et al*., 1988; Dvorak and Zhang, 1990; Miyashita *et al*., 1994; Wang *et al*., 1997). But the suggestion of *Ae. speltoides* is based on the relative distance of this species to the B genome compared to other grass species, which, among the three genomes of bread wheat, is the most diverged from its ancestor and, according to some studies, the most variable (Talbert *et al*., 1995; Huang *et al*., 2002). Thus, this point might also require further clarification (*e.g.* Mori *et al*., 1997).

As for the D genome donor, the ancestor has been firmly determined as the Asian *Ae. tauschii*, which has a very wide distribution (Hillman, 1975; Dvorak *et al*., 1998). Nesbitt (2001) suggests the western *Ae. tauschii* ssp. *strangulata*, which is found in Transcaucasia and the Southeast Caspian region (see Figure 1.2) as the original donor, in contrast to the initial candidate, ssp. *tauschii* of southeast Turkey and west Iran; but extra evidence may be needed to verify this important point.

1.2.7.2 The cultivation and domestications of founder crops

In 1921 the Soviet scientist Vavilov made an important proposal for his time to track the initiation of domestication. He put forward that the “centre of diversity” of a cultivated crop should be where its domestication started, based on the assumption of increasing diversity with time. Following this hypothesis, he has suggested the Fertile Crescent as one of the centres of domestication. This hypothesis is of limited value today, while we have substantial archaeological knowledge compared to Vavilov’s time, and the more important

problem has become to pinpoint the domestication events for each species (Harlan, 1992; Harrris, 1996; Heun *et al.*, 1997; Özkan *et al*., 2002; Allaby and Brown, 2003).

This will be the hardest task of all, if anyhow accomplished. Specific questions that can be asked regarding the history of wheat are:

i) When did the cultivation of einkorn and emmer start?

ii) When did it first yield domesticated crops; how much time elapsed in between the two stages?

iii) Did independent domestications of one species occur, or was diffusion the sole factor of agricultural spread? In other words, were the domestications of these species paraphyletic or monophyletic?

iv) Were both einkorn and emmer domesticated sympatricly and/or simultaneously, as a "package of crops"?

v) Did the domestication of one species contribute to another one's domestication?

vi) Are all domesticated species represented in the present day genepools, or did some go extinct?

vii) What was the relation of the domestication of einkorn and emmer with the domestication processes of other "founder crops" of the Near East, namely barley and the legume species?

As discussed above in section 1.2.5, archaeological evidence gives only a weakly satisfactory answer to these questions. First of all, the discrimination of wild and *early* domestic species is not usually possible from charred grains (van Zeist, 1992), and chaff is usually the only possible evidence (Nesbitt, 2001). Also, gathering of wild species continued for long periods after domestication (as seen from Table 1.3), which might complicate issues.

There is limited evidence from studies on modern species, and these also face problems. The AFLP study of Heun *et al*. (1997), for example, pointed out to a single origin of *T. monococcum*. But such examples are viewed with suspicion by Allaby and Brown, 2003, who show through a computer simulation that the genetic diversity in diphyletic domestic species may converge and lead to flawed conclusions of monophlyly in AFLP analysis. On the other hand, the specific *T. dicoccoides* population giving rise to cultivated emmer, or the cultivated emmer that took part in the fusion leading to bread wheat are not known. Other evidence regarding monophyly vs. polyphyly, or genetic bases of change under cultivation are very limited, or sometimes controversial (see section 1.3.1).

Another area of study has been the domestication related characters. By making numerous crosses between cultivars and searching for any wild type progeny, the recessive mutations controlling the brittle rachis character of domestic einkorn and domestic emmer have been shown to reside in a single gene (Zohary, 1999). On the other hand, there are until now two loci shown to control fragility of the rachis, one on chromosome 2 and another on 3 (Sharma and Waines, 1980; Peng *et al*, 2003; Watanabe and Ikebata, 2000). These results suggest monophyly, but the exact mutation(s) are not known, and thus we are still unable to know whether the mutation is really monophyletic. Neither do we have information regarding the mutation rates in these genes, which is an important parameter for the estimations of domestication rates. Hillman and Davies (1990) have assumed a net forward mutation rate of 10^{-6} among wild population at this locus, but seemingly with no experimental basis.

On the other hand, the same crossing experiments on barley have shown that there may be more than one gene involved, and thus there could have occurred two or more domestication events for barley (Zohary, 1999). This result contrasts with the other studies based on DNA variation and suggesting monophyly.

As for the last question listed above, it is known that especially in the Caucasian region -most probably due to its proximity to the initiation of agriculture and the mountainous landscape- many endemic domesticated species have been cultivated locally (Nesbitt, 2001). Such situations may be even more widespread, which is left to collectors and taxonomists to reveal.

1.2.7.3. The modes of diffusion, spread and breeding

For the initial stage of the spread of agriculture, any answer to the questions listed would contribute to the understanding of human history, as well as the history of wheat. One is the question of the mode of diffusion: During the spread of wheat, did the farmers predominantly share their domesticated crops, by what is called “secondary” or “cultural” diffusion, or did diffusion show correlation with human migratory actions, *i.e.* "primary” or “demic” diffusion (Harris, 1996)? Although the first case seems more likely, there have been attempts to correlate genetic and cultural relationships of human communities with local varieties, in support of the second alternative (Nesbitt and Samuel, 1996; ed. Harris, 1996)

Other questions can also be asked: Was diffusion a single event, or a pulsed series of events? More specific problems are also tackled, such as from which routes domesticated species reach Europe, from the Balkans or *via* Transcaucasia?

As for later stages, wheat history again bears the potential of shedding light on human history; *e.g.* how much of the grain sown was home produce? Did the establishment of regional states (such as the Hittites, Babylonians, Romans, *etc*.) affect the varieties sown (Jones, 1996)? These are all questions which are very difficult to answer by tracking local wheat varieties or the charred grain record, although other archaeological data might also help.

A further issue that could be dealt using a comprehensive wheat dendogram would be regarding wheat breeding in a regional or later global scale. For example, in the history of agriculture, did there ever occur a period of intensive wheat breeding, rapid changes in certain agriculturally adaptive genes?

1.2.7.4. The origin of bread wheat and spelt as two specific cases

The problem about the origin of bread wheat stems from the three facts i) the present day distribution of the claimed D genome donor *Ae. tauschii* ssp. *strangulata* starts from the southeast Caspian, ii) there are no signs of agriculture in this area until 6000 BC

(uncalibrated), iii) free-threshing hexaploid wheat is found in Can Hasan III (6400-5700 BC, uncalibrated) and Cafer Höyük III and IV (7000-6200 BC, uncalibrated) (Nesbitt, 2001). Although a simple explanation would be that the distribution of ssp. *strangulata* could have had a more western border about 10,000 years ago, this has to be resolved more thoroughly.

Another dilemma regarding bread wheat is that it was shown to carry a polymorphism at an RFLP site (*Xpsr920*) where one allele was nearly fixed in wild emmer whereas another allele was fixed in all cultivated emmer (Dvorak and Luo, 2001). Bread wheat populations from Eurasia contained both alleles, with the frequency of the wild emmer-type allele increasing towards the west. This suggested two things, i) all cultivated emmer carrying the same mutant allele points to their monophyletic origin, ii) there should have occurred a second hybridisation between wild emmer and bread wheat to yield the observed polymorphism in bread wheat, bearing in mind the fact that the D genome points out to the monophyletic origin of hexaploid wheat also (Dvorak *et al*., 1998). Dvorak and Luo (2001) propose Turkey as a possible location of this putative hybridisation.

The second case, spelt (*T. spelta*), is a hulled hexaploid wheat. It is widely cultivated in Europe since the Early Bronze Age (Nesbitt and Samuel, 1996). Its relative resistance against cold and against pre-harvest sprouting seems to have bestowed it popularity among European farmers suffering from a cold and wet climate. It is detected in European sites much later than the initial introduction of free-threshing bread wheat in the Neolithic PrePottery B, and the findings of spelt in the Near East are very rare and controversial (Nesbitt, 2001). Thus, initially, spelt was considered to be a European species. It was believed to have arisen as a hybrid between cultivated emmer and free-threshing bread wheat, the possibility of which was shown empirically.

Later, spelt was found to be cultivated in Iran, but this Asian spelt was shown to be only distantly related with European spelt by isozyme variation (Jaaska, 1978). More extensive RFLP studies of the D and A and B genomes by Dvorak *et al*. (1998) and Dvorak and Luo (2001) showed that European spelt was genetically most closest to European bread wheat, Turkish bread wheat, European domesticated emmer and Turkish domesticated emmer, respectively. These studies distinguished between Asian and European spelt. These results also suggest that European spelt was not a product of simple hybridisation, but rather introgression of European cultivated emmer into European bread wheat. Another study using

the high molecular weight glutenin locus by Blatter et al. (2002) has further suggested that bread wheat and European spelt had different ancestry.

In short, the origins and evolution of these two species deserve more thorough investigation and more robust evidence.

1.3. Analysis methods of wheat history

1.3.1. Analysis of contemporary evidence

Contemporary evidence regarding the history of wheat domestication and evolution resides in the genetic relationships between progenitor and domesticated species, and among domesticated varieties. The quest is aimed to find the closest relatives and the modern distributions of these. From such information it may be possible to infer the location of historical domestication events and paths of spread.

Besides, Zohary (1999) points out that these investigations can clarify the issue of monophyly vs. polyphyly through three potential sources of evidence: **i)** founder effects in the domesticated species' genepool as revealed by low intraspecific variation in the domestic species compared to the progenitor; **ii)** presence of mutations in a single domestication gene, while many loci of similar effect are present; **iii)** the absence of domestication of species closely related and/or very similar to the wild progenitor (sibling species).

Morphological, physiological or ecological comparisons are the oldest and most basic tools (*e.g.* Damania *et al*., 1997; Valkoun, 2001). But the confusing effect of homoplasises (parallelisms and convergences), and the difficulty in distinguishing environmental effects form heritable ones, are important handicaps. These disadvantages have been partially compensated by molecular genetic analysis during the last decades. Such analysis uses molecular markers, *i.e.* inherited and easily detectable molecular characters exhibiting polymorphism at a certain taxonomic level, and/or showing linkage to particular phenotypic traits. In addition to the advantages of being inherited, molecular markers are

also superior over other markers due to the ability of being evaluated more objectively (Graur and Li, 2000).

An important and early method has been cytogenetics, developing since mid-20th century. The detection of ploidy levels and homologies through DNA-DNA hybridizations (*e.g.* Badaeva *et al*., 1986) has proved particularly informative (Morrison, 2001). Another early tool has been the study of protein polymorphisms, or isozymes (*e.g.* Nevo *et al*., 1988; Waines, and Payne, 1987).

On the other hand, none of the above mentioned methods have proved to be as versatile and precise as DNA sequence based molecular markers. The best advantage is that it allows study at very different taxonomic levels, from intrapopulation investigations to those among kingdoms. Secondly, differentiations are much more precisely detected, such as genetic diversity among and within species. This allows better phylogenetic reconstruction, which can help in determining founder effects. Sequence data also enables estimation of divergence times between taxa.

Study of genetic polymorphisms on the DNA sequence level has taken start with RFLP analysis (*e.g.* Dvorak *et al.*, 1988; Mori *et al*., 1995). Since the last decade, AFLP and SSR analysis have further broadened the area of overall genetic comparison between wheat species and cultivars (*e.g.* Becker *et al*., 1995; Fahima *et al*., 1998; Bohn *et al*., 1999; Yıldırım, 2001; Bilgiç, 2002). On the other hand, PCR based sequence analysis has been speeding up, perhaps hindered by costs, but much more informative and precise than the previously mentioned methods. The internal transcribed spacer (ITS) region (*eg*. Zhang *et al.*, 2002), high molecular weight glutenin gene (*e.g.* Allaby *et al.*, 1999) and chloroplast ATP synthase genes (*e.g.* Ikeda *et al*., 1992) have been some widely used loci.

Three prominent work can be mentioned here. One is that by Heun *et al.* (1997), which showed by AFLP analysis that a single wild einkorn population in the whole Fertile Crescent –a wild einkorn population form Diyarbakır-Karacadağ in Turkey- was most similar to domestic einkorn. It thus approximately suggested the location of einkorn domestication (assuming limited change in the wild einkorn distribution) and also proposed monophlyly of einkorn domestication. A more recent AFLP study (Özkan *et al.,* 2002) again suggested southeast Turkey as the core area of emmer domestication.

In yet another study, the genepools of all ABD hexaploid species were shown to share a common D genome using RFLP markers, which suggested that there had occurred a single hybridisation event (Dvorak *et al*., 1998) and all ABD hexaploids share a common ancestor.

In addition to studies of genetic variation, workers in this field have examined domestication related traits by classical breeding methods and molecular mapping. Among other points, the results on the brittle rachis character have contributed to the discussion on whether domestic wheat are monophyletic or polyphyletic, which are described in Zohary (1999) (see section 1.2.7.2.) Although these generally suggest monophyly for at least einkorn and emmer, the results are still rather superficial while the genes themselves have not been identified. Also, the result from these studies suggesting a polyphyletic origin for barley is conflicting with a AFLP based study concluding a monophyletic origin in the Jordan Valley for barley (Badr *et al.*, 2000).

Estimation of divergence times based on nucleotide substitution rates (Nei and Kumar, 2000) of wheat species has also been sought by different studies. This would be very useful for wheat history studies, especially if recent events could be discerned. Unfortunately, the results have not been very precise or consistent, as shown in Table 1.4.

Table 1.4. Four publications giving estimated divergence times for *Triticum*, A*egilops* and *Hordeum* (barley). MYA = million years ago.

Authors	**Diverging taxa**	**Estimate (MYA)**	**Locus used**
Ogihara *et al*. (1991)	*Triticum* and *Aegilops*	1.5	Chloroplast Hot Spot region
Ogihara *et al*. (1991)	(*Triticum-Aegilops*) and *Hordeum*	10	Chloroplast Hot Spot region
Ikeda *et al*. (1992)	*Triticum* and *Aegilops*	3	AtpB
Ikeda *et al*. (1992)	(*Triticum-Aegilops*) and *Hordeum*	8	AtpB
Allaby *et al*. (1999)	A, B, G and D	5.0 - 6.9	HMW Glutenin
Allaby *et al*. (1999)	B and G	2.5 – 3.5	HMW Glutenin
Huang *et al*. (2002)	(*Triticum-Aegilops*) and *Hordeum*	10.8 – 12	Acetyl-CoA carboxylase and 3-phosphoglycerate kinase
Huang *et al*. (2002)	A, B, G and D	2.5 – 4.5	Acetyl-CoA carboxylase and 3-phosphoglycerate kinase

Usually the substitution rate calculations are based on Wolfe *et al*. (1987) and Wolfe *et al*. (1989). The last study has tried to avoid over- or underestimations by using multiple sources, but still, paucity of the fossil record allows for very general estimates only. In addition to this general handicap, the use of very limited genetic data in the above listed wheat studies severely decreases their reliability. Thus, although the inferences made by Allaby *et al*. (1999) and Blatter *et al*. (2002) regarding polyphyletic origin of domestic wheat are interesting, it might be sounder to view these with caution.

A second point to be made is that because the *Triticum-Aegilops* differentiation is not very concrete (Morrison, 2001), it should probably be more reasonable to try to estimate divergence times for genomic lineages as Allaby *et al*. (1999) have done, rather than give estimates for the divergence between the two genera.

1.3.2. Problems encountered in genetic analysis of contemporary species

Phylogenetic analysis and other investigations mentioned above are limited by the fact that the relationships among extant species and populations cannot directly represent past, so that they are constrained to make estimations only. One can imagine many factors that could render this estimation erroneous: i) The changing distribution of wild species, ii) increase and decrease in the variation of domestic species, iii) homoplasies between paraphyletic species, iv) contradictions between gene phylograms and species phylograms. This last point should also draw attention to the fact that, the use of limited amount of genetic information, *i.e.* small number of loci studied, usually distorts the picture.

Moreover, it has also been claimed that genetic analysis using methods such as AFLP may not be accurate in determining the exact progenitor species due to the dynamics of post-domestication evolution (Allaby and Brown, 2003).

As the last examples given in the above section implies, evidence from different sources may also seem to conflict with each other. This is due to the fact that the actual processes of evolution and domestication might have been much more complicated than the preferred shortcut, simple and general inferences.

1.3.3. Analysis of archaeological evidence

Plant and material culture remains found in archaeological sites provide an important tool to track domestication. The latter findings can be tools related to crop harvesting, but as mentioned, they are quite scarce. Some experiments have been made to understand the precise usage of these tools –such as whether they were used upon tilled soil, *i.e.* whether were they tools of farmers- making use of the wear effect, but the results are controversial (Hillman and Davies, 1990).

Especially following the use of the technique "floatation" in the 1960's, archaeobotany has developed in an impressing manner (Nesbitt and Samuel, 1996). Most plant remains are found in charred state, but preservation *via* desiccation, mineralization, carbonization, water-logging (Jones et. al., 1996). Nearly all archaeological remains found in Turkey are in charred form (M. Nesbitt, personal communication).

Charring is a very effective way of preservation of plant remains. For wheat species, it must have happened as a result of grains or chaff being unintentionally roasted near ovens, burnt during general fires, or being deliberately burnt for waste disposal, or (following the establishment of animal herding) as part of dung (Nesbitt, 1993).

1.3.4. Problems encountered in morphological analysis of archaeological remains

Morphological analysis of charred wheat has considerably developed since the last two or three decades, and many identification keys have been published (*e.g.* Hubbard, 1992). Despite this progress, charred wheat remains can easily elude identification, especially when only grains are found, which is frequently the case as can be seen from Table 1.3. Also, identifying a portion of archaeological collections also may confuse results (Hillman, 2001).

One of the problems encountered is a consequence of the charring effect, which distorts the shape of the seed, usually causing it to lose water and swell (Nesbitt, 2001).

Moreover, the seed shape is a function of glume shape, which is similar among species, and the shape and size are also affected by the number of grains in a spikelet, which may change.

The following three points summarize the ability or inability of archaeobotanists to identify charred wheat remains.

i) Identification at the ploidy level: The glume base and rachis node morphology is the main source of evidence, while grains are much less informative. The hulled vs. naked character can be determined using the rachis segments and creases on the grains, but identification of the ploidy level is controversial. It can be noted that early identification of naked wheat were all erroneously classified as hexaploids. (Nesbitt, 2001) Thus, many free-threshing wheat remains are identified as *T. aestivum* / *durum*.

ii) Identification at the species level: Leaving aside progenitor-domestic pairs, this is quite a hard task. Intraspecific variation can be a confusing factor for species determination. For example, discriminating between the wild species *T. boeoticum* and *T. urartu*, or the domestic species *T. dicoccum* from *T. timopheevii* is not always possible. The high variation in the archaeological wheat grain record also prompts many investigators to claim the identification of new species, albeit being viewed with suspicion. (Hillman, 2001)

iii) Identification of domestication: This is probably the most desired but simultaneously the most difficult task. Even when the rachis remains are present, distinguishing between the progenitor-domestic pair is usually not possible (van Zeist, 1992). The reason is that here, the main criterion is the smooth vs. ragged-torn abscission scar on the spike, assumed to indicate the way of spikelet disarticulation from a brittle vs. tough rachis; but the harvesting of wild heads while unripe can also produce a ragged scar (Hillman, 2001). Another reason is that the domestics are very similar to ancestor and there is little variation among ancestors (Nesbitt, 2001). In addition, criteria such as seed size – which increased after domestication- selection may not be applicable to early sites. Of course, the charring effect further hinders the use of this criterion.

Even if accurate identification is achieved, the problem of dating may render the investigation controversial. Rather than dating the plant sample, dating is done by using the layer data, which can lead to errors. As a result, the rare reports of domesticated wheat

findings in Pre-Pottery Neolithic A (8500-7500 BC, uncalibrated) are approached with suspicion (Nesbitt, 2001).

Morphological analysis of charred remains today still remains the main tool of archaeobotany, although other chemical criteria are being more and more abundantly used. Thin–layer chromatography, pyrolysis mass spectroscopy, gas chromatography mass spectroscopy and infrared spectroscopy have been used during the last two decades, and the latter method is claimed to be particularly useful (Hillman and Davies, 1990). Meanwhile, ancient DNA study, although risky, has become quite a tempting area.

1.3.5. Ancient DNA analysis

Ancient DNA (aDNA) study actually goes back to the beginning of 1980's (Higuchi *et al*., 1984), but the real start was given by the introduction of the Polymerase Chain Reaction (PCR), allowing small amounts of target DNA to be readily multiplicated, which could then be sequenced (Pääbo, 1989). Random primer amplification or hybridisation is another option, although not as attractive as PCR.

DNA is a quite unstable molecule, readily decaying *via* hydrolysis and oxidation. Still, it is considered that DNA, under favourable conditions, can survive 10,000-100,000 years, especially under cold conditions (Höss *et al*., 1996). There have also been studies claiming to obtain DNA from earlier periods although rare and usually disclaimed (Cooper and Wayne, 1998). In any case, the DNA delivered from ancient tissue has diminished to a very small amount, and has additionally decomposed into fragments usually ranging from 100-500 base pairs. This fragmented form, together with modifications of the structure usually poses problems during PCR, such as jumping between similar templates and yielding chimeric molecules or other kinds of fake sequences (Pääbo *et al.*, 1990; *e.g.* Allaby *et al.*, 1999). In addition to PCR based problems, post-mortem modifications may also be causes of error for phylogenetic analysis (Graur and Li, 2000).

1.3.5.1 Authenticity criteria

Ancient DNA may sound as a very attractive technical subject to molecular geneticists, but its real use is in its contribution to phylogenetic studies. But here, the precious information hidden in aDNA is jeopardized by the risk of contamination from unrelated DNA sequences, modern, or ancient but irrelevant. PCR itself provides an important opportunity for contaminating DNA to appear as ancient and specific. Not surprisingly, many early claims of aDNA extraction have been later disqualified (Cooper and Poinar, 2000). Thus, criteria are obligatory to evaluate the authenticity of aDNA.

Perhaps the most direct evidence regarding the authenticity of aDNA sequences is a possible ancestral state of the sequence obtained (*e.g.* mammoth sequences obtained by Willerslev *et al*., 2003). Still this needs to be regarded with suspicion, as Cooper and Wayne (1998) point out that intense effort to obtain PCR amplification from void samples may eventually yield short chimeras that can be perceived as original. Moreover, in many cases the aDNA may not be "ancient enough" for significant sequence divergence. Thus, other criteria are necessary. Cooper and Poinar (2000) have published a set of criteria which have thereafter been adopted:

i) "Physically isolated work area: To avoid contamination, it is essential that, prior to the amplification stage, all ancient DNA research is carried out in a dedicated, isolated environment. [*e.g.* Yoder and Delefosse, 2002]

ii) "Control amplifications**:** Multiple extraction and PCR controls must be performed to detect sporadic or low-copy number contamination, although "carrier effects" do limit their efficacy. All contaminated results should be reported, and positive controls should generally be avoided, as they provide a contamination risk.

iii) "Appropriate molecular behaviour: PCR amplification strength should be inversely related to product size [*e.g.* Bilgiç, 2002]. Reproducible mitochondrial DNA (mtDNA) results should be obtainable if single-copy nuclear DNA or pathogen DNA is detected. (…) Sequences should make phylogenetic sense [Debruyne *et al.*, 2002].

iv) “Reproducibility: Results should be repeatable from the same, and different, DNA extracts of a specimen. Different, overlapping primer pairs should be used to increase the chance of detecting contamination by a PCR product.

v) “Cloning: Direct PCR sequences must be verified by cloning amplified products to determine the ratio of endogenous to exogenous sequences and damage-induced errors. Overlapping fragments are desirable to confirm that sequence variation is authentic and not the product of errors introduced when PCR amplification starts from a small number of damaged templates. [Bilgiç, 2002]

vi) “Independent replication: Intra-laboratory contamination can only be discounted when separate samples of a specimen are extracted and sequenced in independent laboratories. (…) [*eg*. Rollo *et al.*, 2002; Willerslev *et al*., 2003]

vii) “Biochemical preservation: Indirect evidence for DNA survival in a specimen can be provided by assessing the total amount, composition, and relative extent of diagenetic change in amino acids and other residues. [Amino acid racemization is a frequently used method (*e.g.* Rollo *et al*., 2002)]

viii) “Quantification: The copy number of the DNA target should be assessed using competitive PCR. When the number of starting templates is low (<1,000), it may be impossible to exclude the possibility of sporadic contamination, especially for human DNA studies. [Höss *et al.*, 1996]

ix) “Associated remains: In studies of human remains where contamination is especially problematic, evidence that similar DNA targets survive in associated faunal material is critical supporting evidence. Faunal remains also make good negative controls for human PCR amplifications. [*e.g.* Poinar *et* al., 1998; Rollo *et al*, 2002]”

Following these general rules and trying to abide to as much of them as possible, although perhaps not all, aDNA studies have been continuing and excavating very interesting information. Meanwhile, disclaims regarding published results have not grown much rare (Debruyne *et al*. (2002) and Ovchinnikov and Goodwin (2003) are two recent examples

questioning the authenticity of previously published results on mammoth and human aDNA, respectively).

1.3.5.2. aDNA analysis of charred wheat

Having considered the context and requirements of the study of wheat history, and having noted the obstacles faced by genetic analysis and phylogenetic reconstruction attempts using modern species, and besides the handicaps of archaeobotanical identification, it will be clear that aDNA study of wheat appears as a very appealing issue. There have been studies on charred archaeological samples of other species such as maize (*Zea mays*) (Rollo *et al*., 1991; Goloubinoff *et al*., 1993). The latter study has focused on the post-domestication genetic variation in maize and concluded that there occurred multiple domestications including different wild maize (teosinte) species and/or intense introgression with teosinte.

These studies have been followed by a number of studies conducted on wheat. Allaby *et al*. (1994) and Allaby *et al*. (1999) have claimed to extract ancient DNA from charred 1,000 year old British spelt and 3,000 year old mixed Greek grain, respectively, which they were able to identify. Schlumbaum *et al*. (1998) have amplified DNA from 5,000 year old charred naked grains from Switzerland, and identified them as hexaploid and quite similar to modern bread wheat. The same group (Blatter *et al*., 2002) has worked on 300 year old spelt and 250 year old bread wheat, both desiccated samples from Switzerland. They have found putative historical alleles in spelt. The only study conducted in Near Asia with a closer attention on wheat domestication and early agriculture (*i.e.* the Neolithic period) has been that of Bilgiç (2002). Here, aDNA from Anatolian charred samples (8,000 year old Çatalhöyük and 4,000 year old İmamoğlu Höyük) has been extracted and sequenced. An interesting result has been to find spelt-like motifs in the Çatalhöyük sequence, and the re-identification of seeds -previously assigned as einkorn and emmer- as hexaploid.

All these studies have utilized a single locus, *i.e.* a part of the promoter region of the high-molecular weight (HMW) glutenin protein; a protein which bestows bread wheat its elastic quality for bread-making. The locus has been used also in other genetic relationship studies, and shown to be suitable for species identification (Waines, and Payne, 1987;

Fernandez-Calvin and Orellana, 1990; Allaby *et al*., 1999). This locus has undergone duplication prior to the divergence of the A, B and D genomes, and thus is found in 6 copies per hexaploid individual, which has been referred as an advantage for aDNA amplification (Jones *et al*., 1996).

In the mentioned studies, 200 bp or longer targets in this region have been used. One exception has been Bilgiç (2002), where 107 and 156 bp segments could be amplified from the 8000 year old sample, indirectly suggesting the authenticity of the sequence.

1.3.5.3. Problems with charred wheat aDNA

There are a number of the major difficulties in this area; the primary one being the enhancement of DNA degradation upon charring. Threadgold and Brown (2003) have shown that bread wheat grains lose any amplifiable DNA after being subjected to temperatures higher than 250˚C for durations longer than 2 hours. The authors of this study suggest that seeds charred in big stores may have been affected by milder conditions. But this is probably not the case for many samples of charred archaeological wheat, especially if they burnt intentionally or as part of dung. The shiny appearance of grains is indicative of very high temperatures Bilgiç, 2002), but it is not possible to detect the exact amount of heat exposure from charred grain morphology.

A nearly equally weighted bunch of problems are PCR related ones. They have the same main source: charring. The charring effect probably intensifies those problems related with age. The simplest one is increased *Taq* Polymerase errors as mentioned above, but the more complicated case involves PCR inhibition by charring products. The reason for Taq polymerase inhibition is the presence of the so-called Maillard products (*eg*. alkylpyrazines, furanones and furaldehydes), which arise as a consequence of cross-linking between reducing sugars and amino-acid groups, both entrapping DNA and inhibiting PCR (Poinar *et al*., 1998). These occur spontaneously with age, but much more readily at high temperatures (http://www.fst.reading.ac.uk/people/aamesjm/maillard.htm). The brown colour of heated organic matter occurs as a result of this reaction, and in aDNA studies, Maillard products can be identified by a brownish tint in the DNA extract and by their blue fluorescence under UV

light when the DNA extract is run through an agarose gel (Threadgold and Brown, 2003). DNA entrapped in the form of Maillard products can be released by the addition of the chemical *N*-phenacylthiazolium bromide (PTB) (Poinar *et al*., 1998).

In addition to the charring effect related problems, the age of the sample is naturally important when the tendency towards spontaneous decay is considered. Moreover, conditions of preservation of the grain or chaff, such as the moisture, pH, temperature of the environment, or microbial or fungal decay, will have a determining effect on the success of an aDNA study (Bilgiç, 2002). Depurination and breaks of the strands due to oxidation and hydrolysis will cause aberrant and/or chimeric molecules during PCR (Pääbo *et al.*, 1990).

Contamination by microorganism DNA is another problem (Höss *et al*, 1996), although it may not be as severe as that in animal bones (due to their porous structure (Jones *et al*., 1996)), and wheat specific primers may overcome this problem.

In agreement with the above mentioned points, the amount of charred sample used appears as a crucial factor. But considering that DNA isolation destroys invaluable archaeological material, the decision on the amount of material is certainly not easy.

In the face of all these problems, purification methods that have been developed to deal with contaminants such as those mentioned above (Poinar *et al*., 1998) or that increase DNA extraction efficiency (Höss and Pääbo, 1993) have given extra impetus to aDNA studies, and also to wheat aDNA.

Still, the limited number (including only three groups) of wheat aDNA work conducted hitherto -in contrast to the appeal of this area of study- seems to be, rather than anything else, the reflection of the above mentioned difficulties. The initiator group (R. Allaby and T. Brown, UMIST, Manchester) has not published any charred wheat DNA analysis results since the last five years.

1.4. Objectives of this study

This thesis is aimed to contribute to the wheat aDNA studies conducted until now. One aim has been to isolate and amplify aDNA from new archaeological wheat samples that could be informative in the investigation of wheat history in Asia Minor.

A second kind of study has forced itself upon the difficulties faced with aDNA: Artificial charring of wheat and identification of its effects on PCR.

The third attempt has been to find new loci to be used as genetic markers for wheat aDNA studies. This necessity has multiple reasons behind it. First of all, they stem from my view that the HMW glutenin locus has some handicaps:

i) Being a promoter region limits the accumulated polymorphism amount, and this region may not provide sufficient phylogenetic resolution.

ii) Different amounts of polymorphism occurs in different genomes and different HMW glutenin loci (Allaby *et al*., 1999), probably as a reflection of differing variance among wheat genomes, but possibly also related with the silencing of some of the alleles following polyploidization,

iii) Being related to a gene which most probably has been selected during the evolution of bread wheat, this locus may not be very appropriate to base a comprehensive wheat phylogeny on, baring in mind possible discrepancies between gene trees and species trees (Nei and Kumar, 2000).

iv) If enough DNA is not present in the sample or sufficient number of clones cannot be sequenced, only a limited number of alleles may be retrieved, as in Allaby *et al.* (1999). This presents a limitation for phylogenetic analysis. Using a single copy locus may overcome this disadvantage.

Thus, two approaches have been adopted:

i) To develop a marker directly linked to the brittle rachis character and thus implying domestication status, which could distinguish between morphologically indistinguishable progenitor-domestic remains.

ii) To develop a plasmon (chloroplast or mitochondrial) DNA based marker, which can be used more easily due its high copy number and a higher chance of survival, and secondly which is haploid and can be a good independent support for nuclear DNA based phylogenetic analyses, assuming limited plasmon variation per individual.

One potential problem here is the relatively low rates of substitution of plant mtDNA and chDNA genes, as shown by Wolfe *et al* (1987). Using non-coding DNA for this analysis might compensate for this problem. But it has also been understood that there exists a high rate of structural variation in these genomes (Grauer and Li, 2000). Structural variation, specifically in the form of insertions and deletions in the wheat chloroplast (Ogihara and Ohsawa, 2002), shows a generally irreversible character, and thus can be relatively useful for determining close relationships. But it cannot be used in divergence time detection, because for this purpose, nucleotide substitution rates for this region must be determined specifically.

A second disadvantage of using plasmon loci as markers is that, we are limited to the investigation of only the maternal lineage. Thus, paternal contributions during hybridizations will not reflect themselves in the plasmon genotype, and introgression can not be traced at all. But this potential limitation will not be as severe as in other allogamous species, such as humans (Lewin, 1999).

In any case, use of plasmon loci will eventually help a detailed reconstruction of wheat phylogeny.

CHAPTER II

MATERIALS AND METHODS

All PCR reagents used were products of MBI Fermentas Inc. Lithuania, Roche Ltd.-Germany, Merck & Co. Inc. USA, Sigma USA and Promega USA if otherwise not stated.

2.1. General procedures

2.1.1. PCR amplification

Techne Genius Thermal Cycler, UK, was used as thermocycler. PCR amplifications were carried out in 50 µL for non-radiolabelled (cold) PCRs or in 10 µL for radiolabelled (hot) PCRs, for which agarose gel electrophoresis and polyacryalmide gel electrophoresis (PAGE) were applied, respectively.

Cold PCR mixes contained 1-5 µL of DNA extract, 50 pmol of each forward and reverse primer, 10 nmol of each dNTP, 1.5 mM $MgCl_2$, 1X PCR Buffer (10 mM Tris-HCl, pH 8.3, 50 mM KCl), 1-1.5 unit *Taq* DNA polymerase (MBI or FastStart Taq (Roche)) and the volume was adjusted to 50 µL by PCR water (sterile, double distilled).

Hot (radiolabelled) PCR mixes contained the similar amounts as above except 10 pmol of each primer and 2 nmol of each dNTP species, with the addition of 0.025-0.1 µL of $\propto^{32}$P-dATP or $\propto^{32}$P-dCTP (3000 mCi/mmol), and the volume adjusted to 10 µL. Shields

were used and other appropriate measures were taken to prevent exposure and avoid contamination

PCR cycling conditions were applied as follows: Initial denaturation (94°C for 2 min), 30-40 cycles of denaturation, annealing, synthesis (94°C for 1 min., 50°C (or the appropriate melting temperature of the specific primer pair) for 1.5 min., 72°C for 2 min.), and final extension (60°C for 2 min, 72°C for 15 min). Negative controls (lacking DNA) were carried out to be able to detect any contamination during the PCR preparation. Positive controls were occasionally included in the experiments, but a different thermocycler was used for these. PCR products were visualized either on agarose or polyacrylamide gels.

For degenerate primers, a slightly lower temperature than the Tm of the specific primer pair was preferred to overcome possible failures due to sequence discrepancies. The same approach was adopted for aDNA amplifications, considering possible DNA mutations due to old age, such as depurinations and hydrolysis (Pääbo *et al*., 1990).

Secondary PCRs were also carried out by using previous PCR products as DNA source. The same conditions were applied, but alternatively, PCR cycle numbers were decreased from 40 to 30 to avoid the production of unspecific bands (due to build up of primer-dimers).

2.1.2. Visualization of the amplification products on agarose gels

PCR products, after adding 6X loading dye per sample, were applied on 1 or 2% agarose gels containing EtBr. λ/*Pst*I, and sometimes λ/*Hind*III was used as molecular weight marker; the former is more suited for low molecular weight band detection. Electrophoresis was applied in 1X TBE (Tris, Boric Acid, EDTA) Buffer at 100 V D.C. for 1-1.5 h. PCR products were visualized under UV light.

2.1.3. Preparation of the molecular weight marker for agarose gels

A size marker suitable for the expected sizes in ancient DNA amplifications was prepared by incubating λ bacteriophage DNA with *Pst*I restriction endonuclease. The reaction was prepared by mixing the followings and incubating at 37ºC for 3 hours: 25 µL λ DNA (0.3 µg/µL, MBI Fermentas), 0.5 µL *Pst*I enzyme (20 units/µL, New England Biolabs), 5 µL NEB Buffer3 (10X, NEB), 0.5 µL BSA (100X , NEB), 19.5 µL dd H_2O.

Resultant fragment sizes of the λ /*Pst*I digestion are as follows (in bp):

11,497	5,077	4,749	4,507	2,838	2,560	2,459	2,443	2,140	1,986
1,700	1,159	1,093	805	514	468	448	339	264	247
216	211	200	164	150	94	87	72	15	

2.1.4. Visualization of the amplification products on polyacrylamide gels (PAGE)

PCR reactions were terminated by adding 4 µL stop solution (95% formamide, 0.25% bromophenol blue, 0.025% xylene cyanol, 20 mM EDTA) per 10 µL of reaction volume. PCR products, labelled with $\propto^{32}$P-dATP or $\propto^{32}$P-dCTP, were separated on 6% denaturing polyacrylamide gels (7 M urea, 6% acrylamide, 19/1 (w/w) N, N' Methylene-bis-acrylamide, 1X TBE (90 mM Tris base, 90 mM Boric Acid, 2 mM EDTA, pH 8.0)).

The denaturing polyacrylamide gel solution (60 ml) was polymerized by adding 650 µL 10% APS (ammonium Persulfate) and 25 ml TEMED (N,N,N',N'-Tetramethylethylene diamine) just before casting the gel in the "Sigma Coat" treated glass plates (42X33 cm). The gel was heated under high voltage before loading. Initial denaturation of the PCR products at 95°C for 3 min was applied prior to loading. PCR products (4-6 µL per lane) were loaded alongside with the negative PCR control and molecular weight standards (A and G sequencing products of M13mp18 DNA).

In some cases, the PCR products were loaded by leaving an empty lane in between so that a precise recovery of the PCR product from the gel to be possible. But this practice was abandoned in later experiments while it caused uneven running of the PCR products.

Electrophoresis was applied to the gel at 60 Watt constant power (Biometra High Voltage Power Supply, P6ack P30, Germany) for about 1 hour before the lower dye marker ran off.

Following the termination of the electrophoresis, the two glass plates were cooled down at room temperature, carefully separated, and the gel transferred onto Whatman filter paper (3MM 46X57 cm chromatography paper). The gel was covering with stretch film and dried on a slab gel dryer (SLAB-Gel Dryer SGD 2000) at 76°C for about 1 hour under constant vacuum. Dried gels were exposed to X-ray films (Agfa CP-BU). Exposure time (1-7 days) was determined by the intensity of the radioactivity detected by the Geiger Counter. The X-ray films were developed at the METU Medical Centre.

2.1.5. Cloning of PCR Products

The bands corresponding to expected molecular weights were cut using a sterile lancet from the agarose (minimizing exposure of the bands to UV as much as possible) or polyacrylamide gel. The DNA containing solution was extracted from the agarose gel by sequential freezing and thawing in liquid nitrogen. Alternatively, agarose gel products were melted at 72˚C with the addition of 0.5 unit Taq and 10 nmol dATP, to increase ligation efficiency into the vector. 10 µL of the extracted DNA was directly ligated into pGEM T-Easy vector (Promega) using a 1/3-1/5 diluted vector and following the manufacturer's instructions, with overnight incubation at 4˚C and a subsequent 2-3 hour incubation at 37˚C. *E. coli* DH5α competent cells were prepared and stored in $CaCl_2$ according to Sambrook *et al*. (1989). 40 µL cells was added onto the ligation product, incubated in ice for 30 min, transformed with a 75 second heat shock at 42˚C, stored for another 2 minutes in ice, and grown for 1 hour after addition of 150 µL SOC medium. The cells were then spread onto two LB-ampicillin, XGal, IPTG plates. The colonies were allowed to grow at 37^{o}C for 16-24 hours. Recombinant clones were subjected to blue-white selection, on the basis of inactivation of the lacZ' gene. Selected colonies were PCR checked for presence or absence of insert with M13 or the original primers, and simultaneously grown in LB-ampicillin.The positive colonies were stored as stock at -80 ^{o}C with 1:1 volume addition of glycerol.

2.1.6. Sequencing of the plasmids

The Qiagen Miniprep plasmid isolation kit was used for plasmid purification. The inserts were rechecked with PCR, the DNA concentrations determined. The plasmids were sent for custom sequencing, using forward T7 primers, at the Keck facility of Yale University.

2.2. Ancient DNA isolation and analyses

2.2.1. Plant materials

In this study, archaeological wheat samples from Turkish archaeological sites ranging from the Early Bronze Age to the Ottoman periods (~2300 BC to ~1800 AC) were analyzed. Table 2.1 lists the archaeological samples, the locations of which are shown in Figure 2.1.

Dating of the samples was based on archaeological evidence and is approximate. All are charred carbonized seeds recovered at site either by flotation or by simple collection together with soil. Numbers next to sample names denote the date of excavation that yielded the particular wheat sample. Patnos wheat was a museum specimen for which exact recovery date could not be determined, others were provided either by the excavation heads or archaeobotanists. Identification of the samples have been performed by the following archaeobotanists:

Emel Oybak Dönmez, Ph.D. (Department of Biology, Hacettepe University, Turkey): Yenibademli Höyüğü, İmamoğlu Höyük, Patnos.

Andrew Fairbairn, Ph.D. (School of Archaeology and Anthropology, Australian National University, Australia): Kaman.

Table 2.1. Archaeological wheat samples from Anatolian sites and the number of DNA extractions conducted in this study.

Sample	Location	Species Name	Dating	Period	No. of extractions
Yenibademli Höyüğü 97	Gökçeada	*T. dicoccum*	~2900-2700 BC	Early Bronze Age II	1
İmamoğlu Höyük 86	Malatya	*T. durum / aestivum*	~2300-2000 BC	Early Bronze Age	1
Patnos 61?	Ağrı	*T. durum / aestivum*	~800-700 BC	Urartu	1
Kaman 96_1	Kırşehir	*T. monococcum*	~2300-2000 BC	Early Bronze Age	4
Kaman 96_2	Kırşehir	*T. durum / aestivum*	~2300-2000 BC	Early Bronze Age	4
Kaman 96_3	Kırşehir	*T. durum / aestivum*	~1500-1800 AC	Ottoman	4

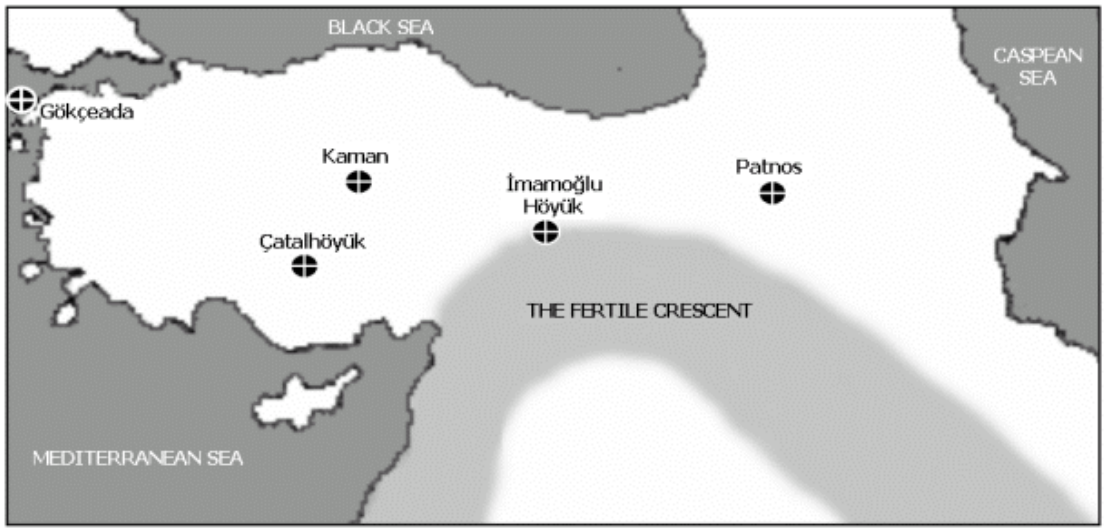

Figure 2.1. The locations of the archaeological sites from which charred wheat samples used in this study were obtained.

2.2.2. Extraction of aDNA

For the ancient wheat seeds, several methods available in the literature for the ancient DNA isolation or newly designed in this study were applied.

During all extractions, one or two extraction blanks were used as control against *in vitro* contamination. These consisted only of the applied chemicals and were processed in parallel to the other samples, with care given to prevent any cross-contamination.

2.2.2.1. CTAB Method

This is a modification by Allaby *et al.* (1999) to a previously developed plant DNA extraction method utilising CTAB (Cethyltrimethylammonium bromide). It is reported to be most promising for DNA isolation from archaeological wheat samples by Bilgiç (2002).

Extraction:

0.5 g of charred seeds were crushed to powder inside UV sterilized aluminium foil, and transferred to two sterile microfuge tubes. 750 µL preheated Buffer 1 (2% w/v CTAB (Cethyltrimethylammonium bromide), 100 mM EDTA pH 8.0, 20 mM Tris-Cl pH 8.0, 1.4 M NaCl) was added to each tube. Two extraction blanks consisting of only Buffer 1 were assembled.

The tube caps were sealed with parafilm, and vortexed vigorously. The mixture was kept at 60ºC in a water bath for at least 4 hours (this is 1 hour in the original procedure), with occasional shaking. Following the incubation, the tubes were centrifuged for 30 min at 15000 rpm at room temperature to spin down the seed debris. About 500 µL of the supernatant was retrieved, the rest being trapped among the debris particles. In cases which even this amount was not reached, an extra 300 µL Buffer 1 was added to the tubes, the centrifugation step repeated, and 250-300 µL of supernatant added to the previous collection. 500 µL (1:1 volume) of chloroform:isoamyl alcohol (24:1) was added to the tubes, and the mixture vortexed. The aqueous and organic phases were separated by 10 min centrifugation

at 15000 rpm. The upper aqueous layer containing DNA was carefully removed. Two volumes of Buffer 2 (1% w/v CTAB (Cethyl-trimethylammonium bromide), 50 mM EDTA pH 8.0, 10 mM Tris-Cl pH 8.0) was added to the aqueous supernatant and the tubes incubated at 4ºC overnight or longer for the precipitation of DNA. The extract was centrifuged for 30 min at 15 000 rpm to precipitate DNA. The supernatant was discarded and the pellet resuspended in 50 µL sterile dd H_2O.

Ethanol (EtOH) precipitation

0.2 volumes of 5 M NaCl and 4 volumes of 100% ethanol were added to the resuspended pellet. The mixture was incubated at –20ºC for 12-24 h. The DNA was collected by centrifugation at 15 000 rpm for 30 min. The ethanol was poured off and the pellet was allowed to air-dry before resuspension in 50 µL sterile dd H_2O. The DNA extract was stored at –20ºC.

2.2.2.2. CTAB-2 method

A DNA isolation procedure created as a hybrid between CTAB and 2XCTAB (Saghai-Maroof *et al*., 1984) and named as CTAB-2 in this study was tested on ancient seeds. The differences from the above described CTAB method are as follows: i) 0.2 volumes of 5 M NaCl and 4 volumes of 100% ethanol were added rather than Buffer 2 as precipitant; ii) following storage at –20ºC, centrifugation at 4ºC, and discarding of the supernatant, a washing step was included, in which 75 - 90% EtOH (higher alcohol concentration may be preferred to avoid DNA loss) was added. A second 15000 rpm centrifugation at 4ºC was conducted, the pellets air dried, and eluted in 50 µL sterile dd H_2O. The DNA extract was stored at –20ºC.

In a further modification attempt, the chloroform:isoamyl alcohol extraction was done twice, as suggested by Threadgold and Brown (2003).

2.2.2.3. Comparison of the CTAB and CTAB-2 methods

To assess the efficiency of each in terms of DNA extraction, a trial with 0.5 g of modern seeds was conducted. The pellets formed during the initial precipitation step were compared by eye.

2.2.2.4. Single seed extraction

This method is reported by Bilgiç to have been used successfully at least once. In this method, carbonized seeds were crushed by sterile 1 ml pipette tip in 50 µL of sterile double distilled (dd) water and the extract used directly for PCR. The aim was to eliminate the possibility of loosing any trace amount of DNA during laborious extraction steps.

2.2.2.5. Column purification of the DNA extracts

A purification attempt was conducted on the CTAB-2 extractions from the three Kaman grain samples, using the Qiagen PCR purification kit, which removes salts and nucleotides lower than 100 bp.

Another purification on some of the DNA samples was performed using Centri-Spin10, Centri-Spin20 and Centri-Spin40 column purification kits (Princeton Separations, Inc.), according to the manufacturer's instructions. These kits purify short sequences from contaminants such as other nucleotides, shorter (*e.g.* > 30 bp) sequences, salts and polysaccharides.

2.2.2.6. Purification using a biotinylated primer

A 5' biotinylated Glu_156A/B reverse primer was used as bait. Yenibademli, İmamoğlu and Kaman (EBA, *T. monococcum*) CTAB-2 extracted samples, the corresponding blank, and a positive control containing 100 ng DNA were used.

The procedure was modified from Hakkı (2002). Promega streptavidin coated magnetic beads were pre-washed twice in 10X PCR Buffer w/MgCl using a magnetic stand. 30 µL of ancient seed extract was heated at 94˚C for 5 min, then cooled down to room temperature for 1 hour 15 minutes. The extract was added to 20 µL of washed beads and mixed. After waiting for 10 min, the suspension was placed on the magnet for 15 min. The supernatant was removed using a micropipette, taking care not to disturb the beads. The beads washed thrice in 10X PCR Buffer w/MgCl, with a 5 min resting step at room temperature after each addition of the buffer. Finally, the beads were suspended in 20 µL Buffer.

Cold PCR was conducted using 3 µL of beads, taking care of the reaction stochiometry, due to the transfer of 10X PCR Buffer inside the bead suspension.

2.2.3. Authenticity criteria

The following authenticity criteria were applied throughout the study:

i) PCR (negative) controls and extraction controls (blanks) were applied for all the extractions and PCRs.

ii) Extractions and PCR setup were performed at a dedicated laboratory and in a sterile hood, were no other biological work was done previously or at the time of ancient DNA experiments.

iii) Dedicated equipment and sterile disposable plasticware was used during DNA extraction.

iv) PCR reactions were carried out at a dedicated thermocycler block in which no any other reactions were performed.

v) All equipment, solutions and the sterile hood were cleaned first with double distilled water, then bleach or ethanol, and subsequently either autoclaved or UV sterilized.

vi) Different length PCR primers of the target region were used.

2.2.4. Çatalhöyük DNA

Çatalhöyük "einkorn" and "emmer" extracts of Hatice Bilgiç, previously shown to give PCR amplification, and the corresponding extraction blank, were used in this study. These were of interest because they had exhibited hexaploid sequence motifs and were later re-identified as bread wheat based on morphological criteria (Bilgiç, 2002). Hence using the same names used in that study has been preferred, but in quotation marks.

2.2.5. Primers

Table 2.2 summarizes the primer pairs used in this part of the study. All primers except Ta16S, Chl_HtSpt and Cox1_1 are adopted from Bilgiç (2002).

2.3. The artificial charring experiment

The aim was to try to follow the effects of the charring process on the morphology of the seeds, to assess the efficiency of different DNA extraction methods and to observe closely the PCR inhibition effect of extracts from charred wheat.

Dağdaş-94 bread wheat was used as plant material. The seeds were enclosed in aluminium foil, and the charring was accomplished in an etuve. The different samples and treatments used are listed in table 2.3.

Table 2.2. Target regions and sequences of the PCR primers utilized in ancient DNA amplifications in this study. Degenerate letters in sequences stand for: Y:C/T; R:A/G; K:G/T; M:A/C, W: A/T, S: C/G.

PCR Primers	Target region	Size (bp)	Sequence (5'→3')
Glu_243	Nuclear-HMW glutenin promoter	243	F: GATTACGTGGCTTTAGCAGAC R: TGCTCGGTGTTGTGGGTGAT
Glu_156_D	Nuclear-HMW glutenin promoter	156	F: CAAAGCTCCAATTGCTCCT R: TTTATAGGGACGTGGTGAAG
Glu_156_A/B	Nuclear-HMW glutenin promoter	156	F: CAAAGC ACCAATTGCTCCT R: TTTATAGGGACGAGGTGAAG
Glu_107	Nuclear-HMW glutenin promoter	107	F: GCTTYTTTTGTGTTGGCAAAYT R: GTTCRKGACMATGGYTGYGT
TrnL	Chloroplast intergenic spacer between tRNA-L 3' exon and tRNA- F	385-418	F: CAAGTCCCTCTATCCCCA R: AACTGAGCTATCCTGACC
TrnL-int	Chloroplast intergenic spacer between tRNA-L 3' exon and tRNA- F	120	F: GCAATGGGTTTAAGATTCA R: CGCCGATACTCTAATAAAA
Chl_HtSpt	Chloroplast intergenic spacer in the "hot spot" region	243	F: GACTTCTTTATTTCTTTATGTGCAA R: ACATTGACTTTCTATTCAAAAA
Ta16S	Chloroplast 16S RNA sequence incorporating an SNP	111	F: GTGTTGGGTTAAGTCTCGCA R: GCCTCATCCTCTCCTTCCTC
Cox1_1	Mitochondrial sequence flanking the 5' region of the cox1 gene	114	F: AGTTCTTCTCTTCCAGCCCC R: GWRMAAATAGAAAGYAGTAAASC

Table 2.3. The samples and DNA extracts used in the artificial charring experiment. N=not weighted

Treatment (°C)	Duration	Grams	No. of seeds	Extraction method
270	11 h	N	10	CTAB, CTAB-2
270	11 h	N	1	Direct single seed
270	6 h	0.107	5	QIAGen DNAEasy
270	1 h	0.101	3	QIAGen DNAEasy
200	20 h	0.101	3	QIAGen DNAEasy
200	15'	0.106	3	QIAGen DNAEasy
None	0'	0.116	3	QIAGen DNAEasy

The direct single seed and CTAB-2 extraction methods were as previously described in section 2.2.2, except that when using CTAB-2, only 100% EtOH but no NaCl solution was added for initial precipitation. The Qiagen DNeasy kit was used following the protocol of the kit.

The extracts, together with positive control DNAs, and using nuclear and chloroplast primers, were used in PCRs under different conditions.

2.4. New marker development

The general aim was to develop short primers useful for species identification in aDNA studies. All primers were designed by and/or checked for Tm compliance and 3'-end complementarities using the online Primer3 program (www-genome.wi.mit.edu/cgi-bin/primer/primer3.cgi/). Primers were subjected to Blastn search to detect any homologous sequences that might confuse results.

2.4.1. Domestication specific marker development

The aim here was to develop a marker able to indicate domestication status in wheat. The genes controlling the brittle rachis character have not been identified and cloned as yet (see Introduction). The primers were designed using the *qSh-1* gene cDNA sequence cloned from *Oryza sativa*, found in GenBank with accession numbers AB071330- AB071333 (corresponding to four different cultivars). The gene is known to control a similar trait in rice, called "shattering".

Homologous sequences were searched in GenBank using TIGR Blastn, and an unidentified *T. aestivum* cDNA sequence (accession # BJ300850) was found. Using the one-to-one homologous regions of the two sequences, two forward primers and one reverse primer were designed. The primers were named as **"**Brittle rachis candidate". Their sequences are as follows:

Br-can-Long Forward: 5’ ctgcgcaacacsagcaa 3’

Br-can-Short Forward: 5’ tggctgttcgaycacttc 3’

Br-can Reverse: 5’ cgttgtargcgaagctcac 3’

“y” denotes a degenerate site with a pyrimidine, and “r” denotes a degenerate site with a purine.

For differential screening with Br_can primers, *Triticum boeoticum*, *T. monococcum*, *T. urartu*, *T. dicoccoides*, *T. dicoccum*, *T. durum*, *Ae. tauschii*, *Ae. cylindrica* and *T. aestivum* sample DNA was used.

Firstly, cold PCR products were screened on agarose gels. Later radiolabelled PCR was done and run with PAGE. A domestication specific band was searched by comparing the banding pattern of domestic species with those of wild ancestral and progenitor species, separately for the einkorn and emmer lineages.

2.4.2. Non-nuclear marker development

Four chloroplast and mitochondrial genome based primers have been developed, the sequences of which are given in Table 2.2 except Cox1-2.

2.4.2.1. Chloroplast primer design

The Ta16S primer was designed targeting a single nucleotide polymorphism containing chloroplast 16S RNA sequence, which was found to distinguish *T. aestivum* from its ancestors *Ae. speltoides, Ae. tauschii*, *T. dicoccum* (Rudnoy *et al*., 2002). The sequence and target region size is given in 2.2.5.

The chloroplast hot spot region for mutations, downstream of the rbcL gene was scanned for a most variable short region using GenBank accession #X62117, X62118,

X62119, and a 243 bp portion between *psaI* and ORF185 was chosen. Primers were designed using homologous regions flanking this sequence:

2.4.2.2. Mitochondrial primer design

Using the only mitochondrial locus present in the GenBank for both wild progenitor and domestic wheat species, a 5' flanking region of the cytochrome c oxidase 1 (*cox1*) gene, two short overlapping primers were designed, covering a highly variable region of total 301 bp. Cox1_1 and Cox1_2 were designed to amplify 114 and 178 bp regions, respectively. The sequences of Cox1_2 are as follows:

Cox1_2_For: 5' cmgstttactrctttctattt

Cox1_2_Rev: 5' grtcytgaatctcyggmgg

2.4.2.3. Testing of TrnL-F intergenic sequence primers

In addition, the primers designed by Bilgiç (2002) but not used in that study were tested. The TrnL primers (in short for the TrnL-F intergenic region amplifying primers) were used to clone the corresponding locus using plant DNA (listed in Table 2.4) prepared previously by Hatice Bilgiç, and can be found in Bilgiç (2002). The samples other than the ones from ICARDA and the Aegean Agricultural Research Institute ones have been prepared by Hatice Bilgiç from unspecified material.

Table 2.4. The species and varieties used in cloning the TrnL-F intergenic sequence. TR: accessions from the Aegean Agricultural Research Institute, Menemen, Turkey. IG: accessions from the International Center for Agricultural Research in the Dry Areas (ICARDA), Aleppo, Syria.

Variety	Species
KCD21	*T. boeticum*
IG116190	*T. urartu*
TR36938	*T. monococcum*
IG45257	*T. monococcum*
30 100 dicoccum	*T. dicoccum*
IG46160	*T. dicoccoides*
Tauschii 940105	*A. tauschii*
Çakmak79	*T. durum*
TR50282	*A. biuncialis*
TR52074	*A. columnaris*
TR51865	*A. cylindirica*

2.4.2.5. Alignment of sequences

Ancient and modern DNA sequences obtained in this study were aligned by the "multiple alignment" option of Clustal X (1.81) program after excising the primer sequences. The "nexus" format files were created for further phylogenetic analysis with PAUP* (Pylogenetic Analysis Using Parsimony) using the corresponding options in Clustal X. The Chromas program was used to detect the precision of the sequence.

2.4.2.6. Phylogenetic analyses

PAUP* version 4.0b10 was used for phylogenetic analysis of the 400 bp TrnF-L intergenic region sequences. The commands were included into the data file according to the User Manual and Command Reference document of the Beta version (Swofford, 1991). Maximum parsimony trees were sought with exhaustive search under the following assumptions, which are discussed in section 3.3.3.1. The gaps in the length mutation region were changed to missing data by changing the character to "?", while the other single

nucleotide gaps were depicted as "-" and assumed as 5th character states. The three length mutations in the region were given binary character states, and included as new characters in the data file.

These three new characters were given a weight of 1 and a weight of 5 in two searches, while all other characters were given equal weight. Similarly, all characters were assigned "unordered" states, while the three new characters were assigned "Dollo" characters. The ancestral state of these three new characters was assigned "0", *i.e.* absent.

The most parsimonuos trees (having the highest consistency indices) were constructed, together with the 100% majority rule consensus tree. The branch lengths for the trees were calculated. A heuristic bootstrap analysis was done with 100 replications. Genetic distances were not calculated.

CHAPTER III

RESULTS AND DISCUSSION

3.1. Ancient DNA isolation, PCR and analysis

3.1.1. The charred grains

All the ancient seeds were very brittle as observed during crushing. They had intense black colour, while Patnos and Ottoman Kaman exhibited the bright black colour indicating exposure to even higher temperatures than the others.

The only exception was the Yenibademli Höyük emmer seed sample; these grains were relatively less black (although not brown) and also contained boreholes indicative of insect attack. Although this property is discussed from a negative standpoint by Bilgiç (2002), a different interpretation is also possible, such that insect attack may suggest the presence of digestible organic molecules that survived charring, among which DNA may also be present.

3.1.2. Comparison of the CTAB and CTAB-2 methods regarding efficiency

The only factor in the CTAB method (2.1.2.1) that is expected to precipitate DNA at the stage of Buffer 2 addition seems to be low temperature, as Buffer 2 is only a dilute version of Buffer 1, the extraction buffer. Temperature alone may not be effective as alcohol,

and this may cause extra DNA loss during purification. This danger was the reason the CTAB-2 protocol has also been developed and used.

To assess the efficiency of each in terms of DNA extraction, a trial with 0.5 g of modern seeds was conducted. The pellet that formed at the first precipitation step by the 2X CTAB method was about ten times larger in size than that formed by CTAB method, as observed by naked eye. This result was not reproduced. Also, because RNAse was not added during the procedure, a substantial amount of the pellet may be comprised of RNA.

Nevertheless, the result suggests a significant loss of DNA by the CTAB method - perhaps leading to complete removal in ancient samples.

3.1.3. Comparison of extraction methods regarding purification

A total number of 3 extractions using the described DNA isolation methods were performed with the Kaman Kalehöyük einkorn (*T. monococcum*) EBA, naked (*T. durum/aestivum*) EBA and naked (*T. durum/aestivum*) Ottoman samples. The İmamoğlu, Patnos and Yenibademli samples were used in only one extraction.

During CTAB extraction, either a faint brown pellet or no precipitate was observed, rather than the usual white DNA pellet. Accordingly, some of the elutes obtained at the final stage carried a brownish colour. This observation was also made by Bilgiç on some of her samples (personal communication).

On the other hand, elutes obtained with the CTAB-2 method generally exhibited a much darker brown colour compared to those from the CTAB extraction. Also in a few samples (İmamoğlu and one Kaman einkorn), a clear elute solution contained a precipitate of a small amount of insoluble and minute brown particles.

A soluble or insoluble brown pigment is not observed in uncharred seed DNA extracts; hence this colour is most probably a reflection of charring-associated Maillard products (see section 1.3.5.3). If this is true, the CTAB method, compared to the unmodified

CTAB-2 method, appears to have a better capacity to purify DNA from products related with charring; although it is less efficient in DNA extraction (see above).

Meanwhile, further modification of the CTAB-2 protocol by including two times chloroform:isoamyl alcohol extraction mostly compensated for the mentioned increase in colour.

It can be noted that, the colour of the Kaman naked Ottoman sample extract was the darkest in all extractions. This peculiar feature may be an outcome of severe burning, which in turn implies deliberate burning. Nesbitt (1993) points out that most of the charred grains from the Ottoman period were obtained as part of burnt dung, while dung burning must have become more common as deforestation increased in this era. This possibility is synonymous to the lack of DNA in the Ottoman sample, albeit it is the most recent one.

The applications of CTAB-2 extracted samples on the Qiagen and Centri-Spin columns as a purification step was able to remove both the soluble coloured pigments and the insoluble particles for most of the samples. In this respect, the most effective columns seemed to be Centri-Spin20 and Centri-Spin40. Still, each column is known to retain some DNA, decreasing efficiency.

It can be noted that, there are other purification methods also used in ancient charred seed DNA studies, such as silica extraction (Höss and Pääbo, 1993), gel purification or electroelution (Bilgiç, 2002), but Bilgiç's results indicate that it is not possible to assign high success to any of these, at least when charred remains are concerned. Moreover, each extra step in the extraction and purification protocol introduces both extra risk of contamination and loss of DNA, and hence is not much preferable.

In summary, the use of the CTAB-2 method with twice chloroform:isoamyl alcohol extraction, followed by column purification may be a simple and efficient approach for ancient DNA extraction from charred wheat seeds.

A last note can be made regarding the duration of initial incubation with Buffer 1. This was 30 minutes in the original procedure (Allaby *et al*, 1994) and Bilgiç (2002) used 1 hour, whereas in this study, 4-8 hours incubations were done. This was prompted by the

observation that DNA yields increase with increasing incubation time for modern seeds. Actually, this may also be a factor increasing the solubilization and amount of charring-related products in the extract.

3.1.4. PCR results and PCR inhibition

Unfortunately, none of these charred seed extracts (CSE), either in crude form or purified, yielded any PCR amplification to be observed by EtBr staining under UV light or *via* radiolabelling. Neither did secondary PCRs yield any result (ignoring the unspecific banding patterns obtained in 2nd PCRs). PCRs with primers targeting chloroplast and mitochondrial sequences have similarly not given any amplification.

Meanwhile the purification attempt with the biotinylated Glu156A/B reverse primer was not successful on both the ancient samples and with the positive control, suggesting that the protocol is not optimized.

Figures 3.1 and 3.2 show two examples of unsuccessful hot PCR and cold PCR amplifications. As can be observed, the strong bands in positive controls indicate that the PCR has worked efficiently. The lack of bands in the blanks and negative controls suggest the absence of contamination for these experiments. Contamination in blanks and negative controls was also observed in some other cases, an example of which is shown in Figure 3.3.

At this point it can be noted that, obviously in this study, only a small number of the authenticity criteria put by Cooper and Poinar have been abided by (see Introduction), but the general –and rational- strategy has been to first obtain amplification and to achieve perfect reproducibility later.

Considering that the use of multiple targets and multiple samples should decrease the role of any stochastic effects, there can be three alternative explanations proposed for the lack of amplification:

i) The charred seeds did not contain any amplifiable DNA.

ii) DNA was lost through the DNA extraction procedures.

iii) A PCR inhibition effect is observed in the PCRs with the charred extracts.

Suspicious of this last possibility, which is not an unusual case (Threadgold and Brown, 2003), certain PCR reactions have been done with a mixture of modern DNA and ancient extracts, called “spiking” (Yoder and Delefosse, 2002). These can also been seen in Figures 3.1 and 3.2. Apparently, this was the case in all the reactions.

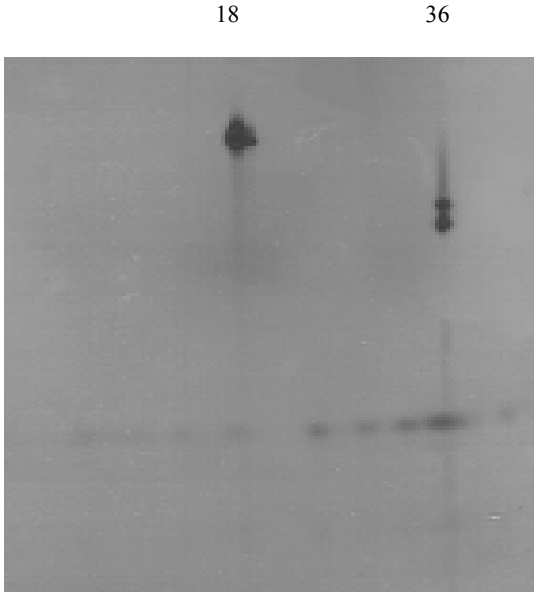

Figure 3.1. An example of a hot PCR on ancient samples. PAGE results with $\propto^{32}$P-dCTP radiolabelled PCRs using CTAB-2 extracted and Centri-Spin column purified Kaman Kalehöyük (KK) EBA einkorn, EBA naked, Ottoman naked, İmamoğlu naked, Patnos naked and Yenibademli emmer extracts, the corresponding blanks, “spiking” samples (containing both charred seed extracts and modern DNA), PCR controls and positive controls. Two sets of amplification have been done with Chl_HtSpt and Cox1_1 primers. ÇH “einkorn” sample and blank were also included in the last lanes of the Cox1_1 set (Lane 39, 40). The positive controls have been loaded at the end of each set and can be observed as the only bands on the gel (Lanes 18, 36). The same amount of modern DNA has not yielded any amplification when mixed with CSE, displaying the PCR inhibition effect (Lanes 16, 17, 34, 35). The appearance of double bands in the Cox1_1 set is most probably due to unspecific banding at low annealing temperature, although such a case was not observed during the primer tests. The appearance of a shadow band following the positive control of the Chl-HtSpt is probably due to leakage, while this lane was left empty.

The amplification with the extra (column) purified extracts were also unable to overcome the PCR inhibiting effect, let alone give amplification. The only effect was the removal of low molecular weight fluorescent compounds, which can be nucleotides or other compounds (see below).

These recurrent observations have lead to the construction of an artificial charring experiment.

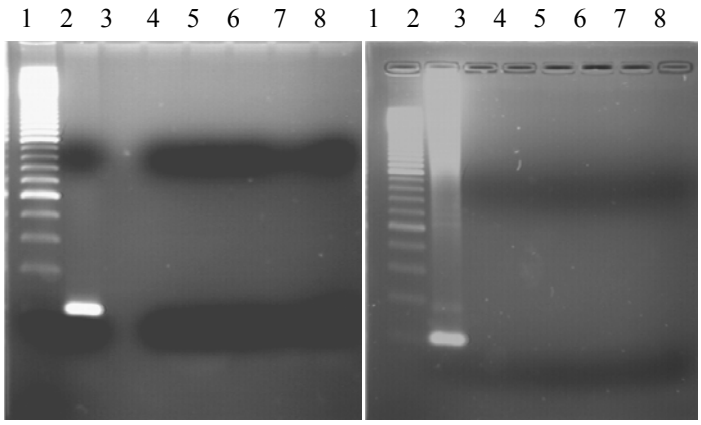

Figure 3.2. An example of a cold PCR and its 2nd PCR. Glu_107 primers used on Kaman Kalehöyük (KK) samples. Run on 2% agarose gels. 1st gel contains 1st PCR, 2nd gel contains 2nd PCR products. Lane 1: 100 bp marker. Lane 2: Positive control. Lanes 4-8 (1st gel) and lanes 3-7 (2nd gel): CTAB extracted two KK EBA einkorn samples, two blanks, negative control, respectively. Lane 8: Negative control of the 2nd PCR. The high molecular weight fluorescence and unspecific bands observed in the positive control in the 2nd gel are characteristic of 2nd PCRs (Bilgiç, 2002).

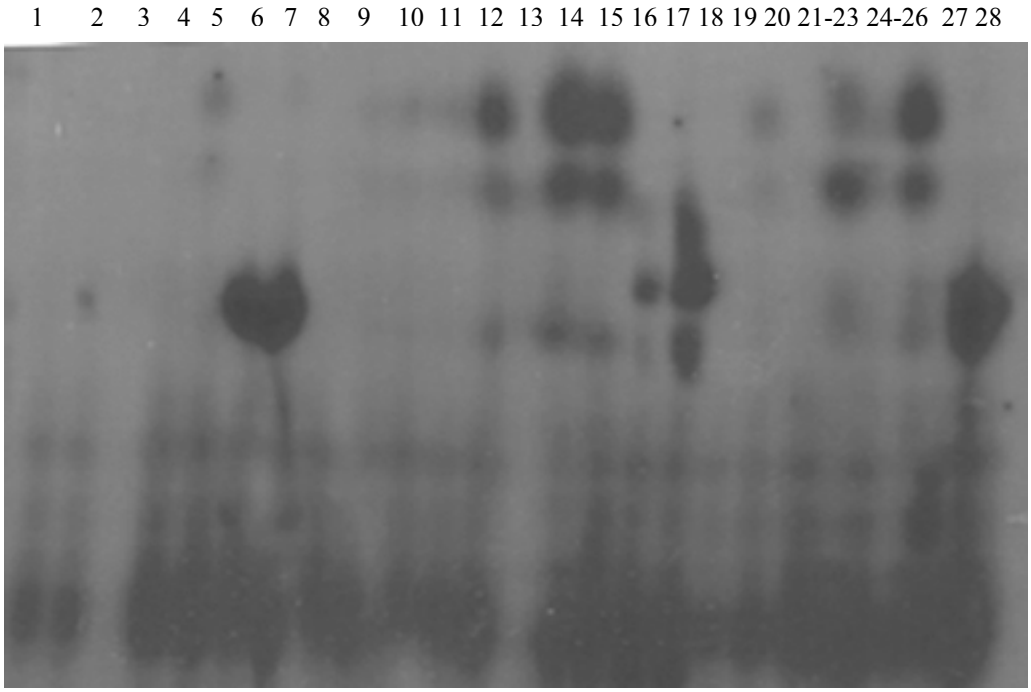

Figure 3.3. An example of contamination in a hot PCR with ancient samples. PAGE results with $\propto^{32}$P-dATP radiolabelled PCRs using CTAB extracted KK EBA einkorn (einkorn-1), CTAB-2 extracted KK EBA einkorn (einkorn-2) and naked wheat, Çatalhöyük (ÇH) "einkorn" and "emmer" DNA samples, the corresponding blanks, PCR controls and positive controls. Glu107 and Ta16S primers were used in this experiment. In respective order: **Lanes 1-5, 21-23, 27:** Glu107 1st PCR - ÇH einkorn, ÇH emmer, KK einkorn-1, KK einkorn-2, KK naked, ÇH blank, KK blank, positive control. **Lanes 6-10:** Glu107 2nd PCR - ÇH einkorn, ÇH emmer, KK einkorn-1, KK einkorn-2, KK naked. **Lanes 11-15, 24-26, 28:** Ta16S 1st PCR - ÇH einkorn, ÇH emmer, KK einkorn-1, KK einkorn-2, KK naked, ÇH blank, KK blank, positive control. **Lanes 16-20:** Glu107 2nd PCR - ÇH einkorn, ÇH emmer, KK einkorn-1, KK einkorn-2, KK naked. The PCR solutions have been loaded with one lane space left between them, which has caused the gel to run unevenly, and also has prevented us from loading the controls of the 2nd PCRs. The bands at the positive controls indicate which bands are of appropriate length, while there are some larger sized unspecific bands in ancient sample lanes. At first look, amplifications with both Ta16S and Glu107 primers using ÇH einkorn and emmer are noticed, but a careful look should reveal that there are also two faint bands at this region in PCR's with ÇH blanks, again for both primers. When the 2nd PCR blanks and controls were loaded to another gel (data not shown), some showed strong amplification clearly indicating contamination. Hence the positive results in the above gel were discarded. It is also interesting that the CTAB extraction of KK einkorn showed a strong PCR inhibition -there was no unspecific amplification in the lanes in which the PCRs with this sample was loaded. This was not the case for the two CTAB-2 extracted KK samples, although the latter showed the PCR inhibition effect in other cases (shown in Figure 3.8).

3.1.5. Çatalhöyük DNA results

In general, the PCRs in which Çatalhöyük "einkorn", "emmer" (see section 2.2.4 for reason for using quotation marks) and the corresponding extraction blank were included did not yield any amplification. In one case, the negative control, and in one case the blank showed contamination, and the results were discarded.

A positive result was obtained with only one "hot" amplification event: Using TrnF-int primers on Çatalhöyük "einkorn". In this gel, the blank, negative control and "ÇH emmer" lanes appeared clean of any bands (Figure 3.4). The band that appeared as the 1st PCR product was of very weak intensity when compared to that in the positive control.

It can be noted that whereas both ÇH samples had given clear amplification in previous trials by Hatice Bilgiç, this time the same "emmer" sample failed to give amplification.

Interestingly, in this gel, there appeared two very similar sized bands, both in the Çatalhöyük sample and the positive control, in which *T. turgidum* DNA was used. The exact size difference -probably about 5 bp- could not be detected due to a problem with the AG molecular size marker. A second PCR was also performed, and the same two bands appeared, with a clean blank and PCR control as before, while the intensity in the Çatalhöyük sample increased. The two bands from ÇH "einkorn" and *T. turgidum* 2nd PCRs were cut from the polyacrylamid gel and cloned. The ÇH "einkorn" bands and the positive control lower (small sized) band were sent to sequencing.

The sequences were aligned with the ClustalX 1.81 program and compared with other species' sequences at this loci. Part of the TrnL-F intergenic region upstream of the 120 bp Int-TrnL target sequence was also included in the alignment. The alignment is given in Figure 3.5.

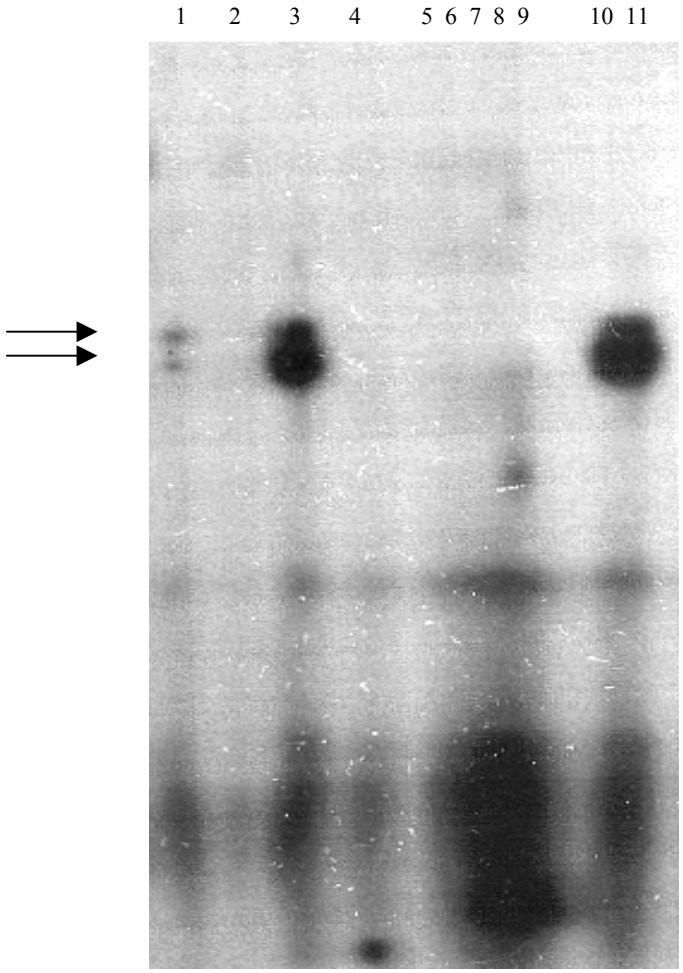

Figure 3.4. A successful hot PCR with Çatalhöyük seed extract. PAGE results with $\propto^{32}$P-dATP radiolabelled PCRs, using Çatalhöyük (ÇH) "einkorn" and "emmer" DNA and TrnL-int primers. Lane 1: ÇH einkorn. Lane 2: ÇH emmer. Lane 3: ÇH einkorn 2nd PCR. Lane 4: ÇH emmer 2nd PCR. Lane 5: ÇH Blank. Lane 6: PCR control. Lane 7: ÇH Blank 2nd PCR. Lane 8: 2nd PCR of 1st PCR's control. Lane 9: PCR control of 2nd PCR. Lane 10: Positive control (*T. turgidum*). Lane 11: Positive control 2nd PCR. Arrows indicate the "upper" and "lower" bands.

Unfortunately, this 120 bp section of the TrnL-TrnF intergenic spacer region essentially did not exhibit any polymorphism between the species. The only exception was the initial residue: In the *T. dicoccoides* line including *T. aestivum*, the nucleotide corresponding to the first nucleotide of the primer sequence was a T instead of the G present in the primer sequence G. The reason for this polymorphism was most probably an

insertion/deletion polymorphism (see section 3.3). But this difference apparently did not alter the PCR performance.

The small-sized "lower bands" from both samples shared the same expected sequence, while the Çatalhöyük "upper band" sequence had a 1 bp deletion, which was actually inside the primer region and probably corresponded to an amplification error (The sequence was read unambiguously at that residue).

```
dicoccum          TATCCTCTTTTTTTT---CTTTTTAT----------------------------CAATG
durum             TATCCTCTTTTTTTT---CTTTTTAT----------------------------CAATG
aestivumGB        TATCCTCTTTTTTTT---CTTTTTAT----------------------------CAATG
dicoccoides       TATCCTCTTTTTTTT---CTTTTTAT----------------------------CAATG
monococcum        TATCCTCTTTTTTTTTTTCTTTTTATCAATGCACTTTTTTTTCTTTTTATCAATGCAATG
boeticum          TATCCTCTTTTTTTTTT-CTTTTTATCAATGCACTTTTTTTTCTTTTTATCAATGCAATG
upperCH           ----------------------------------------------------GCAATG
lowerCH           ----------------------------------------------------GCAATG
lowerTr           ----------------------------------------------------GCAATG

dicoccum          GGTTTAAGATTCATTAGCTTTCTCATTCTACTCTTTCACAAAGGAATGCGAAGAGAACTC
durum             GGTTTAAGATTCATTAGCTTTCTCATTCTACTCTTTCACAAAGGAATGCGAAGAGAACTC
aestivumGB        GGTTTAAGATTCATTAGCTTTCTCATTCTACTCTTTCACAAAGGAATGCGAAGAGAACTC
dicoccoides       GGTTTAAGATTCATTAGCTTTCTCATTCTACTCTTTCACAAAGGAATGCGAAGAGAACTC
monococcum        GGTTTAAGATTCATTAGCTTTCTCATTCTACTCTTTCACAAAGGAATGCGAAGAGAACTC
boeticum          GGTTTAAGATTCATTAGCTTTCTCATTCTACTCTTTCACAAAGGAATGCGAAGAGAACTC
upperCH           GGTTTAAGATTCATTAGCTTTCTCATTCTACTCTTTCACAAAGGAATGCGAAGAGAACTC
lowerCH           GGTTTAAGATTCATTAGCTTTCTCATTCTACTCTTTCACAAAGGAATGCGAAGAGAACTC
lowerTr           GGTTTAAGATTCATTAGCTTTCTCATTCTACTCTTTCACAAAGGAATGCGAAGAGAACTC
                  ************************************************************

dicoccum          AATGGATCTTATCCTATTCATTGAATAGATTTCTTTTTTATTAGAGTATCGGCAAGAAAT
durum             AATGGATCTTATCCTATTCATTGAATAGATTTCTTTTTTATTAGAGTATCGGCAAGAAAT
aestivumGB        AATGGATCTTATCCTATTCATTGAATAGATTTCTTTTTTATTAGAGTATCGGCAAGAAAT
dicoccoides       AATGGATCTTATCCTATTCATTGAATAGATTTCTTTTTTATTAGAGTATCGGCAAGAAAT
monococcum        AATGGATCTTATCCTATTCATTGAATAGATTTCTTTTTTATTAGAGTATCGGCAAGAAAT
boeticum          AATGGATCTTATCCTATTCATTGAATAGATTTCTTTTTTATTAGAGTATCGGCAAGAAAT
upperCH           AATGGATCTTATCCTATTCATTGAATAGATTTCTTTTTTATTAG-GTATCGGCGA-----
lowerCH           AATGGATCTTATCCTATTCATTGAATAGATTTCTTTTTTATTAGAGTATCGGCGA-----
lowerTr           AATGGATCTTATCCTATTCATTGAATAGATTTCTTTTTTATTAGAGTATCGGCGA-----
                  ****************************************** ******** *
```

Figure 3.5. Alignment of modern and ancient Çatalhöyük einkorn Int_TrnL sequences. The asterisks (*) denote consensus nucleotides. The species names are given before the sequences, without genus names. **aestivumGB:** T. aestivum TrnL-F intergenic sequence obtained from GenBank (# AF148757). **Upper CH:** Çatalhöyük Int-TrnL upper band. **Lower CH:** Çatalhöyük Int-TrnL lower band. **LowerTr:** *T. turgidum* (used as positive control in this reaction) Int-TrnL lower band. The CTAAG repeats are underlined.

By observing the alignment result, we can actually accept the two sequences as the same, although there appeared two different sized bands on the denaturing gel. This phenomenon is not very easy to explain. One possible explanation could be that the two (upper and lower bands) were the strands of the same sequence, but the purine/pyrimidine difference between the strands (5/6) allowed for a difference in running speed. A similar interpretation could be that the strands did not totally denature during PAGE, and thus there appeared two bands.

Still another explanation could be that the CTAAG repeats (underlined in Figure 3.5) present at the start of the sequence caused *Taq* polymerase to produce an erroneous sequence, which was later compensated for during re-amplification before cloning. This explanation would suggest that *both* the Çatalhöyük sample and the *T. turgidum* positive control included *T. monococcum* chloroplast DNA (because the repeats are only found in the *T. boeoticum-monococcum* line), and at least the last case is quite unlikely.

3.2. The artificial charring experiment

This experiment was planned after facing successive failures in DNA amplification and the detection of PCR inhibition through "spiking". The general aim was to try to follow the charring process, to assess the efficiency of different DNA extraction methods and to observe closely the PCR inhibition effect of extracts from charred wheat.

3.2.1. Comparison of the charring results with the ancient samples

The first observation that can be made is regarding the colour and brittleness of the artificially charred seeds when compared to authentic ancient seeds, as in Table 3.1.

As can be seen in this table, the black colour is attained by the seeds only after 6 hour exposure to 270°C. Threadgold and Brown (2003) report the highly probable

extermination of amplifiable DNA after 2 hours of exposure to 250°C. This suggests that the possibility of DNA presence in black coloured seeds is very low.

A parallel observation on brittleness was made. In fact, none of the above seeds, including the seeds exposed to 270°C for 11 hours, were as brittle as the ancient ones, which can be classified as "very brittle". This would suggest that the ancient seeds we used have been burnt under more severe conditions.

Table 3.1: Comparison of the colour and brittleness characteristics of modern artificially charred wheat seeds exposed to indicated temperatures for indicated times, and ancient wheat seeds.

Treatment (°C)	**Duration**	**Colour of grains**	**Brittleness**
Ancient seeds	-	Black/shiny black	Brittle/Very brittle
270	11 h	Black (shiny)	Brittle
270	11 h	Black (shiny)	Brittle
270	6 h	Black	Brittle
270	1 h	Brownish black	Partially brittle
200	20 h	Dark brown	Partially brittle
200	15'	Light brown	Solid
None	0'	Yellow	Solid

Of course, when comparing the changes in artificially charred wheat and archaeological charred wheat, and reaching a conclusion, we are actually making two implicit assumptions. One is that the structure and colour does not change upon long term underground burial. This might be a weak assumption, especially regarding brittleness, which might have increased with pressure and carbonization.

The second assumption is that the archaeological wheat grains have had the same constitution and thus changed the same way under heat as modern grains. Considering that

seed coat has thinned as a consequence of selection to reduce seed dormancy (although at early stages of domestication) (Zohary and Hopf, 1994), ancient seed coats might perhaps been slightly more thick on average, so that they might still have preserved some DNA while turning black upon charring.

Artificially charred seeds were also subjected to DNA extraction. The 11 hour 270°C charred seeds were used in CTAB-2 extraction, which yielded a light brown coloured extract as expected, and also in a "direct single seed extraction" attempt by crushing a single seed in sterile water. The results can be seen in Figure 3.6.

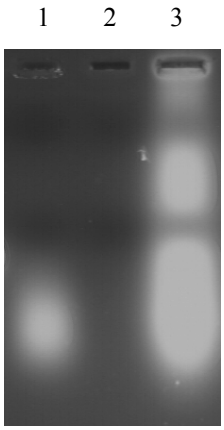

Figure 3.6. Artificially charred seed (11 hour 270°C) extracts, run on 1 % agarose gel and observed under UV light. 15 μL loaded per lane. **Lane 1:** CTAB-2 extract of 10 seeds. **Lane 2:** Blank extract of CTAB-2 extraction. **Lane 3:** Direct single seed extract. A molecular weight marker was not added, but the lower bromophenol blue dye (corresponding to 200-300 bp for 1% gels) indicates that the fluorescent smear in Lane 1 is between 10-100 bp, and the upper xylene cyanol dye (corresponding to 3000-4000 bp for 1% gels) indicates that the high molecular weight smear is between 4000-200 bp.

As can be observed, there exists a high molecular weight fluorescent ingredient in the single seed extract. This may be the pure DNA, DNA that has undergone some chemical

modification (such as binding other molecules) fluorescent Maillard products, or a combination of these. The observation that this HMW smear has been removed upon CTAB-2 extraction suggests that there is at least no DNA in pure form. Still, the low molecular weight smear is present in both Lane 1 and 3, and this might be DNA, but in a very fragmented form, which probably would not yield amplification of 100 bp or larger targets.

The same (putative DNA) smear can be detected in gels of PCRs with ancient samples, where the fluorescent material has been carried over from in the PCR reaction (see Figure 3.8 as an example). This also points out to the difficulty of DNA amplification from the ancient samples used.

From this general comparison it can be inferred that the possibility of PCR-amplifiable DNA presence inside black and brittle wheat grains is appreciably low. Having reached this conclusion, and considering the explicit results of Threadgold and Brown (2003), further quest for determination of the exact temperature and duration of exposure that will wipe out all amplifiable DNA was abandoned. Instead, the PCR inhibiting effect was focused on.

3.2.2. Maillard products and the PCR inhibition effect

There are two main causes for this effect in aDNA studies. One is the salts and similar PCR inhibitors transferred from the soil or along with dung, but this does not seem to be a major problem in our case, while the grains are quite clean. The second and much more probable reason for *Taq* polymerase inhibition is the presence of the so-called Maillard products (described in section 1.3.5.3).

The investigation on the PCR inhibition effect has not been carried to put forward exactly how much of extract, from what kind of charred seed can inhibit PCR under what conditions. But a general conclusion can be drawn from the reproduced experiments for artificial charred seed extracts (CSE):

i) The PCR inhibition effect depends upon the amount of charring. Figure 3.7 gives an example. There is certainly no *Taq* activity in Lane 1 containing 270°C 6 hour CSE, while the inhibition is overcome in the 270°C 1 hour CSE. On the other hand, the primer dimers in Lanes 3 and 4 indicate that PCR is working while there is no amplifiable DNA (although there could be some in the 200°C 15 minutes sample, according to Threadgold and Brown (2003)). It can be seen that the effect is quite strong when compared with the intensity of positive control bands. It can also be added that the *Taq* inhibiting performance of ancient seed extracts was similar to that of 11 hours 270°C charred seeds.

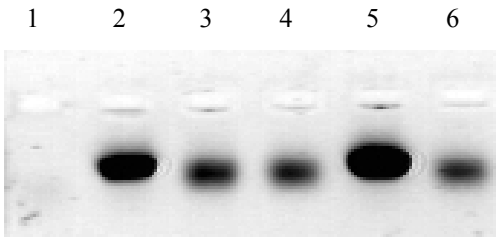

Figure 3.7. Differential PCR inhibition depending upon duration of charring. Inhibition of PCR with TrnL-int primers using artificial CSE of seeds exposed to different doses of heat. Run on 1 % agarose gel and observed under UV light (inverse colours). Positive control used in PCR. **Lane 1:** 100 ng DNA (1 µL) + 270°C 6 hour charred SE (2 µL). **Lane 2:** 100 ng DNA (1 µL) + 270°C 1 hour charred SE (2 µL). **Lane 3:** 200°C 20 hours charred SE (2 µL). **Lane 4:** 200°C 15 minutes charred SE (2 µL). **Lane 5:** Extraction control (2 µL). **Lane 6:** PCR control. The artificially charred seeds and the extraction control seeds (same amount of uncharred Dağdaş-94 bread wheat) were subjected to Qiagen Dneasy extraction kit (*via* column purification). The faint and slightly shorter bands are putative primer-dimers.

ii) The inhibition effect cannot be easily overcome through conventional column purification methods. Figure 3.8 shows the results of two PCRs with nuclear and chloroplast primers on KK DNA. Although column purification has removed the most of the <100 bp flourescent material, it apparently cannot deal with the PCR inhibiting elements and cannot heal the effect.

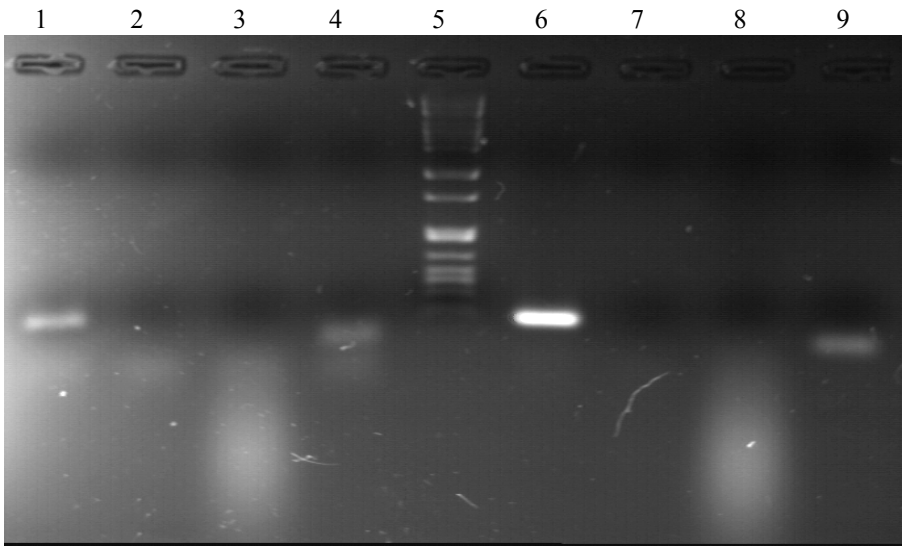

Figure 3.8. PCR inhibition of ancient charred seed extract and extra purified extract from KK. Run on 1 % agarose gel and observed under UV light. **Lane 1-4**: Amplifications with Glu107 primers. **Lane 6-9:** Amplifications with chloroplast Ta16S primers. **Lanes 1 & 6:** 100 ng DNA (Positive control). **Lane 2 & 7:** 100 ng DNA + Qiagen PCR purification kit purified, CTAB-2 extracted KK CSE. **Lane 3 & 8:** 100 ng DNA + CTAB-2 extracted KK CSE. **Lane 4 & 9:** PCR controls. **Lane 10:** λ/PstI marker. The faint lower band in PCR controls are primer dimers.

iii) In addition to PCR inhibition, there exists a more complex effect of artificial CSE: Disappearance of DNA mixed with charred seed extract, when observed under UV light after running on agarose gels (Figure 3.10). This phenomenon might be due to a) the degradation of DNA by a CSE component, which does not seem to be a very high probability; b) the prevention of visualisation of intact DNA, perhaps *via* a CSE component disrupting EtBr incorporation between the DNA strands.

iv) The PCR inhibition effect shows positive correlation with the relative concentration of artificial CSE. In most cases, 1 volume of DNA (100 ng/µL) was inhibited by the addition of 3 volumes of extract (0.25 grams of charred seed eluted in 50 µL water), irrespective of extraction or purification method, while 1:1 v/v ratios were less effective in PCR inhibition (Fig. 3.9). This suggests a competitive relation between DNA and the extract component(s), but other explanations are also possible (Fig 3.9). It also implies that the low amount of DNA in CSE can never be amplified under these conditions.

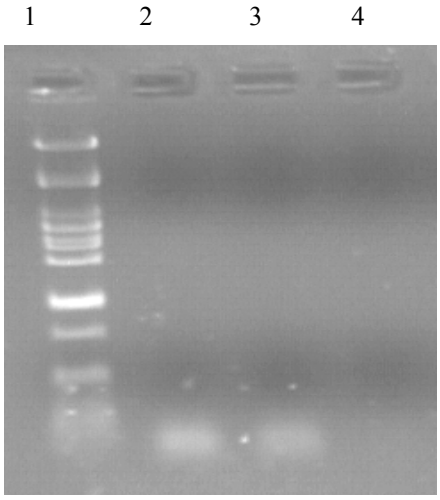

Figure 3.9. Differential PCR inhibition with different concentrations of CSE. Run on 1 % agarose gel and observed under UV light. Glu_107 primers were used. **Lane 1:** λ/PstI marker. **Lane 2**- positive control: 50 ng DNA (1 µL). **Lane 3:** 50 ng DNA (1 µL) + artificial CSE (1 µL). **Lane 4:** 50 ng DNA (1 µL) + artificial CSE (3 µL).

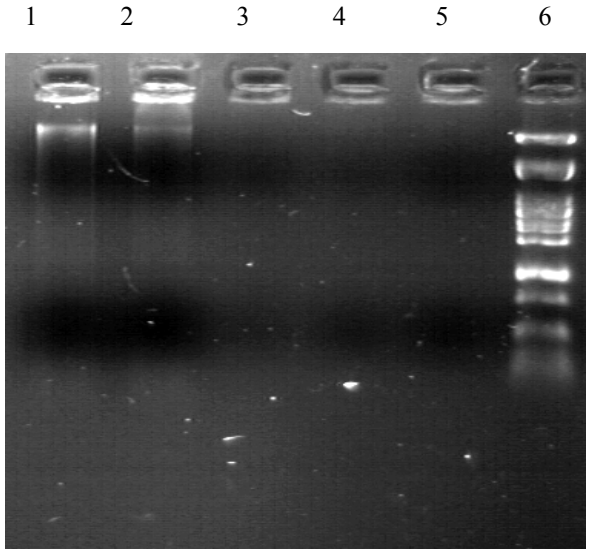

Figure 3.10. Disappearance of intact DNA when mixed with artificial CSE. Run on 1 % agarose gel, observed under UV light. **Lane 1:** 5 µL Intact DNA. **Lane 2:** 5 µL Intact DNA + 5 µL artificial CSE. **Lane 3:** 5 µL Intact DNA + 15 µL artificial CSE. **Lane 6:** λ/PstI marker. All three solutions/mixtures were incubated at room temperature for 7 hours before loading on gel. Intact DNA is 2X CTAB extracted *T. aestivum* DNA (50 ng/ µL). The CSE is from bread wheat exposed to 270°C for 11 hours and extracted by the CTAB method (and has very faint light brown colour).

As a whole, the above observations suggest that the charring products, or Maillard reaction products, do not present a very easily definable effect on PCR, which may be due to the chemical complexity of the reactions themselves This issue requires more systematic study.

3.3. New marker development

3.3.1. General performance of newly designed primers

Most of the designed plasmon primers, namely Ta16S, Chl_HtSpt, Cox1_1 and Cox1_2 worked well, producing single, correct sized and intense bands on agarose gels. The high amplitudes (compared to nuclear primers) can be considered as a reflection of the high copy number of plasmon genomes per cell. Amplification with Br_can_long and Br_can_short gave both multiple bands, probably due to the presence of similar sequences in the wheat genome. Some of the data are given in 3.1.4 and 3.2.2 as gel photos.

The two primer pairs, Cox1_1 and Cox1_2 had been designed to amplify two short and overlapping regions. But the use of Cox1_1 Forward and Cox1_2 Reverse primers as a pair, to cover an area about 300 bp (changing with species) was not satisfactory. It gave inconsistent results, such that single bands, multiple (unspecific) bands or no bands appeared under the same PCR conditions. This can be related to the fact that the region is quite polymorphic, additionally, that the reverse primer had degenerate nucleotide sites.

3.3.2. Brittle rachis candidate marker development

Both hot and cold PCRs were conducted with Br_can_long and Br_can_short (Brittle rachis candidate) primers on domesticated and wild species, the latter also containing wild progenitors such as *T. dicoccoides*. Unfortunately, no domestication specific band, present or absent in all wild progenitor or domesticated species (wild einkorn *vs.* domesticated einkorn; *T. urartu* and wild emmer *vs*. domesticated emmer, durum and bread

wheat) was found (see Figure 3.11). This can be explained in different ways, assuming the primer and PCR conditions appropriate:

i) There was no homologue of the *Oryza sativa* "*shattering*" locus in wheat, and the amplified bands were totally irrelevant.

ii) There was a homologue, but that gene was not among those involved in the "*brittle rachis*" character.

iii) The homologue target was really one of those controlling *Br*, but the actual polymorphism was not one that would be reflected as a DNA length polymorphism able to be detected on a gel. A single nucleotide null mutation, barring the production of the gene products responsible for spikelet abscission is quite probable. Actually, one could only detect an insertion/deletion type mutation by this way, but other types would not reveal themselves so easily.

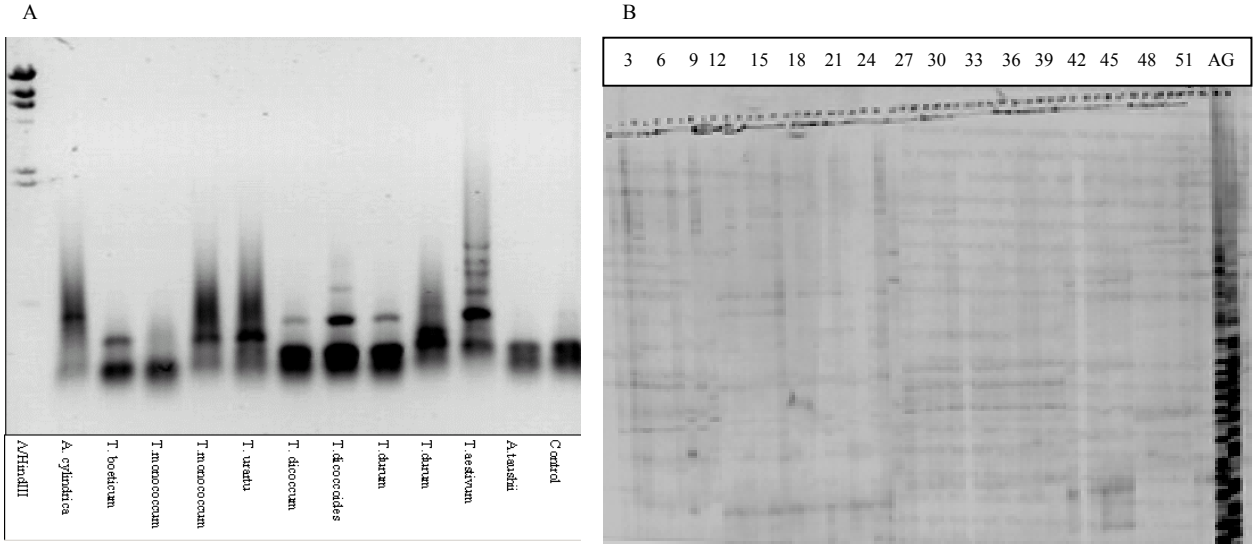

Figure 3.11. Cold (A) and hot (B) PCRs using Br_can_long, conducted with 9 species (11 samples) and 8 species (52 samples), run on agarose and polyacrylamide gels, respectively. In Figure B the species are as follows: Lanes 1-8: *T. urartu.* Lanes 9,10, 35-39: *T. monococcum.* Lanes 11, 12: *T. dicoccum*. Lanes 13-24: *T. dicoccoides*. Lanes 25-34: *T. boeoticum*. Lanes 40-45: *T. durum*. Lanes 46-50: *Ae. tauschii*. Lanes 51,52: *T. aestivum*.

The three alternatives seem equally probable. A bulk segregating analysis *via* RAPD or AFLP, using a segregating population for *Br* and seeking a putative abscission enzyme, would certainly have been more effective, but would still not guarantee success. The recessive state of the mutation renders the search for a polymorphism difficult. Furthermore, the fact that the genetic control mechanism is not clearly understood (see section 1.3.1) hindered the continuation of the study.

3.3.3. The TrnL-F intergenic locus

3.3.3.1. Phylogenetic analysis using the TrnL-F intergenic locus

The reason behind cloning this locus was the amplification of the Çatalhöyük "einkorn" sample with the TrnL-int primers, although this shorter (120 bp) region later turned out to be non-polymorphic (see 3.1.5). Nevertheless, a phylogenetic reconstruction with obtained the sequences was attempted. The multiple sequence alignment of the 385-418 bp region obtained by Clustal X (1.81) is as shown in Figure 3.12.

```
tauschiiGB        AG-TCCCTCTATCCCCAAATCCTCTTTTATTCCCTAACTATACTATATTATTATTAT 56
tauschii          AGGTCCCTCTATCCCCAAATCCTCTTTTATTCCCTAACTATACTATATTATTATTAT 57
cylindirica       AG-TCCCTCTATCCCCAAATCCTCTTTTATTCCCTAACTATACTATATTATTATTAT 56
biuncialis        AG-TCCCTCTATCCCCAAATCCTCTTTTATTCCCTAACTATACTATATTATTATTAT 56
speltoidesGB      AG-TCCCTCTATCCCCAAATCCTCTTTTATTCCCTAACTATACTATATTATTATTAT 56
dicoccoides       AG-TCCCTCTATCCCCAAATCCTCTTTTATTCCCTAACTATACTATATTATTATTAT 56
dicoccum          AG-TCCCTCTATCCCCAAATCCTCTTTTATTCCCTAACTATACTATATTATTATTAT 56
durum             AG-TCCCTCTATCCCCAAATCCTCTTTTATTCCCTAACTATACTATATTATTATTAT 56
aestivumGB        AG-TCCCTCTATCCCCAAATCCTCTTTTATTCC-TAACTATACTATATTATTATTAT 55
?monococcum       AG-TCCCTCTATCCCCAAATCCCCTTTTATTCCCTAACTATACTATATTATTATTAT 56
urartu            AG-TCCCTCTATCCCCAAATCCTCTTTTATTCCCTAACTATACTATATTATTATTAT 56
boeticumGB        AG-TCCCTCTATCCCCAAATCCTCTTTTATTCCCTAACTATACTATATTATTATTAT 56
boeticum          AG-TCCCTCTATCCCCAAATCCTCTTTTATTCCCTAACTATACTATATTATTATTAT 56
monococcum        AG-TCCCTCTATCCCCAAATCCTCTTTTATTCCCTAACTATACTATATTATTATTAT 56
                  ** ****************** ********** ***********************
```

Figure 3.12: Alignment of cloned modern TrnL-F intergenic sequences. The species names are given before the sequences, without genus names. The postscript "**GB**" denotes sequences obtained from GenBank for *T. aestivum, Ae. speltoides, T. boeticum,* and *Ae. tauschii* with accession numbers AF148757, AF519112, AF519168, AF519113, respectively. **monococcum** and **?monococcum** are the einkorn samples TR36938 and IG45257 respectively. The polymorphic region starting at 71st - 72nd bp has been realigned by eye. The relative species are listed together. The asterisks denote consensus nucleotides. The transversion site is enclosed in a box. The two repeat motifs at the main indel site are named as X and Y (shown on the second block with subscripts). Y_{1A} and Y_{1B} are two subdivisions of Y_1.

```
                                  Y1A  Y1B
                    =========Y1=========__X2__ =========Y2=======__X3_____X1_

tauschiiGB        TTATCCTCTTTTTTTT---CTTTTTATCAATG------------------------CAATG 88
tauschii          TTATCCTCTTTTTTTT---CTTTTTATCAATG------------------------CAATG 91
cylindirica       TTATCCTCTTTTTTTT---CTTTTTATCAATG------------------------CAATG 90
biuncialis        TTATCCTCTTTTTTTT--------TATCAATG------------------------CAATG 85
speltoidesGB      TTATCCTCTTTTTTTT---CTTTTTAT----------------------------CAATG 85
dicoccoides       TTATCCTCTTTTTTTT---CTTTTTAT----------------------------CAATG 85
dicoccum          TTATCCTCTTTTTTTT---CTTTTTAT----------------------------CAATG 85
durum             TTATCCTCTTTTTTTT---CTTTTTAT----------------------------CAATG 85
aestivumGB        TTATCCTCTTTTTTTT---CTTTTTAT----------------------------CAATG 84
?monococcum       TTATCCTCTTTTTTTT---CTTTTTAT----------------------------CAATG 85
urartu            TTATCCTCTTTTTTTTT--CTTTTTATCAATGCACTTTTTTTTCTTTTTATCAATGCAATG 115
boeticumGB        TTATCCTCTTTTTTTTTT-CTTTTTATCAATGCACTTTTTTTTCTTTTTATCAATGCAATG 116
boeticum          TTATCCTCTTTTTTTTTT-CTTTTTATCAATGCACTTTTTTTTCTTTTTATCAATGCAATG 116
monococcum        TTATCCTCTTTTTTTTTTCTTTTTATCAATGCACTTTTTTTTCTTTTTATCAATGCAATG 117
                  ***************        ***                              ***


tauschiiGB         GGTTTAAGATTCATTAGCTTTCTCATTCTACTCTTTCACAAAGGAATGCGAAGAGAACT 147
tauschii           GGTTTAAGATTCATTAGCTTTCTCATTCTACTCTTTCACAAAGGAATGCGAAGAGAACT 150
cylindirica        GGTTTAAGATTCATTAGCTTTCTCATTCTACTCTTTCACAAAGGAATGCGAAGAGAACT 149
biuncialis         GGTTTAAGATTCATTAGCTTTCTCATTCTACTCTTTCACAAAGGAATGCGAAGAGAACT 144
speltoidesGB       GGTTTAAGATTCATTAGCTTTCTCATTCTACTCTTTCACAAAGGAATGCGAAGAGAACT 144
dicoccoides        GGTTTAAGATTCATTAGCTTTCTCATTCTACTCTTTCACAAAGGAATGCGAAGAGAACT 144
dicoccum           GGTTTAAGATTCATTAGCTTTCTCATTCTACTCTTTCACAAAGGAATGCGAAGAGAACT 144
durum              GGTTTAAGATTCATTAGCTTTCTCATTCTACTCTTTCACAAAGGAATGCGAAGAGAACT 144
aestivumGB         GGTTTAAGATTCATTAGCTTTCTCATTCTACTCTTTCACAAAGGAATGCGAAGAGAACT 143
?monococcum        GGTTTAAGATTCATTAGCTTTCTCATTCTACTCTTTCACAAAGGAATGCGAAGAGAACT 144
urartu             GGTTTAAGATTCATTAGCTTTCTCATTCTACTCTTTCACAAAGGAATGCGAAGAGAACT 174
boeticumGB         GGTTTAAGATTCATTAGCTTTCTCATTCTACTCTTTCACAAAGGAATGCGAAGAGAACT 175
boeticum           GGTTTAAGATTCATTAGCTTTCTCATTCTACTCTTTCACAAAGGAATGCGAAGAGAACT 175
monococcum         GGTTTAAGATTCATTAGCTTTCTCATTCTACTCTTTCACAAAGGAATGCGAAGAGAACT 176
                   ************************************************************


tauschiiGB         CAATGGATCTTATCCTATTCATTGAATAGATTTCTTTTTTATTAGAGTATCGGCAAGAAA 207
tauschii           CAATGGATCTTATCCTATTCATTGAATAGATTTCTTTTTTATTAGAGTATCGGCAAGAAA 210
cylindirica        CAATGGATCTTATCCTATTCATTGAATAGATTTCTTTTTTATTAGAGTATCGGCAAGAAA 209
biuncialis         CAATGGATCTTATCCTATTCATTGAATAGATTTCTTTTTTATTAGAGTATCGGCGAGAAA 204
speltoidesGB       CAATGGATCTTATCCTATTCATTGAATAGATTTCTTTTTTATTAGAGTATCGGCAAGAAA 204
dicoccoides        CAATGGATCTTATCCTATTCATTGAATAGATTTCTTTTTTATTAGAGTATCGGCAAGAAA 204
dicoccum           CAATGGATCTTATCCTATTCATTGAATAGATTTCTTTTTTATTAGAGTATCGGCAAGAAA 204
durum              CAATGGATCTTATCCTATTCATTGAATAGATTTCTTTTTTATTAGAGTATCGGCAAGAAA 204
aestivumGB         CAATGGATCTTATCCTATTCATTGAATAGATTTCTTTTTTATTAGAGTATCGGCAAGAAA 203
?monococcum        CAATGGATCTTATCCTATTCATTGAATAGATTTCTTTTTTATTAGAGTATCGGCAAGAAA 204
urartu             CAATGGATCTTATCCTATTCATTGAATAGATTTCTTTTTTATTAGAGTATCGGCAAGAAA 234
boeticumGB         CAATGGATCTTATCCTATTCATTGAATAGATTTCTTTTTTATTAGAGTATCGGCAAGAAA 235
boeticum           CAATGGATCTTATCCTATTCATTGAATAGATTTCTTTTTTATTAGAGTATCGGCAAGAAA 235
monococcum         CAATGGATCTTATCCTATTCATTGAATAGATTTCTTTTTTATTAGAGTATCGGCAAGAAA 236
                   ************************************************** *****


tauschiiGB         TCTTGGTTATTCACTCTATTTTTAAGTTTTATTTAAGTAAACCATGCACAATGCATAGGA 267
tauschii           TCTTGGTTATTCACTCTATTTTTAAGTTTTATTTAAGTAAACCATGCACAATGCATAGGA 270
cylindirica        TCTTGGTTATTCACTCTATTTTTAAGTTTTATTTAAGTAAACCATGCACAATGCATAGGA 269
biuncialis         TCTTGGTTATTCACTCTATTTTTAAGTTTTATTTAAGTAAACCATGCACAATGCATAGGA 264
speltoidesGB       TCTTGGTTATTCACTCTATTTTTAAGTTTTATTTAAGTAAACCATGCACAATGCATAGGA 264
dicoccoides        TCTTGGTTATTCACTCTATTTTTAAGTTTTATTTAAGTAAACCATGCACAATGCATAGGA 264
dicoccum           TCTTGGTTATTCACTCTATTTTTAAGTTTTATTTAAGTAAACCATGCACAATGCATAGGA 264
durum              TCTTGGTTATTCACTCTATTTTTAAGTTTTATTTAAGTAAACCATGCACAATGCATAGGA 264
aestivumGB         TCTTGGTTATTCACTCTATTTTTAAGTTTTATTTAAGTAAACCATGCACAATGCATAGGA 263
?monococcum        TCTTGGTTATTCACTCTATTTTTAAGTTTTATTTAAGTAAACCATGCACAATGCATAGGA 264
urartu             TCTTGGTTATTCACTCTATTTTTAAGTTTTATTTAAGTAAACCATGCACAATGCATAGGA 294
boeticumGB         TCTTGGTTATTCACTCTATTTTTAAGTTTTATTTAAGTAAACCATGCACAATGCATAGGA 295
boeticum           TCTTGGTTATTCACTCTATTTTTAAGTTTTATTTAAGTAAACCATGCACAATGCATAGGA 295
monococcum         TCTTGGTTATTCACTCTATTTTTAAGTTTTATTTAAGTAAACCATGCACAATGCATAGGA 296
                   ************************************************************
```

Figure 3.12: (continued)

Figure 3.12: (continued)

The general characteristic of the region is a reflection of the features of plant chloroplast DNA as described by Wolfe *et al*. (1987) and Ogihara *et al*. (1991): Low nucleotide polymorphism and frequent structural changes. The strong conservation in the majority of residues in this region may hint that it is a coding one, but the sequence does not match any cDNA sequences in the GenBank as revealed by Blast search; in addition, it contains many inner stop codons.

The main polymorphic portion of the region starts from the 71st - 72nd basepairs and contains a number of insertion/deletions. The insertion/deletions are comprised of two repeated motifs, which have been given names X and Y, and are shown on Figure 3.12. As for the single nucleotide polymorphisms at this locus, most are non-informative (see the next section).

Under these circumstances, it is better to limit the phylogenetic analysis to maximum parsimony. First of all, when the data is limited, all methods tend to give erroneous results in terms of congruence with the real topology. There may not be a very significant difference between the maximum parsimony, maximum likelihood and the distance based methods, and MP is at least as good as the other methods (Nei and Kumar, 2000). Second and more importantly, for this locus, a classical nucleotide substitution model will not convey enough information to construct a useful tree, whereas the use of the insertion/deletions of the repeats as (partially irreversible) shared derived characters will be quite convenient to approach a true tree topology (Graur and Li, 2000; Nei and Kumar, 2000).

Thus, the X and Y repeats were assigned specific character states. PAUP* has proved to be quite helpful in this work. First of all, it was necessary to make use of the single nucleotide gaps (considering the scarcity of polymorphisms), but avoid assigning new character states (*i.e.* a 5^{th} base) to the gaps in the indel region. Hence the designation of the gaps comprising the repeats were changed from the gap signal "-" to the missing data signal "?", while the remaining gaps were treated as 5^{th} bases. Secondly, a new set of characters signifying the presence or absence of gaps/repeats (in binary format) were added to the data matrix using the PAUP* "interleave" option. Three such repeat regions were included: Y_{1A}, X_2 and $Y_2 + X_3$.

One assumption regarding these three new characters was to accept a null "000" ancestral state. This seems straightforward, and is supported by the finding by Ogihara and Ohsawa (2002) that Y_2 and X_3 appears as a duplication event when compared to other wheat species. But this choice actually did not affect the tree topology (data not shown).

The three new characters were also assumed "Dollo" characters, "*i.e.* characters for which derived character states originate only once" and any homoplasy means reversal (Swofford, 1991). It is hard to tell whether these insertion/deletions are truly irreversible, or freely reversible (an asymmetric stepmatrix was not preferred because it would require far too many assumptions than our knowledge on the history of the region could compensate for). But this definition of "Dollo" characters nevertheless seems to suit the case of length mutations such as duplications, which is what we probably had. In contrast, the rest of the characters were assumed to be "unordered. *i.e.*, any character state is permitted to transform directly into any other state" (Swofford, 1991), which more or less complies with single

nucleotide polymorphisms. Actually, comparisons showed that changing the "Dollo" assumption to "all unordered" did not alter the tree topology (data not shown).

Ae. biuncialis was used as outgroup; analysis using microsatellite markers have shown that it is the most distant to the other wheat species (Bilgiç, 2002), and this distance was also apparent in the sequence alignment.

The most significant choice was regarding *a priori* character weights: All nucleotide characters were given equal weight, while the three gap characters were given five fold weight in one search (Figure 3.13) and equal weighting in another (Figure 3.15). Five fold weighting seems to be better suited than equal weighting -albeit totally subjective- when the general probabilities of insertion/deletion *vs*. single nucleotide substitution are taken into consideration.

There are no calculations of chloroplast DNA length mutation frequencies compared to nucleotide substitution frequencies in the literature as far as I know. But Ogihara and Ohsawa (2002) have detected 14 pronounced length mutations among *all Triticum-Aegilops* species, and also in this specific region, there are 3 putative length mutations in contrast to 9 single nucleotide mutations. These considerations can partially justify the choice of extra weighting of the length mutation characters. Actually, this choice allowed a pronounced difference in tree topologies and much better discrimination between the three main clades, when compared to trees constructed from equally weighted data (Figure 3.15).

Lastly, the GenBank sequences for *Ae. tauschii* and *Ae. boeoticum* were not included in the analysis, considering their inconsistencies with the sequences obtained in this study. The reason might be intraspecific polymorphisms or sequencing errors on either side. An alternative would be to use "Multistate taxa=uncertain" option in PAUP*, but this would not contribute to the clarity of the tree.

One of the 3 maximum parsimony trees obtained by exhaustive search with 1:5 character weighting is given in Figure 3.13 (heuristic or branch-and-bound search were not necessary, as the amount of data and the number of taxa were was limited). The consistency index (CI) was found to be , the same as the other three trees. This number is calculated as the ratio of minimum amount of change that the character set may show on *any* conceivable

tree (*m*) to the length (number of steps) required by the character set on the tree being evaluated (*s*) (Swofford, 1991).

The branch lengths, proportional to the number of inferred changes according to the reconstruction are given in Table 3.2.

The 100% majority-rule consensus tree of the three is given in Figure 3.14. Three of the main branches were supported by all three trees. But the results of a bootstrap analysis with 1000 replicates for the 100% consensus tree's branches were different: the highest bootstrap value was 86, and this was much lower for the other branches, while 95% is the conventional threshold value for significance (Nei and Kumar, 2000). However this picture is not unexpected, because the main characters dictating the tree topology were relatively very few in number (thus were not represented in many replicate data sets).

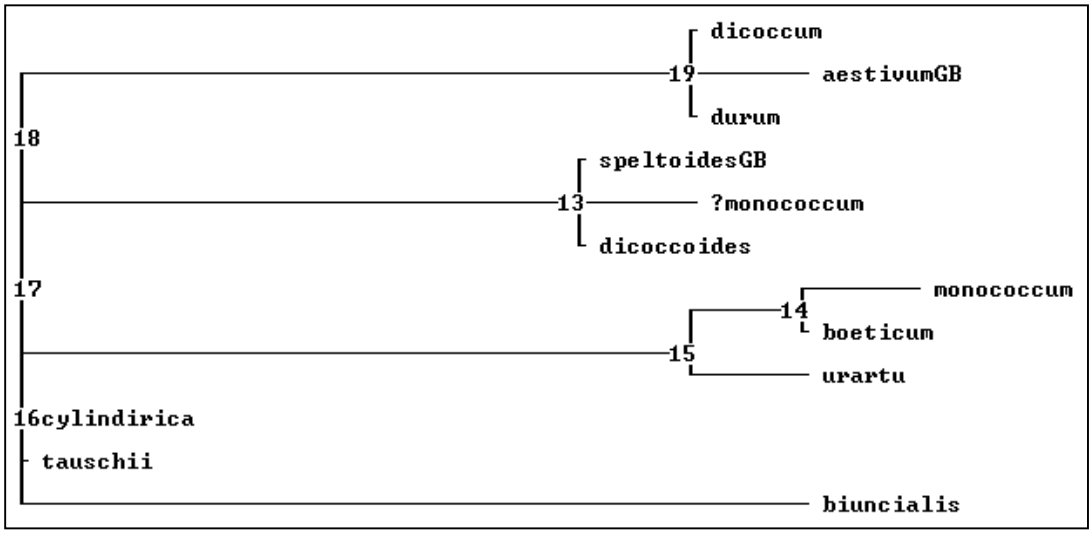

Figure 3.13. One of the three maximum parsimony trees obtained *via* exhaustive search under 5:1 weighting. The PAUP* program has been used under the assumptions described in the text. The length of this tree is 24 and the Consistency index (CI) is 0.9583. Branches with maximum length zero are collapsed to produce polytomies (*T. cylindrica* and *T. tauschii*). The numbers on the nodes are not branch lenghts but node numbers.

Table 3.2. The branch lenghts and linkages for the tree in Figure 3.12. The lenghts are calculated for the unrooted tree.

Node	Connected to node	Assigned branch length	Minimum possible length	Maximum possible length
18	17	0	5	5
19	18	6	0	1
dicoccum (1)	19	0	0	0
aestivumGB (2)	19	1	1	1
durum (3)	19	0	0	0
13	18	5	0	1
speltoidesGB (4)	13	0	0	0
?monococcum (5)	13	1	1	1
dicoccoides (6)	13	0	0	0
16	17	0	0	1
15	16	6	6	6
14	15	1	1	1
monococcum (7)	14	1	1	1
boeticum (8)	14	0	0	0
urartu (11)	15	1	1	1
cylindirica (9)	16	0	0	0
tauschii (10)	16	0	0	0
biuncialis (12)*	17	7	6	7
Sum		29		

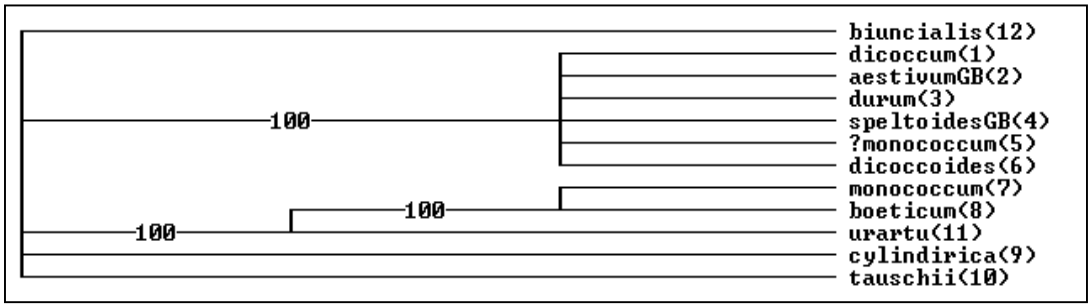

Figure 3.14. The 100% majority-rule consensus tree of the 3 maximum parsimony trees obtained under 5:1 weighting. The branches labelled with "100" are shared in all trees, showing that the main splits are supported by all three trees.

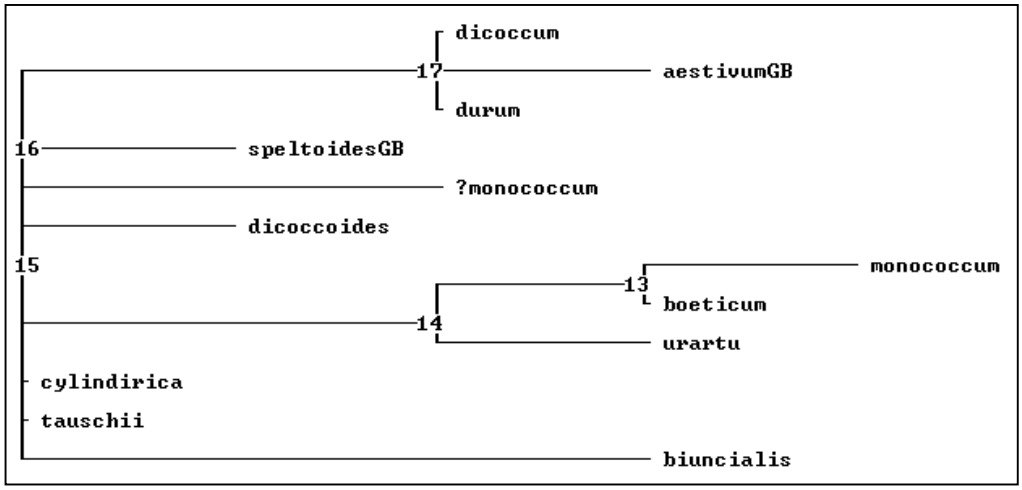

Figure 3.15. One of the four maximum parsimony trees obtained *via* exhaustive search under with equal weighting of characters. The length of this tree is 24 and the Consistency index (CI) is 0.9167. As observed, the branches of Ae. *speltoides,*?monococcum and *T. dicoccoides* are collapsed into polytomies, differing from the tree in Figure 3.12.

The distance matrices calculated by PAUP* are not shown because the program cannot make use of gaps, and thus they are solely based on three informative sites and yield distances totally incongruent with the real species tree.

In the equally weighted tree in Figure 3.15, *T. dicoccoides, Ae. speltoides* and the unidentified sample "?monococcum" are not included in the same clade with the domestic species (*T. dicoccum, T. durum* and *T. aestivum*). Moreover, the branches leading to these three are collapsed, resulting in a polytomy. This topology is thus both imprecise and not congruent with the true species tree.

In the unequally weighted trees (Figure 3.12 and 3.13), on the other hand, there are three main clades observed, which correspond to three haplotypes. One of the three clades comprises *Ae. speltoides*, *T. dicoccoides* and the AB(D) genome domestic species, and can be called the S/B haplotype. This grouping strongly supports the reports of Miyashita *et al*. (1994) and Wang *et al*. (1997) that the maternal donor of tetraploid progenitor and hexaploid domestic species has been *Ae. speltoides* but not *T. urartu*, and *T. dicoccum* but not *Ae. tauschii*, respectively (see Figure 1.1).

The other two clades contain the species with the A genome (*T. urartu*, *T. boeoticum*, *T. monococcum*) and the D genome (*Ae. tauschii* and *Ae. cylindrica*). This last classification also suggests that the maternal donor of *Ae. cylindrica* (genome CD) has been a D genome carrier, although a comparison with a C genome carrier is needed to verify this hypothesis.

A surprising observation made at this point is that the sample classified as *T. monococcum* (IG45257) and designated as "?monococcum" in the above figures is grouped together with the S/B clade instead of the expected A genome clade. Actually, Bilgiç (2002) has reported from her microsatellite study that this sample (along with some others obtained from ICARDA) is misclassified, and she has excluded it from further analysis. The above result confirms her conclusion, and suggests that the sample was a S/B genome carrier (a hulled sample such as *T. dicoccum* is quite probable).

3.3.3.2. Analysis of TrnL-F polymorphisms

Leaving aside the repeat polymorphisms in the second alignment block in Figure 3.12, there are a total of nine single nucleotide polymorphisms throughout the region. Six of them are insertion/deletions, two are transitions and one is a transversion. Except the transversion and two insertions (the first and second T insertions in Y_1 in Figure 3.12), all are non-informative –*i.e.* not contributing to phylogenetic reconstruction (Page and Holmes, 1998); they could be autapomorphies or consequences of sequencing errors (*e.g.* one could be suspicious of the G insertion at the second nucleotide in "tauschii" of not being an authentic polymorphism; it is not shared by any other sequence, including "tauschiiGB").

The sharing pattern of the A→T transversion is complex. The ancestral nucleotide seems to be A, and T is common in all domestic AB(D) genome wheat, thus the transversion appears to be an apomorphism implying domestic status. In this case, the presence of T in the *Ae. biuncialis* sequence (which is most distant) can be explained as a convergence. On the other hand, there exist inconsistencies between GenBank and my sequences for both *Ae. tauschii* and *T. boeoticum*: T is present in the GenBank sequences and A in the sequences obtained in this study. This suggests that there might have been a biased error during Taq

amplification or sequencing; and when the presence of AAATTT repeats at the surroundings is considered, this does not appear to be a very low possibility. The only way to know for sure would be to sequence more samples.

The insertion/deletion polymorphisms starting from the 71st - 72nd basepairs are much more informative, as discussed in the former section. The plesiomorphic state here can be assumed as one containing the regions Y_1 and X_1, (shown on Figure 3.12). Following this assumption, we can reason that there have occurred at least two major insertion events, or in other words, duplications. One is the insertion of X_2 in the D genome lineage (*Ae. tauschii* and *Ae. cylindirica*, the other, the insertion of $X_2 + Y_2 + X_3$ in the A genome lineage (*T. urartu, T. boeoticum, T. monococcum*). Ogihara and Ohsawa (2002) have also determined the $Y_2 + X_3$ duplication in *T. monococcum*. The repeat motifs at this region must have prepared a convenient infrastructure for such events.

Of course, the actual history of this region might have also been different. For example, the ancestral state might have been one containing X_2; in that case the S/B lineage would have undergone a deletion, and in the A genome lineage, there would have occurred a duplication of the $Y_1 + X_2$ region resulting in the addition of $Y_2 + X_3$. The alignment motif of the *Ae. biuncialis* (genome UM) sequence is different from the others, but it also seems to support this explanation (one must mind that the alignment of the TATCAATG in the $Y_1 + X_2$ regions is totally subjective, and this sequence could also be aligned with $Y_2 + X_3$). Nevertheless, this latter alternative might not be as parsimonious as the former one. To decide on the real ancestral state, again a more comprehensive study -including other *Aegilops L.* species- is required.

A last observation about this region is the T insertions at the end of Y_1 in the A genome lineage. They increase in correlation with genetic distance and are most probably authentic. If these insertions are further confirmed, they can be used as a single nucleotide polymorphism marker for distinguishing between wild (*T. boeoticum*) and domestic (*T. monococcum*) species.

In general, this most polymorphic part of the TrnL-F locus seems to be suitable to be used as a marker for distinguishing between S/B, A and D maternal lineages. Furthermore, the length polymorphism can be easily detected on a polyacrylamide gel using primers

flanking the region, thus allowing for shortcut identification with a short sequence; but one must not ignore the fact that such a path will not be reliable for aDNA studies, for at least today. Also, the slippery character of this repeat-rich region may allow for simple DNA polymerase errors. But perhaps the most important point is that, although the discrimination between S/B, A and D is usually not very hard, the critical problem is to determine species and domestication status.

If the transversion at 290th bp in the *Ae. tauschii* sequence, and the T insertion number polymorphism in repeat Y_1 are shown to be valid for all populations of these species, then these mutations can be used as reliable markers showing domestication status. Primers targeting an approximately 250 bp region would usually suffice and be suitable for aDNA studies.

Still, it will be hard to conclude that the locus is a perfect candidate for studies of wheat phylogeny.

CHAPTER IV

CONCLUSION

The history of wheat (genus *Triticum L.*) domestication and evolution deserves attention as a multidisciplinary area where both biology-evolution and social dynamics have played roles. In return, the investigation of this history holds the potential to answer many questions regarding processes such as speciation, domestication, accumulation of variation, and the start, diffusion and widespread establishment of agriculture. When the prominent position of wheat as a contemporary crop is taken into account, this area of study attains even more importance.

As a new method in this area, the popularity of ancient DNA studies in revealing biological and social history has been rising, albeit with fluctuations. Risks of aDNA study of charred plant material, in addition to the general risks of aDNA studies such as contamination, have limited the use of this method on ancient wheat remains.

The intention of this thesis work has been to contribute to aDNA studies, especially aDNA of charred wheat grains, as the dominant form found in Anatolian archaeological sites, in the context of investigation of wheat history.

In this study, 4 independent extraction attempts aimed to extract DNA from charred wheat grains -of different species from different periods and different Anatolian archaeological sites- have been conducted; namely Kaman Kalehöyük Early Bronze Age domestic einkorn and naked wheat, Ottoman period naked wheat, Yenibademli Höyük domestic emmer, İmamoğlu Höyük Early Bronze Age naked wheat and Patnos Urartu period naked wheat. Some of these extracts have been subsequently purified *via* column extraction.

All extracts have been used directly in more than 20 “cold” and 10 “hot” PCR reactions. Unfortunately, neither of the samples used were able to yield any PCR amplification products, while the results for negative controls, blanks and positive controls were as generally expected.

This result does not appear to be an unexpected outcome when the reports and cold attitudes by other former wheat aDNA investigators are considered (as discussed in section 1.3.5.3). Just the opposite is true, and success seems to be a low possibility. The main reason for this is, in addition to or even more important than the age of DNA, the effect of charring on seed DNA and the seed itself.

Thus, one general attempt has been made to compare and optimise different DNA extraction and purification methods. A conservative method (in terms of DNA preservation) such as CTAB-2 developed in this study, strengthened by twice chloroform:isoamyl extraction and followed by column purification may be efficient, but seems to be insufficient to overcome PCR inhibition of the (presumed) DNA extract.

Another main focus of this study has been to determine the effects of charring *via* an artificial charring experiment. The results have shown that much of the DNA in charred wheat exposed to ≥ 200°C for even short time periods, which should not be an unusual case for many charred seeds, is degraded considerably (to pieces less than 50 bp) and loses its capacity to amplify. Secondly, the experiment has suggested that the Maillard reaction products produced during charring can be very effective in PCR inhibition, and would certainly not allow any small amounts of DNA to be amplified by PCR: Moreover, these products cannot be eliminated by many kit purification methods, and furthermore, they may exert complex effects on DNA, such as prevention of visualisation or even degradation of DNA.

Faced with these obstacles, charred wheat grain aDNA studies may not seem to have a very bright future. Archaeobotanists may not prefer to share their priceless material with molecular geneticists for such risky experiments. But in fact, some further work done in this field could shed light on issues which have been considered obscure and have been a potential source of risk.

One is the need to continue to seek for the exact threshold point where wheat seeds lose their amplifiable DNA. For this, of course, the optimisation of charred seed DNA purification and amplification methods must be accomplished. It can be noted that trying two important methods that were not used in this study (the labour-intensive silica-extraction method and PTB addition against PCR inhibition) systematically on charred wheat samples would indeed be informative and perhaps indicate more directly the dimensions of the obstacles faced. As a second step, morphological or chemical criteria can be developed indicating whether this threshold has been surpassed or not in the ancient samples. This will allow the incorporation of only those ancient seeds which have a high potential of containing DNA. In addition to saving time and resources, this approach would also allow preservation of irreplaceable ancient material.

The last general aim of this study has been to develop new molecular markers to be used in wheat aDNA investigations. Their contribution will consolidate aDNA results and make them more precise. In addition, loci with high number of single nucleotide substitution polymorphisms may allow divergence time calculations, which will be priceless in wheat history studies.

A quest for a marker implying domestication status, for which the best candidate in wheat is the *Brittle rachis* gene, has been conducted making use of a putative homologous gene in rice and relatively superficially, and it has not yielded any results, at least for now. On the other hand, the discovery of the mentioned gene(s) is a very interesting and important issue. It should be conducted as a part of a more comprehensive study using a segregating population for both einkorn and emmer. The results will not only be meaningful for those working on wheat aDNA, but also for wheat domestication studies in general.

Chloroplast and mitochondrial markers are convenient due to their haploid character and high copy-number. On the other hand, the low number of single nucleotide polymorphisms in contrast to the high number of length mutations found in these genomes constitutes a partial disadvantage, first due to the size limitation in aDNA studies, and secondly because only *nucleotide substitution* rates can be used to estimate divergence times between taxa.

The TrnL-F intergenic locus analysed in this study also exhibited the features characteristic of chloroplast DNA. 10 clones from different wheat species were obtained for this locus. The analysis of sequence polymorphisms and phylogenetic reconstruction using maximum parsimony allowed discrimination between A (*T. urartu, T. boeoticum, T. monococcum*) , S/B (*Ae. speltoides, T. dicoccoides, T. dicoccum, T. durum, T. aestivum*) and D (*Ae. tauschii, Ae. cylindrica*) genome maternal lineages, and it was possible to determine maternal donors of hybrid wheat species such as *T. dicoccoides*, *T. aestivum* and *Ae. cylindrica.* The analysis also allowed identifying a previously misclassified wheat sample.

But this locus did not prove very handy for further identification at the species level, but still, a number of single nucleotide polymorphisms which could indicate domestication status were found. Still, these results need to be further confirmed.

On the other hand, a short (120 bp) portion of this intergenic region amplified from an ancient wheat DNA sample -Çatalhöyük "einkorn" from Bilgiç (2002)- turned out to be totally uninformative. This troublesome result calls for caution in primer design in phylogenetic studies.

It is hoped that the results of this study will contribute to further investigation of charred wheat aDNA, especially in Turkey where archaeobotany holds an enormous but neglected potential, and in general to aDNA studies, where there is a similar situation present.

REFERENCES

Allaby R.G. and Brown T.A. (2003) "AFLP data and the origins of domesticated crops" in *Genome* 46(3):448-53

Allaby, R.G., Jones, M.K. and Brown, T.A. (1994) "DNA in Charred Grains from the Iron Age Hillfort at Danebury, England", *Antiquity*, 68:126-132.

Allaby, R.G., Banerjee, M. and Brown, T.A. (1999) "Evolution of the High Molecular Weight Glutenin Loci of the A, B,D and G Genomes of Wheat", *Genome*, 42:296-307.

Audic, S. and Beraud-Colomb, E. (1997) "Ancient DNA is Thirteen Years Old", *Nature Biotech.*, 15:855-858.

Badaeva, E.D., Shkutina, F.M., Bogdevich, I.N. and Badaev, N.S. (1986) "Comparative Study of *Triticum aestivum* and *T. timopheevii* Genomes Using C-Banding Technique", *Plant Syst. Evol*., 154:183-194.

Badr, A., Müller, K., Schafer-Pregl, R., El rabey, H., Effgen, S., Ibrahim, S.S., Pozzi, C., Rohde, W., Salamini, F. (2000) "On the origin and domestication history of barley" in *Mol. Biol. Evol.* 17:499-510.

Bahrman, N, Zivy, M. and Thiellement H. (1988) "Genetic relationships in the Sitopsis section of *Triticum* and the origin of the B genome of polyploid wheats" *Heredity*, 61:473-480.

Becker, J., Vos, P., Kuiper, M., Salamini, F. and Heun, M. (1995) "Combined mapping of AFLP and RFLP markers in barley." *Mol Gen Genet* 249: 65-73.

Belyayev, A, Raskina, O, Korol, A, Nevo, E. (2000) "Coevolution of A and B genomes in allotetraploid Triticum dicoccoides" in *Genome* 43:1021-6.

Bilgiç, H. (2002) "Genetic relationship of wild and primitive wheat species from turkey based on microsatellite markers and ancient DNA Analysis", *PhD. Thesis*, Middle East Technical University, Ankara, Turkey.

Blatter ,R. H. E., Jacomet, S. , Schlumbaum A. (2002) "Spelt-specific alleles in HMW glutenin genes from modern and historical European spelt (*Triticum spelta* L.)" *Theor. Appl. Genet.* 104:329-337.

Bohn, M., Utz, F. and Melchinger, A.E. (1999) "Genetic Similarities among Winter Wheat Cultivars Determined on the Basis of RFLPs, AFLPs and SSRs and Their Use for Predicting Progeny Varience", *Crop. Science*, 39:228-237.

Braun, H-J. (2000) "Prospects of Turkey's wheat industry, breeding and biotechnology", Proceedings of the *Symposium of the Problems in Crop Production in Central Anatolia*, 8-11 June 1999, Konya, Turkey. Ed: H. Ekiz, Publication of Bahri Dağdaş International Center of Winter Type Cereals, pg.1-13. ISBN:975-487-071-3.

Briggle, L.W. (1967) "Morphology of the Wheat Plant." in *Wheat and Wheat Improvement,* Am. Soc. of Agronomy Monograph. Retrieved from: http://www.members.shaw.ca /oldwheat/Resources/PhDwheatpaper.pdf (23/08/2003)

Brown, T.A. and Brown, K.A. (1992) "Ancient DNA and the archaeologist" *Antiquity*, 66:10-23.

Brown, T.A., Allaby, R.G., Brown K.A., O'Donoghue and Sallares, R. (1994) "DNA in Wheat Seeds from European Archaeological Sites", *Experientia*, 50:571-575.

Brown, T.A., Allaby, R.G., Brown, K.A. and Jones M.K. (1993) "Biomolecular Archaeology of Wheat", *World Archaeometry*, 25:64-73.

Cooper, A. and Poinar, H.N. (2000) "Do it right or not at all" in *Science* 289: 1139.

Cooper, A. and Wayne, R. (1998) "New uses for old DNA" in *Current Opinion in Biotechnology* 9:49-53.

Damania A.B., Pecetti, L., Qualset, C.P., Humeid, B.O., (1997) "Diversity and geographic distribution of stem solidness and environmental stress tolerance in a collection of durum wheat landraces from Turkey", *Genet. Res. Crop Evol*., 44:101-108

Debruyne, R., Barriel, V., Tassy, P. (2003) "Mitochondrial cytochrome b of the Lyakhov mammoth (Proboscidea,Mammalia): new data and phylogenetic analyses of Elephantidae" in *Molecular Phylogenetics and Evolution* 26:421-434

Diamond, J. (1997) "Location, location,location: The first farmers" *Science* 278: 1243-1244.

Dorofeev, V.F. and Migushova, E.F. (1979) "Flora of cultivated plants." *1. Wheat. Klos*, Leningrad.

Dvorak, J and Zhang, H-B. (1990) "Variation in repeated nucleotide sequences sheds light on the phylogeny of the wheat B and D genomes" *Proc. Natl. Acad. Sci. USA*, 87:9640-9644.

Dvorak, J., McGuire, P.E. and Cassidy, B. (1988) "Apparent sources of the A genome of wheats inferred from polymorphism in abundance and restriction fragment length of repeated nucleotide sequences." *Genome* 30: 680-689.

Dvorak, J., di Terlizzi, P., Zhang, H.-B., Resta, P. (1992) "The evolution of polyploid wheats: identification of the A genome donor species" *Genome*, 36:21-31.

Dvorak, J., Luo, M.C., Yang, Z.I. and Zhang, H.B. (1998) "The Structure of the *Aegilops tauschii* genepool and the Evolution of Hexaploid Wheat", *Theo. Appl. Genet*., 97:657-670.

Dvorak, J. and Luo, M-C. (2001) "Evolution of free-threshing and hulled forms of Triticum aestivum: old problems and new tools" in Wheat Taxonomy: The Legacy of John Percival, *The Linnean Special Issue 3*, The Linnean Society of London, Academic Press, London, GB pp.127-136.

Fahima, T., Röder, M.S., Grama A. and Nevo, E. (1998) "Microsatellite DNA Polymorphism Divergence in *Triticum dicoccoides* Accessions Highly Resistant to Yellow Rust", *Theor. Appl. Genet*., 96:187-195.

FAO, 2000, FAOSTAT Agriculture Database (http://apps.fao.org/cgi-bin/nph-db.pl?subset=agriculture)

Feldman, M., Lupton, F. G. H. and Miller, T. E. (1995) "Wheats." in *Evolution of Crop Plants*, Second Edition, Smart J. and Simmonds, N.W. (eds.) Longman Scientific & Technical.

Fernandez-Calvin, B. and Orellana, J. (1990) "High molecular wight glutenin subunit variation in the sitopsis section of Aegilops. Implications for the origin of the B genome of wheat" *Heredity*, 65:455-463.

Gepts, P. (2002) http://agronomy.ucdavis.edu/gepts/pb143/pb143.htm

Gregova, E., Tisova, V., Kraic, J. (1997) "Genetic variability at the Glu-1 loci in old and modern wheats (*Triticum aestivum* L.) cultivated in Slovakia" in *Genetic Resources and Crop Evolution*, 44:301-306.

Gökgöl, M. (1939) *Türkiye Buğdayları*. Tarım Bakanlığı İstanbul Yeşilköy Tohum Islah ve Yetiştirme Enstitüsü Yayını 14 Devlet Matbaası, İstanbul.

Golenberg, E.M., Gianassi, D.E., Clegg, M.T., Smiley, C.J., Durbin, M., Henderson, D. and Zurawski, G. (1990) "Chloroplast DNA sequence from a Miocene Magnolia species" *Nature* 344:656-658.

Goloubinoff, P., Pääbo, S., Wilson, A.C. (1993) "Evolution of maize inferred from sequence diversity of an adh2 gene segment from archaeological specimens"in *Proc. Nat. Acad. Sci. USA,* 90:1997-2001.

Graur, D., Li, W.-H. (2000) *Fundamentals of Molecular Evolution,* 2nd edition, Sinauer Associates, Sunderland Massachusetts.

Hakkı, E.E. (2000) "A novel method for development microsatellite markers and sequence analysis of wheat disease resistance gene analogs." *PhD. Thesis*, Middle East Technical University, Ankara, Turkey.

Harlan, J.R. (1971) “Agricultural origins: Centers and noncenters” *Science* 174: 469-474.

Harlan, J.R. (1987) “The early history of wheat: Earliest traces to the sack of Rome” in L.T. Evans and W. Peacock (eds.) *Wheat Science Today or Tomorrow*. Cambridge, U.K.: Cambridge University Press.

Harlan, J.R. (1992). “Crops and Man.” *American Society of Agronomy, Inc., and Crop Science Society of America, Inc*., Madison, Wisconsin.

Harlan, J.R. and Zohary, D. (1966) “Distribution of wild wheats and barley.” *Science* 153: 1074-1080.

Harris, D. (1996) “Introduction: Themes and concepts in the study of early agriculture” in D. Harris (ed.) *The origins and spread of agriculture and pastoralism in Eurasia.,* Smithsonian Institution Press, Washington D.C. pp. 1-11

Harris, D. (ed.) (1996) *The origins and spread of agriculture and pastoralism in Eurasia.* Smithsonian Institution Press, Washington D.C.

Helbaek, H. (1964) *Anatolian Studies*, 14:121-123.

Heun M., Schafer-Pregl, R., Klawan, D., Castagna, R., Accerbi, M., Borghi, B. and Salamini, F. (1997). “Site of einkorn wheat domestication identified by DNA fingerprinting.” *Science*, 278: 1312-1314.

Higuchi, R., Bowman, B., Freiberger, M., Ryder, O.A. and Wilson, A.C. (1984) “DNA sequences from the quagga, an extinct membr of the horse family” *Nature*, 312:282-284.

Hillman, G. (1996) “Late Pleistone changes in wild plant-foods available to hunter gatherers of the Northern Fertile Crescent: possible preludes to cereal cultivation” in D. Harris (ed.), *The origins and spread of agriculture and pastoralism in Eurasia,* Smithsonian Institution Press, Washington D.C.pp.159-203

Hillman, G. (2001) “Archeology, Percival, and the problems of identifying wheat remains ” in Wheat Taxonomy: The Legacy of John Percival, *The Linnean Special Issue 3*, The Linnean Society of London, Academic Press, London, GB pp.27-36.

Hillman, G. C. (1975) Appendix A. “The Plant Remains from Tell Abu Hureyra: A Preliminary Report.” *Proceedings of the Prehistoric Society* 41: 70-73.

Hillman, G. C. and Davies, M.S. (1990) “Measured domestication rates in wild wheats and barley under primitive cultivation, and their archeological implications.” *J. World History* 4(2): 157-222.

Hillman, G. C., Mason S., de Moulins, D. & Nesbitt, M. (1996), “Identification of archaeological remains of wheat: the 1992 London Workshop”, *Circaea, J. Assoc. Environ. Arch.* 12:195-209.

Hillman, G.C., Wales, S., McLaren, F., Evans, J. and Butler, A. (1993) “Identifying problematic remains of ancient plant foods: a comparison of the role of chemical, histological and morphological criteria” *World Archaeology*, 25:94-121.

Hofreiter, M., Serre, D. Poinar, H. N. Kuch, M. & Pääbo, S. (2001) “Ancient DNA” *Nature Rev. Genet.* 2, 353-359.

Höss, M. and Pääbo, S. (1993) “DNA extraction from Pleistocene bones by a silica-based purification method” in *Nucleic Acids Research*, 21(16):3913-3914.

Höss, M., Dilling, A., Currant, A., Pääbo S. (1996) “Molecular phylogeny of the extinct ground sloth *Mylodon darwinii*” in *Proc. Natl. Acad. Sci.* 93: 181-185.

Huang, Q, Borner, A, Roder, S, Ganal, W. (2002) “Assessing genetic diversity of wheat (Triticum aestivum L.) germplasm using microsatellite markers” in *Theor Appl Genet*. 105(5): 699-70.

Huang, S., Sirikhachornkit, A., Su, X., Faris, J., Gill, B., Haselkorn, R., Gornicki, P. (2002) “Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the

Tirticum/Aegilops complex and the evolutionary history of polyploid wheat" in *Proc. Natl. Acad. Sci.* 99: 8133-8138.

Hubbard, R.N.L.B. (1992) "Dichotomous keys for the identification of major Old World crops" in *Rev. Paleobotany and Palynology*, 73: 105-115.

Ikeda, TM, Terachi, T, Tsunewaki, K. (1992) "Variations in chloroplast proteins and nucleotide sequences of three chloroplast genes in Triticum and Aegilops" in *Jpn J Genet.* 67(2): 111-23.

Incirli, A. and Akkaya, M.S. (2001) "Assessment of Genetic Relationships in Durum Wheat Cultivars Using AFLP Markers." *Genetic Resources and Crop Evolution* 48: 233-238.

Ishii, T., Mori, N. and Ogihara, Y. (2001) "Evaluation of allelic diversity at chloroplast microsatellite loci among common wheat and its ancestral species" *Theor. Appl. Genet*. 103:896-904.

Jones, M., Brown, T., Allaby, R. (1996) "Tracking early crops and early farmers: the potential of biomolecular archaeology" in D. Harris (ed.), *The origins and spread of agriculture and pastoralism in Eurasia,* Smithsonian Institution Press, Washington D.C. pp. 93-100

Karagöz, A. (1996) "Agronomic practices and socioeconomic aspects of emmer and einkorn cultivation in Turkey" in Hulled Wheats Proceedings of the First International Workshop on Hulle Wheats 21-22 July 1995, Pascoli, Italy, Padulosi, S., Hammer, K. and Heller J. eds. IPGRI, Rome, Italy pp. 172-177.

Kimber G and Sears ER. (1987) "Evolution in the genus *Triticum* and the origin of cultivated wheat" In: *Wheat and wheat improvement*, 2nd Ed (Heyne, E.G. ed.). Ameican Society of Agronomy, Madison, WI. pp. 154-164.

Kimber, G. and Feldman, M. (1987) *Wild Wheat, an introduction*. Special Report 353, College of Agriculture, University of Missouri, Columbia. 146 pp.

Lewin, R. (1999) *Modern İnsanın Kökeni*, (translator Özüaydın, N.) 5th edition, TÜBİTAK, Ankara, Turkey.

Miyashita N.T., Mori, N., Tsunewaki, K. (1994) "Molecular variation in the chloroplast DNA regions in ancestral species of wheat" in *Genetics* 137: 883-889.

Mori, N, Miyashita, NT, Terachi, T, Nakamura, C. (1997) "Variation in coxII intron in the wild ancestral species of wheat." in *Hereditas*;126: 281-8.

Mori, N., Liu, Y.G. and Tsunewaki, K. (1995) "Wheat Phylogeny Determined by RFLP Analysis of Nuclear DNA. 2. Wild Tetraploid Wheats", *Theor. Appl. Genet*., 90:129-134.

Morrison, L.A. (1993) "*Triticum-Aegilops* systematics: taking an integrative approach. in Biodiversity and wheat improvement" in A.B. Damania, ed. John Wiley and Sons, London. Pp.59-66.

Morrison, L.A. (2001) "The Percival Herbarium and wheat taxonomy: yesterday, tday and tomorrow" in Wheat Taxonomy: The Legacy of John Percival, *The Linnean Special Issue 3*, The Linnean Society of London, Academic Press, London, GB pp.65-80.

Nei, M., Kumar, S. (2000) *Molecular Evolution and Phylogenetics,* Oxford University Press, New York.

Nesbitt, M. (1993) "Ancient crop husbandry at KK-Kalehöyük: 1991 archeaobotanical report" in H.I.H. Prince Takahito Mikasa (ed.), *BMECCJ* VII: *Essays on Anatolian archaeology*, Weisbaden, Harrosowitz Verlag, pp.75-97.

Nesbitt, M. (2001) "Wheat evolution: integrating archaeological and biological evidence" in Wheat Taxonomy: The Legacy of John Percival, *The Linnean Special Issue 3*, The Linnean Society of London, Academic Press, London, GB pp.37-59.

Nesbitt, M. and Samuel, D. (1996), "From staple crop to extinction? The archaeology and history of the hulled wheats" *in Hulled Wheats: Proceedings of the First International Workshop on*

Hulled Wheats 21-22 July 1995, Pascoli, Italy, Padulosi, S., Hammer, K. and Heller J. eds. IPGRI, Rome, Italy pp. 41-100.

Nevo, E., Beiles, A. and Krugman, T. (1988) "Natural selection of allozyme polymorphisms: a microgeographical differentiation by edaphic, topographical, and temporall factors in wild emmer wheat (*Triticum dicoccoides*)" *Theo. Appl. Genet*. 76:737-752.

Ogihara, Y. and Ohsawa, T. (2002) "Molecular analysis of the complete set of length mutations found in the plastomes of *Triticum-Aegilops* species" in *Genome* 45: 956-962.

Ogihara, Y, Terachi, T, Sasakuma, T. (1991) "Molecular analysis of hot spot region related to length mutations in wheat chloroplast DNAs. I. Nucleotide divergence of genes and intergenic spacer regions located in the hot spot region" in *Genetics* 129:873-884.

Ovchinnikov, I.V. and Goodwin, W. (2003) "Ancient human DNA from Sungir?" in *Journal of Human Evolution*, 44: 389-392.

Özkan, H., Brandolini, A., Schafer-Pregl, R., Salamini, F. (2002) "AFLP Analysis of the collection of tetraploid wheats indicates the origin of emmer and hard wheat domestication in southeast Turkey" in *Mol. Biol. Evol.* 19(10): 1797-1801.

Pääbo, S. (1985) "Preservation of DNA in ancient Egyptian mummies" *J. Arch. Sci.* 12:411-417.

Pääbo, S. (1989) "Ancient DNA: extraction, characterisation, molecular cloning and enzymatic amplification" *Proc. Natl. Acad. Sci. USA*. 86:1939-1943.

Pääbo, S., Irwin, D.M. and Wilson, A.C. (1990) "DNA damage promotes jumping between templates during enzymatic amplification" in *J. Biol. Chem.,* 265: 4718-4721.

Peng, J., Ronin, Y., Fahima, T., Röder, M.S., Li, Y., Nevo, E., Korol, A. (2003) "Domestication quantitative trait loci in Triticum dicoccoides, the progenitor of wheat" in *Proc. Natl. Acad. Sci.USA,* 100:2489-2494.

Poinar H.N., Hofreiter M., Spaulding W.G., Martin P.S., Stankiewicz B.A., Bland H., Evershed R.P., Possnert, G., and Pääbo S. (1998) "Molecular coproscopy: dung and diet of the extinct ground sloth Nothrotheriops shastensis" in *Science*, 281:402-406

Qualset, C.O. (2001) "Wheat evolution and genetics: Discovery and use of egnes in crop improvement" in Wheat Taxonomy: The Legacy of John Percival, *The Linnean Special Issue 3*, The Linnean Society of London, Academic Press, London, GB pp.37-59.

Rieseberg, L.H. (2001) "Polyploid evolution: Keeping the peace at genomic reunions" in *Current Biology*, 11:R925–R928

Rollo, F., Ubaldi, M., Ermini, L., Marota, I. (2002) "Ötzi's last meals: DNA analysis of the intestinal content of the Neolithic glacier mummy from the Alps" in *Proc. Nat. Acad. Sci. USA,* 99: 12594-12599.

Rollo, F., Venanzi, F.M. and Amici, A.(1991) "Nucleic acids in mummified plant seeds: biochemistry and molecular genetics of pre-Columbian maize" *Genet. Res., Camb*. 58:193-201.

Rudnoy, S., Bratek, Z., Paldi, E., Racz, I., Lasztity, D. (2002) "Chloroplast 16S rRNA sequences from different Triticum species" in *Acta Biologica Szegediensis,* 46(3-4):47-48.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) "Molecular Cloning: A Laboratory Manual", Cold Spring Harbour Press, Cold Spring Harbour.

Schlumbaum, A., Jacomet, S. & Neuhaus J.M. (1998) "Coexistence of Tetraploid and Hexaploid Naked Wheat in a Neolithic Lake Dwelling of Central Europe: Evidence from Morphology and Ancient DNA" *J. Arch. Sci.* 25:1111-1118.

Sharma, H.C. and Waines, J.G. (1980) "Inheritance of tough rachis in crosses of Triticum monococcum and T. boeoticum" in *J. Hered.* 7: 214-216.

Smith, B.D. (1995) *The Emergence of Agriculture*, Scientific American Library, New York, USA.

Swofford, D. L. (1991) “PAUP: Phylogenetic Analysis Using Parsimony, Version 3.1. Computer program distributed by the Illinois Natural History Survey, Champaign, Illinois.

Talbert, LE, Blake, NK, Storlie, EW, Lavin, M. (1995) “Variability in wheat based on low-copy DNA sequence comparisons” in *Genome* 38: 951-7.

Threadgold, J. and Brown, T.A. (2003) “Degradation of DNA in artificially charred wheat seeds” in *Journal of Archaeological Science*, 30: 1067-1076.

Valkoun, J. (2001) “The geographical distribution of wild wheats in their historical setting and current context” in Wheat Taxonomy: The Legacy of John Percival, *The Linnean Special Issue 3*, The Linnean Society of London, Academic Press, London, GB pp.81-94

Van Slageren, M.W. (1994) “Wild Wheats: A monograph of *Aegilops L.* and *Amblyopyrum* (Jaub. & Spach) Eig (Poaceae)” Veenman Drukkers, Wageningen, Netherlands.

Van Zeist, W. (1992) “The Origin and Development of Plant Cultivation in the Near East” in *Japan Review*, 3: 149-163.

Waines, J.G. and Payne, P.I. (1987) “Electrophoretic Analysis of High-molecular-weight Glutenin Subunits of *T. monococcum*, *T. urartu*, and the A genome of Bread Wheat”, *Theo. Appl. Genet*., 74:71-76.

Wang, G.Z., Miyashita, N.T., Tsunewaki, K. (1997) **“**Plasmon analyses of *Triticum* (wheat) and *Aegilops*: PCR-single-strand conformational polymorphism (PCR-SSCP) analyses of organellar DNAs” in *Proc. Natl. Acad. Sci. USA, .* 94: 14570-14577

Watanabe, N., Sugiyama, K., Yamagishi, Y., Sakata, Y. (2002) ”Comparative telosomic mapping of homoeologous genes for brittle rachis in tetraploid and hexaploid wheats” in *Hereditas* 137(3): 180-185.

Watson, L. and Dallwitz, M. J. (1992 onwards). “Grass Genera of the World: Descriptions, Illustrations, Identification, and Information Retrieval; including Synonyms, Morphology,

Anatomy, Physiology, Phytochemistry, Cytology, Classification, Pathogens, World and Local Distribution, and References." Version: 18th August 1999.

Wendel, J. F. (2000) "Genome evolution in polyploids", *Plant Molecular Biology,* 42: 225–249

Wheat Taxonomy, http://www.ksu.edu/wgrc/Taxonomy/taxintro.html

Willerslev, E., Hansen, A. J., Binladen, J., Brand, T.B., Gilbert, M.T.P., Shapiro, B., Bunce, M., Wiuf, C., Gilichinsky, D.A., Cooper, A. (2003) "Diverse Pleistone and Animal Genetic Records form Holocene and Pleistocene Sediments" in *Science*, science.1084114.

Wolfe, K. H., Li, W-H., Sharp, P.M. (1987) "Rates of Nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs" in *Proc. Natl. Acad. Sci. USA,* 84: 9054-9058.

Wolfe, KH, Gouy, M, Yang, YW, Sharp, PM, Li WH. (1989) "Date of the monocot-dicot divergence estimated from chloroplast DNA sequence data" in *Proc Natl Acad Sci. USA,* 86(16): 6201-5.

Yıldırım, F. (2001) "AFLP Markers in Molecular Characterization and Genetic Relationship Analysis of *Triticum* ssp. in Turkey and Bulk Segregant Analysis for Determination of Linked Yellow Rust Resistance Marker", MSc. Thesis, Middle East Technical University, Ankara, Turkey.

Yoder, A.D. and Delefosse, T. (2002) "The Rise and Fall and Rise of Ancient DNA Studies" in *Ancient DNA*, McGraw-Hill, Yearbook of Science and Technology, pp. 9-14.

Zhang, W., Qu, J., Gu, H., Gao, W., Liu, M., Chen, J., Chen, Z. (2002) "Studies on the origin and evolution of tetraploid wheats based on the internal transcribed spacer (ITS) sequences of nuclear ribosomal DNA" in *Theo. Appl. Genet*., 104 (6-7): 1099-1106.

Zohary, D. (1996) "The mode of domestication of the founder crops of Southwest Asian agriculture" in D. Harris (ed.), *The origins and spread of agriculture and pastoralism in Eurasia,* Smithsonian Institution Press, Washington D.C. pp. 142-158.

Zohary, D. (1999) “Monophyletic vs. polyphyletic origin of the crops on which agriculture was founded in the Near East” in *Genetic Resources andCrop Evolution*, 46: 133-142.

Zohary, D. and M. Hopf. (1994) “The origin and spread of cultivated plants in West Asia, Europe, and the Nile valley.” *Domestication of plants in the Old World* 2nd ed. Oxford Univ. Press, New York.