Molecular Phylogeny of Tetraploid *Elymus* Species

Revealed by the DMC1 Gene

Xiaodi Zhang

A Thesis Submitted to Saint Mary's University, Halifax, Nova Scotia
in Partial Fulfillment of the Requirements for
the Degree of Master of Science in Applied Science

August, 2009, Halifax, Nova Scotia, Canada

Approved: Dr. Genlou Sun,
Supervisor

Approved: Dr. David Cone,
Supervisor committee

Approved: Dr. Andrew MacRae,
Supervisor committee

Approved: Dr. Rodger C. Evans,
External examiner

Date: July. 2009

Library and Archives Canada

Published Heritage Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et Archives Canada

Direction du Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-58309-8
Our file Notre référence
ISBN: 978-0-494-58309-8

Certification

Molecular Phylogeny of Tetraploid *Elymus* Species

Revealed by the DMC1 Gene

By Xiaodi Zhang

A Thesis Submitted to Saint Mary's University, Halifax, Nova Scotia,

in Partial Fulfillment of the Requirements for the

Degree of Master of Science in Applied Science

October 9, 2009, Halifax, Nova Scotia

Examining Committee:

Approved:	Dr. Rodger C. Evans Department of Biology, Acadia University
Approved:	Dr. Genlou Sun, Senior Supervisor Department of Biology
Approved:	Dr. David Cone, Supervisory Committee Member Department of Biology
Approved:	Dr. Andrew MacRae, Supervisory Committee Member Department of Geology
Approved:	Dr. Susan Bjornson, Program Representative
Approved:	Dr. Pawan Lingras, Graduate Studies Representative

Abstract

Molecular Phylogeny of Tetraploid *Elymus* Species

Revealed by the DMC1 Gene

By Xiaodi Zhang

Abstract: Although the relationship of genera of Triticeae has been analyzed in many previous studies, polyploid genera have not been as extensively studied as diploid ones within a phylogenetic framework. The present study focuses on tetraploidy of Triticeae *Elymus* species, which *are* all polyploid and distributed worldwide. Five basic genomes (St, H, Y, P, and W) in different combinations have so far been found in the *Elymus*. To estimate the phylogenetic relationship of tetraploid *Elymus* with genome content of StH or StY, data from nuclear single copy gene DMC1 (disrupted meiotic cDNA) of 39 tetraploid *Elymus* accessions and their potential genome donor diploid species were used. For StH species involved, the result revealed close relationship of *Pseudoroegneria* and *Hordeum* to *Elymus* StH species by multiple origins and recurrent hybridization. For StY species involved, it is suggested that St and Y genomes may be related on this locus. A direct contribution from P genome to StY genome through introgression still remains a possibility.

June 20, 2009

Acknowledgements

I would like to express the deepest appreciation to my supervisor Dr. Genlou Sun. This thesis would not have been possible without his guidance, encouragement and financial support

I would like to thank my committee members, Dr. David Cone and Dr. Andrew MacRae for their supportive suggestions and comments in this project.

I thank other faculty and staff members of the Biology Department at Saint Mary's University for their help: Dr. Jeremy Lundholm, Jing Yang, Carmen Cranley, Heidi de Boer and Janet White. I also would like to express my gratitude to Stanley King for his friendly help with my study and life.

I thank my family and friends for their supports and concerns, especially my parents, Fei Han and Lin Ling.

Table of Contents

PART 1. Molecular phylogeny of tetraploid *Elymus* species 1

1. Introduction 1

1.1. Polyploidy 1

1.2. *Elymus* and its genome contents 3

1.2.1. Poaceae, Triticeae, *Elymus* 3

1.2.2. Genome contents of *Elymus* 5

1.3. Molecular Phylogenetics 5

1.3.1. Chloroplast DNA 6

1.3.2. Nuclear DNA 7

1.4. Previous studies of tetrapolyploid *Elymus* Species 9

1.4.1. StH and StY species of *Elymus* and cytogenetic studies 9

1.4.2. Molecular phylogenetic analysis of StH and StY *Elymus* species 11

1.5. The DMC1 gene 12

1.6. The purpose of this study 13

2. Materials and Methodology 14

2.1. Materials 14

2.2. Methods 18

2.2.1. DNA isolation 18

2.2.2. DNA amplification 19

2.2.3. Cloning and sequencing 21

2.2.4. Rapid screening for plasmid inserts 21

2.2.5. Data analysis 22

3. Results 24

4. Discussion 41

4.1. Origin of StH species 41

4.2. Genome content and classification of *Hystrix* 43

4.3. The relationship of the Y genome to other genomes 45

4.4. Utility of DMC1 in phylogenetic studies of tetraploid *Elymus* 47

4.5. Summary 49

PART 2. Phylogenetic relationship between *Hordelymus* and *Hordeum* 51

1.Introduction.....51

2. Materials and methods 53

2.1. Plant materials.....53

2.2. DNA amplification and sequencing.....58

2.3. Data analysis 58

3. Results.....58

3.1. Sequence variation 58

3.2. Phylogenetic analysis of the *RPB2* sequences 62

4. Discussion 66

4.1.The relationship of *Hordeum* and *Hordelymus* genome 66

4.2. *Hordelymus* and Ta (*Taeniatherum*) , Ns (*Psathyrostachys*), the unknown genome 67

4.3.Summary 70

PART 3. References 71

Content of Tables

Table 1: Taxa used in this study (Part 1) 15

Table 2: PCR recipe and protocol (Part 1) 20

Table 3. Taxa used in this study (Part 2) 54

Content of Figures

Figure 1: Example of amplified products (Part 1) 25

Figure 2: Consensus tree in the analysis of the DMC1 data set for 66 sequences using maximum parsimony method . (Part 1) 29

Figure 3: Partial complete alignment using software Clustal X (Part 1). 34

Figure 4: Strict consensus tree derived from DMC1 sequence data of 45 sequences using maximum parsimony method (Part 1) 37

Figure 5. Strict consensus tree constructed using Neighbor Joining Method (Part 1). 39

Figure 6. Example of partial alignment of the amplified sequences of *RPB*2 from *Hordelymus* and other genomes in Triticeae (Part 2). 61

Figure 7. Strict consensus tree of the 456 most parsimonious trees derived from *RPB2* sequence data using maximum parsimony method (Part 2) 64

PART 1. Molecular phylogeny of tetraploid *Elymus* species

1. Introduction

The estimation of relationships within the hierarchy of individuals, populations and species is of fundamental importance in understanding evolutionary patterns and processes. The amount and dynamics of genetic variability has much influence on the evolution of a species. At higher taxonomic levels understanding the relationships among species can provide a retrospective window of how they evolved. The study of relationships among plants has been revolutionized in recent years by advances in molecular genetic techniques (Peter *et al.*, 1999). Plant systematics combines molecular analysis and traditional taxonomy to understand the grouping of plants and the evolutionary process by examining their phylogenetic relationships at all levels.

Elymus has been a model in plant systematic research not only because it is of enormous economical importance all over the world, but also because it has undergone a complicated evolutionary process including events such as hybridization, polyploidization and introgression *et cetera*. A better understanding of the evolution and systematics of *Elymus* will contribute significantly to the understanding of the evolutionary framework of plants.

1.1. Polyploidy

Polyploidy, the presence of more than two genomes per cell (Soltis and Soltis, 2000a), is an important evolutionary force and specialization process. Distinguishable types of

polyploids are defined using Stebbins' classification (Stebbins, 1950). Allopolyploids are defined as those polyploids that have been generated through processes of interspecific hybridization and chromosome doubling; autopolyploids are those polyploids that have come from conspecific parents. In other words, allopolyploidy is the combination of two or more differentiated genomes and autopolyploidy is the duplication of a single but complete genome (Masterson, 1994; Soltis and Soltis, 1999, 2000b).

Recent estimates suggest that polyploidization is a common phenomenon, especially in plants. Seventy percent of all angiosperms have experienced one or more episodes of polyploidization (Masterson, 1994; Wendel, 2000). The level of polyploidy may be even higher in ferns, possibly as high as 95% (Grant, 1981). Many economically important domesticated plants such as wheat, potatoes and cotton are polyploid. It has also occurred in many insects, fish, and amphibians (Mangus and Bengt, 2007).

Polyploids often occupy different types of habitats, which can be different from those of their diploid parents, and have been proposed to be superior colonizers to diploids, showing a higher level of adaptability (Stebbins, 1950). Many aspects of the genetic systems of polyploids may contribute to the success of polyploid plants. These characteristics range from the molecular level to the population level and include increased heterozygosity, reduced inbreeding depression and an associated increase in selfing rates, increased genetic diversity through multiple formations of a polyploid species, genome rearrangements, and ancient polyploidy and gene silencing (Soltis and Soltis, 2000a; Hedrick, 1987; Wolf *et al.*, 1989; Segraves *et al.*, 1999).

In the past fifty years, many questions about plant polyploidy have not been answered, and many new questions have arisen, such as factors involved in the origin and establishment of polyploids in nature. More research related to polyploidy is in great demand and the application of molecular techniques could provide critical insights into the understanding of parentage, formation, and immediate genetic consequences of polyploidy.

1.2. *Elymus* and its genome contents

1.2.1. Poaceae, Triticeae, *Elymus*

In many studies of evolution, phylogeny, or taxonomy, tribes of the grass family Poaceae are often utilized as research material because they include a number of plants of major economic importance and they have a complicated evolution process worldwide.

Triticeae, one tribe of Poaceae, includes major crop plants such as wheat, barley, rye and several hundred other related species of grasses. Some of the genera within this tribe are commonly used in the production of grains. While most crop plants are annuals, Triticeae contains a majority of perennials, such as species in *Elymus*, *Pseudoroegneria* and *Leymus,* as well as the partially perennial genus *Hordeum* which includes both perennial and annuals. Hitchcock's (1951) treatment of the tribe Triticeae had been widely accepted in North America for many years. Nowadays, the most commonly used is Löve's treatment (1984) and this was adopted for the present study.

The genus *Elymus* is the largest and most widely distributed genus in Triticeae. *Elymus* comprises approximately 150 polyploid species worldwide on most continents

according to the classifications of Löve (1984) and Dewey (1984). All *Elymus* species are polyploid and share close relationships with other genera within Triticeae (Klimyuk *et al.*, 1997). As a well established group, they have successfully adapted to different arctic, subtropical and temperate environments, although *Elymus* is traditionally viewed as a northern temperate genus. These species inhabit various ecological niches, including grasslands, semi-desert areas, among bushes in mountain slopes and valleys (Löve, 1984; Dewey, 1984; Lu, 1993b). More than half of all known *Elymus* species (approximately 80) originated from Asia, which is regarded as a center of high diversity. North America is the other prominent habitat of *Elymus* species, where about 50 species are found with low barriers to interspecific hybridization (Mason-Gamer *et al.*, 2002; Zhang *et al.*, 2000).

Elymus is of significant economic importance and many *Elymus* species have been used as forage grasses or introduced to new areas for potential forage usage, for example, *E. trachycaulus* and *E. sibiricus*; some species have been used in breeding by interspecific or intergenetic hybridization like *E. canadensis*. Some valuable genes for crop improvement are present in certain *Elymus* species, such as resistance to wheat greenbug, wheat streak mosaic virus, and barley yellow dwarf virus (Lu, 1993b).

Due to its worldwide distribution and great economic value, it is necessary to gain a better understanding of the classification of *Elymus* and its mode of polyploidization and speciation.

1.2.2. Genome contents of *Elymus*

Cytogenetic studies of genome-pairing data have profound effects on the classification of the Triticeae (Dewey, 1984; Löve, 1984). In cytogenetic analyses, the behavior of chromosome pairing at meiosis can indicate the degree of similarity among genomes. In the Triticeae, cytogenetic data have been widely used in the delimitation of genera, in the determination of the origins of polyploids, and have had enormous influence on classification within the tribe (Mason-Gamer *et al.*, 2002). Data given by meiotic studies, together with other sources of information such as morphology, geographic distribution and isozyme variation, can facilitate a better understanding of phylogenetic relationships in a group of organisms (Lu *et al.*, 1992).

Five basic genomes (St, H, Y, P, and W) in different combinations have so far been found in the genus. The St genome was supposedly donated by *Pseudoroegneria*. The H genome is very common in North American, South American and European *Elymus* species and was donated by the genus *Hordeum*. The P and W genomes are thought to be derived from *Agropyron* (Gaertn) and *Australopyrum* (Tzvelev) A Löve, respectively (Jensen, 1990a; Torabinejad *et al.*, 1993). Presently, the origin of the Y genome remains unknown, although it is a common *Elymus* genome in Central and East Asia (Dewey, 1984; Liu *et al.*, 2007).

1.3. Molecular Phylogenetics

During the past 10–15 years, molecular phylogenetics has served to dramatically shape our views of organismal relationships and evolution. This impact has been felt at all

taxonomic levels of the hierarchy of life. Advances in molecular biology and their application have provided us with the opportunity to explore the rate and patterns of genetic variation within plant species.

Within the last few years, more sophisticated molecular markers have been discovered and applied in plant systematics and evolutionary research, such as chloroplast and nuclear DNA. An overwhelming majority of plant phylogenies are based on chloroplast DNA and nuclear ribosomal DNA. These organelle or multicopy nuclear sequences have many desirable qualities and are relatively easier to access, but there are noticeable disadvantages and single copy nuclear genes are selected to compensate for those disadvantages (described on following pages).

1.3.1. Chloroplast DNA

The most widely used source of data in plant molecular systematic analyses is chloroplast DNA (cpDNA). As a molecular tool, the primary advantage of cpDNA is its relatively simple inheritance. The characteristics of cpDNA include haploid and maternal inheritance. Also, a high number of copies of a particular gene within cpDNA guarantees successful restriction site analysis as well as amplification. Ironically, however, the advantage of cpDNA in being haploid and maternally inherited makes it difficult for use in the analysis of hybrid plants or those of polyploid origin; as cpDNA is uniparentally inherited and haploid, it reveals only half of the parentage in hybrid or polyploid plants (Olmstead *et al.*, 1994; Soltis and Soltis, 1998; Randall *et al.*, 2004).

1.3.2. Nuclear DNA

Another widely used gene sequence in plant systematics is nuclear DNA, which is biparentally inherited. The advantages of nuclear genes are as follows: Firstly, nuclear genes have faster evolution rates than organellar genomes. It has been revealed that synonymous substitution rates of nuclear genes are up to five times greater than those of chloroplast genes and 20 times greater than those of mitochondrial genes (Wolfe *et al.*, 1987; Gaut 1998). Nuclear genes have experienced independent evolution events such as chromosomal interaction, hybridization, introgression and non-homologous gene conversion, *et cetera*; thus, data from multiple independent regions makes it possible and practical to obtain a more accurate picture of phylogenetic relationships. This provides better efficiency of sequencing analysis effort due to the presence of more detected variations. Secondly, nuclear genes are divided into several regions with different functions, such as 5' untranslated regions (5' UTRs), exons, introns and 3' UTRs and all these regions are capable of being utilized for research. That is to say, the presence of a large number of multiple independent loci may be used for phylogenetic inference. Confidence in a given phylogenetic tree can be increased by the corroboration of independent dataset analysis (Small *et al.*, 1998; Liu *et al.*, 2001; Cronn *et al.*, 2003). The high copy number for most nuclear genes (excepting low/single copy genes) facilitates manipulation of DNA by both restriction site and PCR-based analysis. Due to all these advantages, comparative phylogenies between nuclear and chloroplast/ mitochondrial sequences have become a powerful tool to identify the mode of polyploidization in some groups (Rauscher *et al.*, 2004) such as rice (Ge *et al.*, 1999), *Elymus* (Mason-Gamer *et al.*,

2001) and *Silene* (Mangus and Bengt, 2001). The greatest disadvantage of using of nuclear genes in phylogeny studies is that the presence, absence or variable strength of concerted evolution still remains as a result of the presence of both multiple copies per array and multiple arrays per genome. Concerted evolution is the tendency of the different genes in a gene family or cluster to evolve in concert. In other words, each gene locus in the a comes to have influence to others in the same group (Small *et al.*, 2004).

Among nuclear genes, nuclear ribosomal DNA (rDNA) is the most widely used nuclear gene in phylogeny studies thus far (Kellogg *et al.*, 1995; Hsiao *et al.*, 1995; Zhang *et al.*, 2008; Liu *et al.*, 2006). However, due to the presence of concerted evolution, the rDNA sequences obtained by sequencing of amplification products may fail to reveal the complexity of nuclear rDNA content, and may in fact preferentially reveal paralogous rDNA sequences in different taxa or accessions (Bailey *et al.*, 1999; Hartmann *et al.*, 2001; Small *et al.*, 2004; Suh *et al.*, 1993). This problem is exacerbated in allopolyploid and hybrid taxa in which multiple divergent rDNA loci are expected to exist on different donors and ought to be defined.

Low/single copy nuclear sequences are able to compensate for these limitations. The most desirable property of low copy sequences apart from other kinds of nuclear genes is their low frequency of concerted evolution. This characteristic makes low copy nuclear genes ideal candidates for identifying donors of hybrids or polyploids (Sang, 2002). Low copy nuclear genes have been used as molecular markers to identify the origins of hybrid or polyploid taxa, such as *Glycine* (Doyle *et al.*, 1999), *Clarkia* (Ford *et al.*, 1999), *Oryza* (Ge *et al.*, 1999) and *Elymus* (Mason-Gamer *et al.*, 2001).

The disadvantages and possible difficulties of using nuclear genomes are in isolating and identifying orthologous genes, which are representatives of speciation events. Other relevant issues include the possibilities of concerted evolution and/or recombination among paralogous sequences and the presence of intraspecific, intrapopulational and intraindividual variation (Soltis and Soltis, 1993, 1998). However, low/single copy nuclear genes are still the most convincing and suitable molecular marker nowadays, especially for studies on polyploidy.

1.4. Previous studies of tetrapolyploid *Elymus* Species

As mentioned above, *Elymus*, one of the most used model plants in taxonomy and phylogenetics, is a genus with polyploid species. Much research has been done to explore the origin of this polyploidy and its relationship to other genera. Polyploidy in *Elymus* includes tetrapolyploidy (StH, StY, StP) and hexaploidy (StHP, StYP, StStH, StHY, *et cetera.*) (Dubcovsky *et al.*, 1997). So far, tetrapolyploid species with StH and StY genome content are the most studied.

1.4.1. StH and StY species of *Elymus* and cytogenetic studies

StH and StY genome species are the most common tetraploids in *Elymus*. The StH genome species are found from the Arctic to the subtropics and from the seacoast to above 5000m in the Himalayas and their morphology is diverse. This group consists of approximately 50 species and is well adapted to enormously divergent climates and habitats (Sun *et al.*, 2007). Cytogenetic studies suggest that this group is an

allopoplyploid arising from the polyploidization of *Hordeum* and *Pseudoroegneria.* The H genome donor is *Hordeum*, which is a genus of about 30 species of annual and perennial grasses, native throughout the temperate northern hemisphere, temperate South America, and also South Africa. The St genome donor is *Pseudoroegneria*, which contains approximately 20 diploid or tetraploid species found in central Asia, the Middle East, northern China and western North America (Watson *et al.*, 1992; Bothmer *et al.*, 1995). Genome-pairing data suggest that nearly all of the native North American species of *Elymus* are StH allotetraploids (Dewey, 1983*a*, 1983*b*, 1984; Mason-Gamer *et al.*, 2002).

Another important group of tetraploid *Elymus* is the StY genome group. About 30 *Elymus* species containing the StY genome are found restrictedly in temperate Asia, where more than half of the known *Elymus* species originated. These StY-genome species are classified into four different sections including Sect. *Elymus* L. (ca. five species), Sect. *Goulardia* (Husnot) Tzvelev (ca. twenty-four species), Sect. *Cline-lymiopsis* (Nevski) Tzvelev (one species), and Sect. *AnthoJachne* (Steud.) Tzvelev (ca. ten species). This classification reflects the large morphological variation of the group (Lu *et al.*, 1992). A large dataset from cytological analyses of artificial hybrids among the StY genome *Elymus* species indicates that the degree of chromosome pairing in the hybrids gradually decreases with increase in geographical distance from the locality of their parent. That is to say, StY genome tetraploid species have been modified to a large extent and exhibit relatively high genetic diversity (Lu 1993b, 1994; Liu *et al.*, 2006).This phenomenon has not been found in the StH genomes, which show relatively

high homology among different tetraploid *Elymus* species distributed in Asia (Lu *et al.*, 1992).

1.4.2. Molecular phylogenetic analysis of StH and StY *Elymus* species

While molecular phylogenetics has greatly deepened the understanding of the evolution of the diploid members of the tribe Triticeae, studies addressing the condition of polyploidy, including that of the *Elymus* species, are limited in number. Cytogenetic techniques were used extensively in polyploid *Elymus* to clarify their origin and relationship; as a result, currently most hypotheses about the origins of the tetrapolyploid taxa are based on morphology and patterns of genome pairing (Dewey, 1984; Löve, 1984; Lu, 1993a, 1993b).

Chloroplast and nuclear genes have been used to reveal the evolutionary history and origin of *Elymus* species. Data sets from chloroplast DNA (Mason-Gamer *et al.*, 2002; McMillan *et al.*, 2004; Xu *et al.*, 2004) agreed that *Pseudoroegneria* is the maternal genome donor of *Elymus.* Nuclear gene data sets from granule-bound starch synthase I (GBSSI) (Mason-Gamer *et al.*, 2001), the phosphoenolpyruvate carboxylase gene (Helfgott *et al.*, 2004,) and the RPB2 gene (Sun *et al.*, 2008) were in agreement that *Pseudoroegneria* and *Hordeum* are genome donors of *Elymus.* As mentioned above, little is known about the phylogeny of the StY genome *Elymus* species. The relationship of the Y genome to other genomes still remains a puzzle despite many cytogenetic and molecular level attempts to clarify the connection. Analyses using nuclear ribosomal internal transcribed spacer and chloroplast *trnL-F* sequence datasets suggested that the St

and Y genomes may have the same origin (Liu *et al.*, 2006). However, results from the single copy gene RPB2 did not agree on this point and suggested that the Y genome originated from a distinct diploid species; W and P genomes may be sisters to the Y genome (Sun *et al.*, 2008).

There is a great demand to discover the relationship of the Y genome to the St genome and other potentially related genomes. Molecular phylogeny analyses of tetrapolyploid *Elymus* species and their related genera will provide opportunities for understanding their phylogenetic relationships in the evolutionary process. Compared to other molecular markers, such as cpDNA and ITS sequences, low/single copy genes may prove to be the most suitable molecular tools due to their great advantages. Few single copy nuclear genes have been applied in plant phylogenetics; in the phylogeny of *Elymus,* three single copy nuclear genes have been reported, including the genes for granule-bound starch synthase (Mason-Gamer, 2001), phosphoenolpyruvate carboxylase (Helfgott *et al.*, 2004), and the RPB2 gene (Sun *et al.*, 2007, 2008) but more information from another single copy gene is required.

1.5. The DMC1 gene

Disrupted meiotic cDNA (DMC1) is active during recombination-mediated homology-dependent chromosome pairing and in strand exchanges resulting in chiasmata (Klimyuk *et al.,* 1997). Recent evidence suggests that the DMC1 gene may also be expressed during mitosis in *Arabidopsis* (Doutriaux *et al.*, 1998).

The DMC1 gene has been discovered in fungi, animals and higher plants (Stassen *et al.*, 1997). Only a few DMC1 sequences are available from higher plants. Complete DNA sequences exist in *Arabidopsis* Schur (Klimyuk *et al.*, 1997) and *Hordeum* L. (Peterson and Seberg, 2003). In *Arabidopsis*, the gene consists of 15 exons, has a total length of 2790 bp and encodes 344 amino acids. The *Hordeum* sequence lacks intron 1 and has a total length of 3640 bp (Peterson and Seberg, 2002).

DMC1 is structurally and evolutionarily related to the *Escherichia coli recA* gene. A number of studies explored the relationship between the *recA*-like genes DMC1, *RAD51, radA*, and *radB*; DMC1 has been mapped to the top arm of chromosome 3 between the molecular markers m560B2 and g4711 (Stassen *et al.*, 1997; DiRuggiero *et al.*, 1999).

The single-copy nuclear gene DMC1 (disrupted meiotic cDNA) has been applied to a few phylogenetic analyses of Triticeae species, including diploid *Hordeum, Psathyrostachys* and tetrapolyploid *Hordelymus* species (Peterson and Seberg, 2004a, b, 2008).

1.6. The purpose of this study

Exploring the origins and phylogenetic relationships of tetraploidy is significant for conservation of biodiversity and utilization of the genetic resources in *Elymus*, which is well known for its economic importance, not only because many *Elymus* species have been used as forage grasses but because they possess valuable genes for breeding usage. Success in establishing the phylogenetic framework of *Elymus* species will extend the understanding of the evolution of tribe Triticeae.

Compared to limited previous studies on tetraploid species on a molecular level (Liu *et al.*, 2006; Mason-Gamer *et al.*, 2002; Xu *et al.*, 2004), the present study chose a single copy gene, DMC1, as the molecular tool for phylogenetic analysis; this is the first attempt to use DMC1 on large number of tetraploid Triticeae. The objectives of this study were: 1) to investigate the phylogenetic relationship between tetraploid species and their potential donor diploids; 2) to attempt to reveal genetic relationships of the Y genome to other genomes in Triticeae; 3) to clarify the debated classifications of a few species and, finally; 4) to test the utility of the DMC1 gene in the phylogenetic analysis of polyploid *Elymus* species.

2. Materials and Methodology

2.1. Materials

Thirty one tetraploid *Elymus* accessions and eight potential genome donor species were used in this study. The name, genomic constitutions and accession numbers are given in Table 1. DMC1 Sequences of a few Accessions directly from GenBank are shown using their GenBank accession numbers (most of them are diploid). Seeds were kindly provided by the Germplasm Resources Information Network (GRIN) of the United States Department of Agriculture (USDA) and the Swedish University of Agricultural Sciences. Seeds were germinated on absorbent paper in Petri dishes. Germinated seeds were transplanted to a sand–peat mixture in a greenhouse.

Table 1: **Taxa used in this study**

Species	Seeds Accession No.	Genome	GenBank Accession No
Elymus gayanus E. Desy	PI 13828	SH	
	PI 636675	SH	
Elymus multisetus (J. G. Sm.) Burtt Davy	PI 20963	SH	
Elymus lanceolatus (Scribn. & Smith) Gould	PI 236663	SH	
*Elymus virescens (*Piper) Gould	H10584	SH?	
Elymus violaceus (Hormem.) Feilberg	H10588	SH?	
Elymus cordilleranus	H6486	SH?	
Elymus trachycaulus (Link) Gould ex Shinners	PI 537323	SH	
Elymus confusus E.Desy	PI 598463	SH	
Elymus coreanus Honda	PI 531578	SH?	
	W614259	SH?	EU366411
Elymus komarovii (Roshev.) A. Löve	ZY3161	SH?	EU366414.1
Elymus barbicallus (Ohwi) S. L. Chen	PI 499380	SH	
Elymus scabrifolius (Doll) J. H. Hunz	PI 537310	SH	

Table 1 (Continued)

Species	Seeds Accession No.	Genome*	GenBank Accession No
Elymus caninus (L.)L	H3169	SH	
Elymus wawawaiensis J. Carlson ex Barkworth	PI 610984	SH	EU366410.1
Elymus dentatus (Hook. F.) Tzvelev	PI 628702	SH	
Elymus canadensis L	PI 531567	SH	EU366406
Elymus hystrix L	H5495	SH ?	
	PI 531616	SH?	EU366416 EU366415
Elymus fibrosus (Schrenk) Tzvelev	H10399	SH	
Elymus wiegandii Femald	PI 531708	SH	
Elymus semicostatus (Nees ex Steud.) Melderis	PI 207452	SY	
Elymus panormitanus (Parl.) Tzvelev	PI 254866	SY	
Elymus longearistatus (Nevski) Tzvelev	PI 401280	SY	
Elymus macrochaetus (Nevski) Tzvelev	PI 618796	SY	
Elymus gmelinii (Ledeb.) Tzvelev	PI 610898	SY	
Elymus caucasicus (K. Koch) Tzvelev	PI 531573	SY	
Elymus fedtschenkoi Tzvelev	PI 564927	SY	

Table 1 (Continued)

Species	Seeds Accession No.	Genome*	GenBank Accession No
Elymus strictus (Keng) A. Löve	PI 499476	SY	
	PI 634293	SY	
Pseudoroegneria spicata(Pursh) Tzvelev		St	AF277245
Agropyron cristatum(*Linneaus.*) Gaerth	PI 383534	P	
Hordeum capense		H	AY592986
Hordeum bogdanii		H	AY137412
Psathyrostachys juncea (Fish) Nevski		Ns	EU366427
Psathyrostachys fragilis (Boiss.) Nevski		Ns	AF277263
Bromus arvensis L			DQ247821
Bromus sterilis L			AF277234

Note: The genome designations are according to Wang *et al*., (1994).

2.2. Methods

2.2.1. DNA isolation

After being frozen in liquid nitrogen, young leaves were ground with a mortar and pestle. The powder was distributed to 2-mL tubes. Seven hundred and fifty microliter of lysis solution (50mM Tris–HCl, pH 7.6; 100mM NaCl; 50mM EDTA; 0.5% SDS) was added to each tube with 10mM β-mercaptoethanol. The mixture was incubated at room temperature for 10-15 min. Four hundred and fifty microliter of phenol-tris-chloroform (pH=7.5) was added; the different phases were then mixed gently and then separated by centrifugation at 13000 rpm for 5 min. The upper phase was transferred with a cut-off 5-mL tip into a new tube. Four hundred and fifty microliter chloroform-isoamyl alcohol (24:1) was added. Tubes were inverted and centrifuged at max speed for 5 minutes. The upper phase was transferred to a new 1.5-mL tube. The nucleic acids in the aqueous phase were precipitated by adding 0.6 volumes of cold isopropanol. The tube was incubated for 20 minutes at -20°C. The pellet was collected by centrifugation at 13000 rpm for 10 minutes and then washed with 70% cold ethanol and dried. The dry pellet was dissolved in 400μL of TE (10 mM Tris, pH 8.0; 1 mM EDTA, pH 8.0) solution and stored at room temperature for at least 20 hours before further use. 50μg/mL RNase was added. The tube was then incubated at 37°C for 30 minutes. In order to purify the raw genomic DNA that was extracted, a purifying procedure was performed. In this process, steps from above were repeated, from adding Phenol-Tris-Chloroform to resuspension in TE buffer (100μL TE buffer as final volume). The purity and concentration of the DNA was assessed spectrophotometrically by calculating the A260/A280 ratio to determine

protein impurities. The DNA yield was calculated from the A260 for clean DNA samples (A260/A280 between 1.8 and 2.0) (Junghans *et al.*, 1990; Sun *et al.*, 1997).

2.2.2. DNA amplification

Primers TDMC1e10F (5'-TGCCAATTGCTGAGAGATTTG-3') and TDMC1e15R (5'-AGCCACCTGTTGTAATCTGG-3') (Peterson and Seberg, 2000, 2002) were used for PCR amplification of the region spanning from exon 10 to exon 15 of the total DMC1 gene, numbered according to the *Arabidopsis* DMC1 gene structure (Klimyuk *et al.*, 1997; Peterson and Seberg, 2002). PCR amplification was performed under standard conditions. The recipe and protocol are shown in Table 2. DNA *Taq* polymerase and reaction buffer were from New England Biolabs (Pickering, ON). PCR products were checked by gel electrophoresis on 1.0% agarose gels, stained in ethidium bromide (1μg/mL) and visualized under UV light.

The amplicons were purified using the QIAquick PCR Purification Kit (Qiagen Inc, Mississauga, ON: Cat. No.28104) following manufacture's instruction.

Table 2: Mixture contents (per tube) for Polymerase Chain Reaction in 20µL volume; PCR recipe for screening colonies in 15µL volume and their protocol

Recipe

Reagent	DNA amplification	PCR for screening colonies	Final concentration
DNA template	4.00 µL	3.00 µL	1.5ng/ µL(30ng)
TDMC1e10F	2.00 µL	1.50 µL	0.2 µM
TDMC1e15R	2.00 µL	1.50 µL	0.2 µM
Reaction buffer	2.00 µL	1.50 µL	1X
dNTP	4.00 µL	3.00 µL	0.2mM
***Taq* polymerase**	0.15 µL	0.11 µL	0.75U
dH_2O	5.85 µL	4.39 µL	
Total volume	20 µL	15 µL	

Protocol

Denaturing at 95°C for 5 minters
Denaturing at 95°C for 1 minutes
Annealing at 55°C for 2 minutes
Extension at 72°C for 2 minutes } 40 cycles
Final extension at 72°C for 7 minutes
Incubation at 4°C

2.2.3. Cloning and sequencing

The purified PCR product was cloned into a pGEM-easy T vector (Promega Corporation, Madison, WI, USA). The ligation reaction was set up as follows: 5μL 2X Rapid Ligation Buffer; 1μL pGEM®-T Easy Vector (50ng); 3μL PCR product; 1μL T4 DNA Ligase (3 Weiss units/μL). The reaction was mixed by pipetting and incubated at room temperature for 1 hour. Two microliter of each ligation reaction was then transferred to a 1.5-mL microcentrifuge tube with 50μL JM109 High Efficiency Competent Cells (Promega Corporation, Madison, WI, USA). The mixture was incubated on ice for 20 minutes. The cell mixture was heat-shocked for 45–50s in a water bath at exactly 42°C without shaking and immediately transferred to ice for 2 minutes. Nine hundred and fifty microliter SOC medium was added to each tube at room temperature and incubated for 1.5 hours at 37°C with shaking (~150rpm). One hundred microliter of each transformation culture was sprayed onto duplicate LB/Ampicillin/ IPTG/X-Gal plates (100μL of 100mM IPTG and 20μL of 50mg/mL X-Gal may be spread over the surface of an LB-Ampicillin plate and allowed to absorb for 30 minutes at 37°C prior to use). The plates were incubated at 37°C overnight (16–24 hours).

2.2.4. Rapid screening for plasmid inserts

Six to ten white colonies were randomLy selected and each was transferred to 50μL of LB broth with 0.1mg/mL antibiotics by picking the colony with either a sterile loop or pipette and twirling that in the liquid in the tube. These solutions were incubated in well plates at room temperature for 40 minutes or at 37°C for 20 minutes. Afterwards, fifteen

microliter of colony PCR reaction were set to check the presence of the insert using the TDMC1e10F/ TDMC1e15R primers (the recipe is shown in table 2). The protocol was the same as shown in Figure 1. If the PCR product was confirmed to be amplified successfully by TDMC1e10F/TDMC1e15R, the clone was confirmed to be recombinant and 8-10μL of the remaining solution was transferred to 5 mL LB broth and incubated at 37°C overnight with antibiotics (Ampicillin) at 0.1mg/mL.

Plasmid DNA was isolated using the Sigma GenElute™ Plasmid Miniprep Kit (Oakville, Ontario, Canada) according to the manufacturer's instructions. Plasmid DNA was sent to Macrogen Inc. (Seoul, Korea) and sequenced commercially using primers pair SP6 and T7. To eliminate *Taq* or sequencing error and provide enough candidates for two copies of DMC1, five to eight independent clones of the same species were sequenced.

2.2.5. Data analysis

Automated sequence outputs were visually inspected with chromatographs. The sequences of both strands from the same species were confirmed and compared to each other using BLASTn (http://blast.ncbi.nlm.nih.gov/Blast.cgi) on the NCBI website. Segments of nucleic acid sequence that may be of vector origin were checked and removed in every sequence by a public sequence analysis tool on the NCBI (National Center for Biotechnology Information of USA) website called VecScreen (http://www.ncbi.nlm.nih.gov/VecScreen/VecScreen.htmL). Full sequences were joined by comparing sequencing results from two different extension sequencing processes

using an online sequence analyzer (All-in-One Seq Analyzer) (Naohiro, 1998). The following steps were applied only to the remaining sequences.

Multiple sequence alignments were performed using ClustalX (Thompson *et al.*, 1997). Phylogenetic analysis using the maximum parsimony (MP) method was performed with the computer program PAUP* version 4 beta 10 Win (Swofford *et al.*, 2003). The most parsimonious trees were obtained by performing a heuristic search using the Tree Bisection-Reconnection (TBR) option with MulTrees on, and 10 replications of random addition sequence with the stepwise addition option. All characters were specified as unweighted. Multiple parsimonious trees were combined to form a strict consensus tree. Overall character congruence was estimated by the consistency index (CI), rescaled consistency index (RCI) and then the retention index (RI). In order to infer the robustness of clades, bootstrap values with 1,000 replications (Felsenstein, 1985) were calculated by performing a heuristic search using the TBR option with MulTrees off. To complement the parsimonious analysis, the neighbor-joining (NJ) method using the Tajima-Nei distance (Saitou *et al.*, 1987) was also used for phylogenetic reconstruction.

All the digital results were kept in Dr. Sun's lab in Saint Mary's University. For each species, one consensus sequence will be sent to GeneBank after all the work in the end.

3. Results

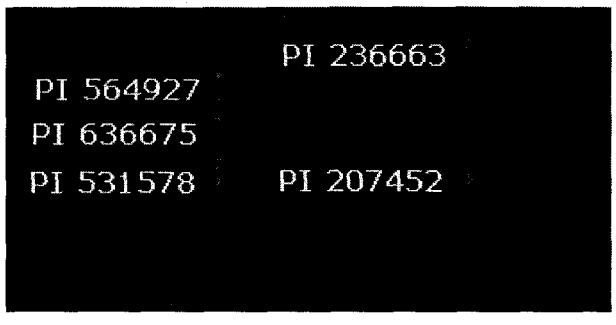
For amplification of each accession in the present study, a single bright band was obtained in electrophoresis. No visible differences such as length variation or multiple bands were observed (Figure 1A).

After purification, for each accession, the amplification product was cloned and eight to ten white colonies were randomly picked and transferred. Colony PCR reaction was set to check the presence of the recombinant. It turned out that some colonies were recombinant, but others were questionable as no DMC1 amplicon was detected (Figure 1B). For some accessions, amplification and cloning was repeated due to a lack of confirmed recombinant. For twenty six tetraploid species and one diploid species which were worked with on the bench, more than 400 colonies from cloning products were screened in total.

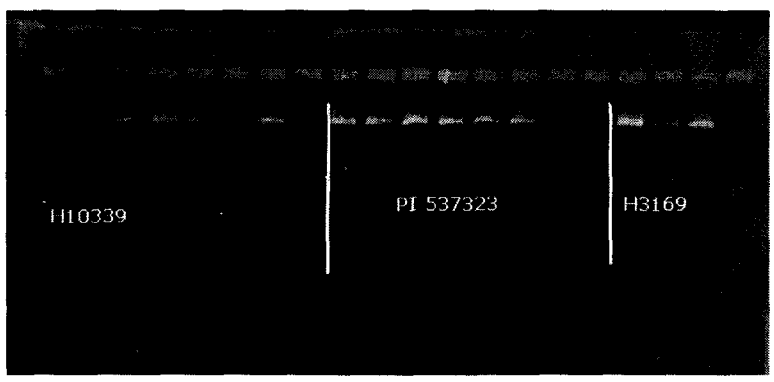
Plasmid DNA was extracted for all recombinants for each accession. The quality of plasmid DNA extraction was checked by electrophoresis. Nearly 180 clones were sent for sequencing by single extension.

Sequencing quality was checked by visually examining peaks in images of results. Within each sequence, if peaks were continuously overlapped or unclear, the sequence was excluded for further analysis. Approximately 160 original sequences were accepted for continued analysis. The length of the amplified DMC1 sequences varied from 995bp to 1225bp in all accessions.

Figure 1: Example of amplified products using the TDMC1e10F/ TDMC1e15R primers. A) DMC1 amplified product examples; single bright bands were obtained. B) Examples of colony PCR screening of inserted plasmid. The PCR was set to check for the presence of the insert; recombinants were confirmed by the presence of their PCR products.


PI 236663
PI 564927
PI 636675
PI 531578
PI 207452

1A


H10339
PI 537323
H3169

1B

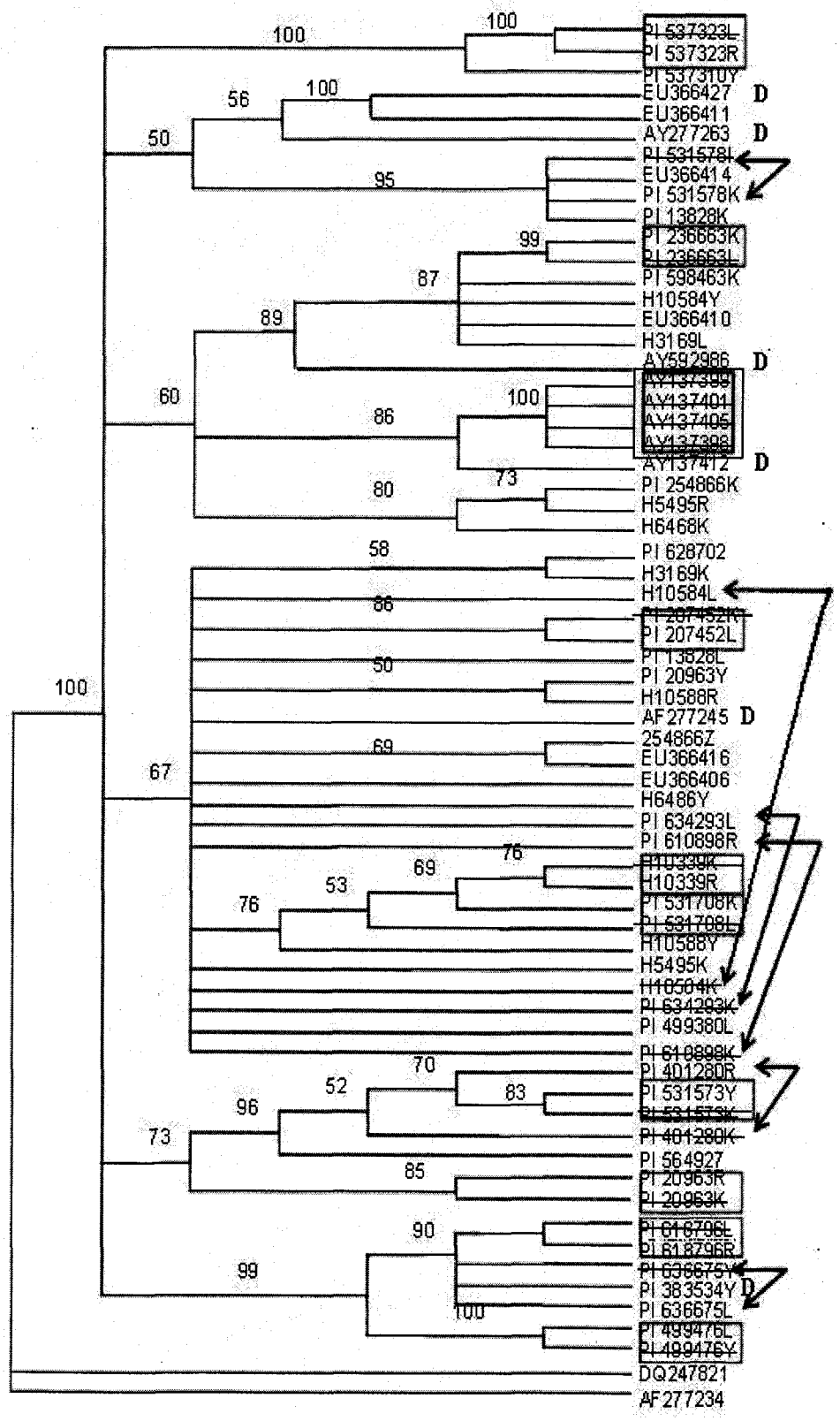
For each accession, distinct sequence groups were organized and majority rule consensus sequences were constructed. If the identity of different sequences from the same accession was equal to or higher than 98%, monophyletic groups of clones obtained from a single accession differing by autapomorphic substitutions only were regarded as a single sequence group; autapomorphic substitutions were interpreted as polymerase errors. If a subset of two or more clones from a single accession had one or more parsimony informative substitutions in common (thus distinguishing them from other clones from the same accession), they were regarded as a single, separate sequence group (Mangus and Bengt, 2007). For each sequence group, sequencing was performed in the reverse direction for those with highest sequence quality and identity in the first round. Consensus full sequences were constructed by hand within each group and taken as distinct gene copy candidates for every accession. Among 26 tetrapolypoid *Elymus* accessions involved, eleven accessions possessed only one copy of the DMC1 gene; the other fifteen accessions were interpreted to have potentially different genetic copies. A phylogenetic tree was drawn containing 41 consensus sequences from groups for all accessions and 13 sequences from GenBank (most of them are diploid accessions' sequences).

According to Mangus and Bengt (2007), if phylogeny analysis reveals that different gene copy candidates from the same accession have a parallel sister group relationship in the same clade, they should be treated as the same copy. Thus, if candidate sequences from the same accession appeared in different clades, revealing a nonsister group relationship among other sequence(s), they were taken as distinct gene copies. Based on

this principle, for all candidate sequences of each accession, their placement and relationship in the phylogenetic tree were checked. If there was a sister group relationship among candidates, only one of them was left and the other was excluded from further analyses; this selection was random (Figure 2). This step was crucial to the study as a whole as it allowed for maximal potential to track all different genetic copies for every single accession.

After this analysis, 19 sequences were excluded and 47 sequences were left. Phylogenetic analysis was done once more to check any sister group relationships within sequences for each accession. In the second run, PI 10588R and PI 20963Y were excluded as well. 45 sequences remained and the following analysis and final results were based on them.

Figure 2: Consensus tree in the analysis of the DMC1 data set for 66 sequences using heuristic search with TBR branch swapping. Numbers on branches are bootstrap values. *Bromus arvensis* and *Bromus sterilis* are used as outgroups. Sequences in each box are candidates of distinct gene copies, which turned out to share a sistergroup relationship in a subclade Sequences within pairs of blue arrows share a sistergroup relationship in a clade. For both types, one of the candidates for each accession was randomly excluded (strikethroughs). Sequences in double box are *Hordeum* diploid accessions' sequences from the GenBank database used as an aid in analysis. Based on the present tree, only one of these is sufficient for analysis and the rest were excluded. The letter 'D' following a sequence name means this sequence is from a diploid accession. Letters following accession numbers (such as L, R or Y) are to distinguish between different candidates.


PI 537323L
PI 537323R
PI 537310Y
EU366427
EU366411
AY277263
PI 531578L
EU366414
PI 531578K
PI 13828K
PI 236663K
PI 236663L
PI 598463K
H10584Y
EU366410
H3169L
AY592986
AY137399
AY137401
AY137405
AY137408
AY137412
PI 254866K
H5495R
H6468K
PI 628702
H3169K
H10584L
PI 207452K
PI 207452L
PI 13826L
PI 20963Y
H10588R
AF277245
254866Z
EU366416
EU366406
H6486Y
PI 634293L
PI 610898R
H10339K
H10339R
PI 531708K
PI 531708L
H10588Y
H5495K
H10584K
PI 634293K
PI 499360L
PI 610898K
PI 401280R
PI 531573Y
PI 531573K
PI 401280K
PI 564927
PI 20963R
PI 20963K
PI 616796L
PI 618796R
PI 636675Y
PI 383534Y
PI 636675L
PI 499476L
PI 499476Y
DQ247821
AF277234
D
D
D
D
D
D

After the trial phylogeny analysis, six accessions were confirmed to have two distinguishable gene copies detected; these six accessions possessed StH genome content.

Complete alignment was conducted among the 45 sequences, and showed an indel in two StH species, *E. trachycaulus* (PI 537323) and *E. scabrifolius* (PI537310), at position 225-257. Two smaller indels were also detected; one was at position 268- 282 between *E. hystrix* (EU366415) and *E. cordilleranus* (H6468) and the other was at position from 309bp to 317bp, among *A.cristatum* (PI383534), *E. macrochaetus* (PI618796) and *E. strictus* (PI499476) (Fig 3A and Fig 3B).

The 45 distinct copies of sequences were used for phylogenetic analysis using the maximum parsimony method with *Bromus arvensis* and *Bromus sterilis* as outgroups (Figure 4) . The data matrix contained 1119 characters, of which 654 were constant, 262 were parsimony uninformative, and 203 were parsimony informative. As a result, trees were generated with a consistency index (CI) of 0.775, a retention index (RI) of 0.820, and a rescale consistency index (RCI) of 0.635. The consensus parsimonious tree is shown in Figure 4 with bootstrap values. In the phylogeny tree, the genome type for each clade was determined by the presence of diploid accessions with commonly recognized genome content. Four clades were indicated by the letters Ns, St, H and P respectively.

In the Ns clade, *E. coreanus* (EU366411) is grouped with *P. juncea* (EU366427) and *P. fragilis* (AF277263) and furthermore, *E. coreanus* is closely linked to *P. juncea* with 100% support. *E.coreanus, E. gayanus* and *E. komarovii* formed a sister group with a

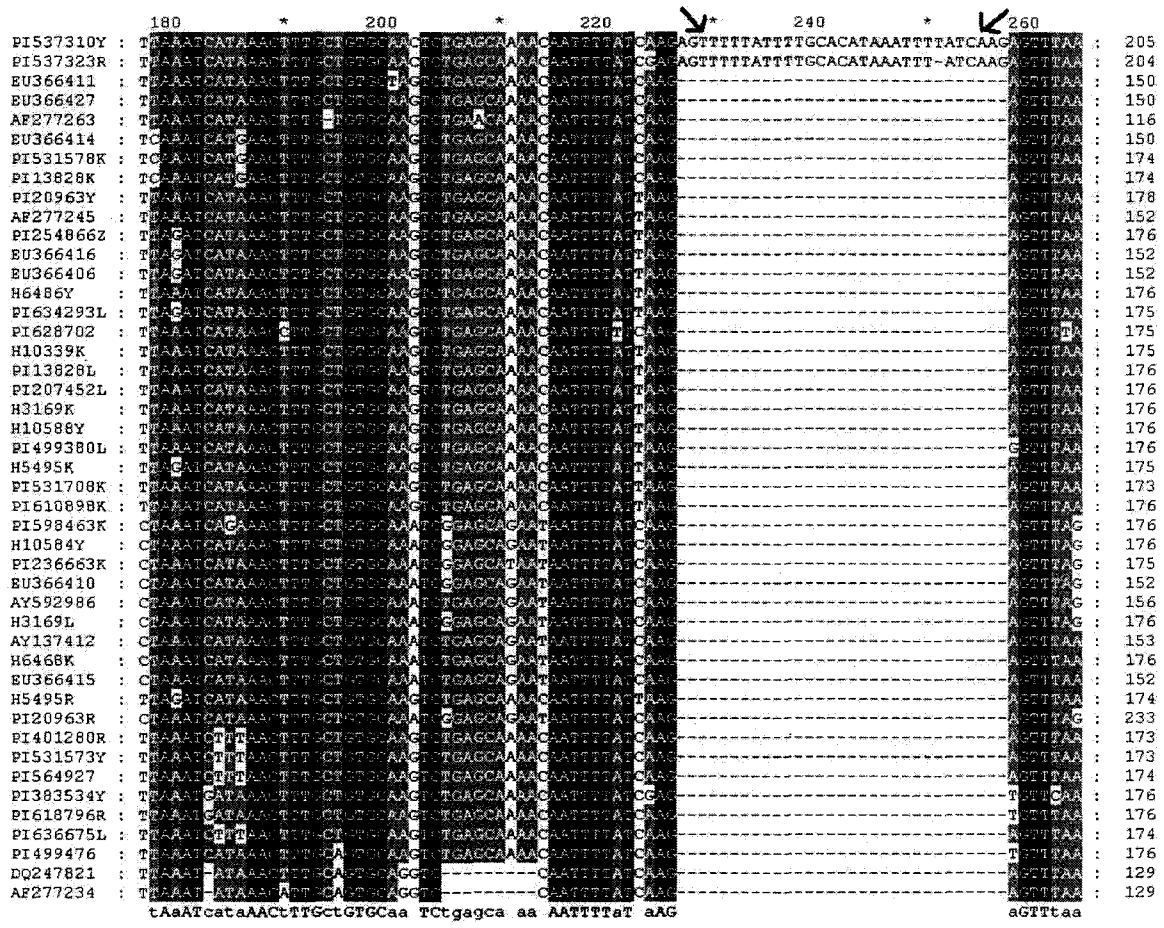
support value of 87%; one of two distinct gene copies of *E. gayanus* was covered in this clade.

In the St clade, twenty-one *Elymus* species were clustered together with 61% bootstrap value support. Within this clade, two sister groups were observed. In one of them, *P. spicata* (AF277245) was grouped with 17 *Elymus* tetraploid species with 98% bootstrap value, which include four StY species: *E. panormitanus* (PI 254866), *E. strictus* (PI 634293), *E. semicostatus* (PI 207452) and *E. gmelinii* (PI 610898). The other subclade, with 75% support, includes *E. multisetus* (PI 20963) and three StY genome species: *E. longearistatus* (PI 401280), *E. caucasicus* (PI 531573) and *E. fedtshenkoi* (PI 564927).

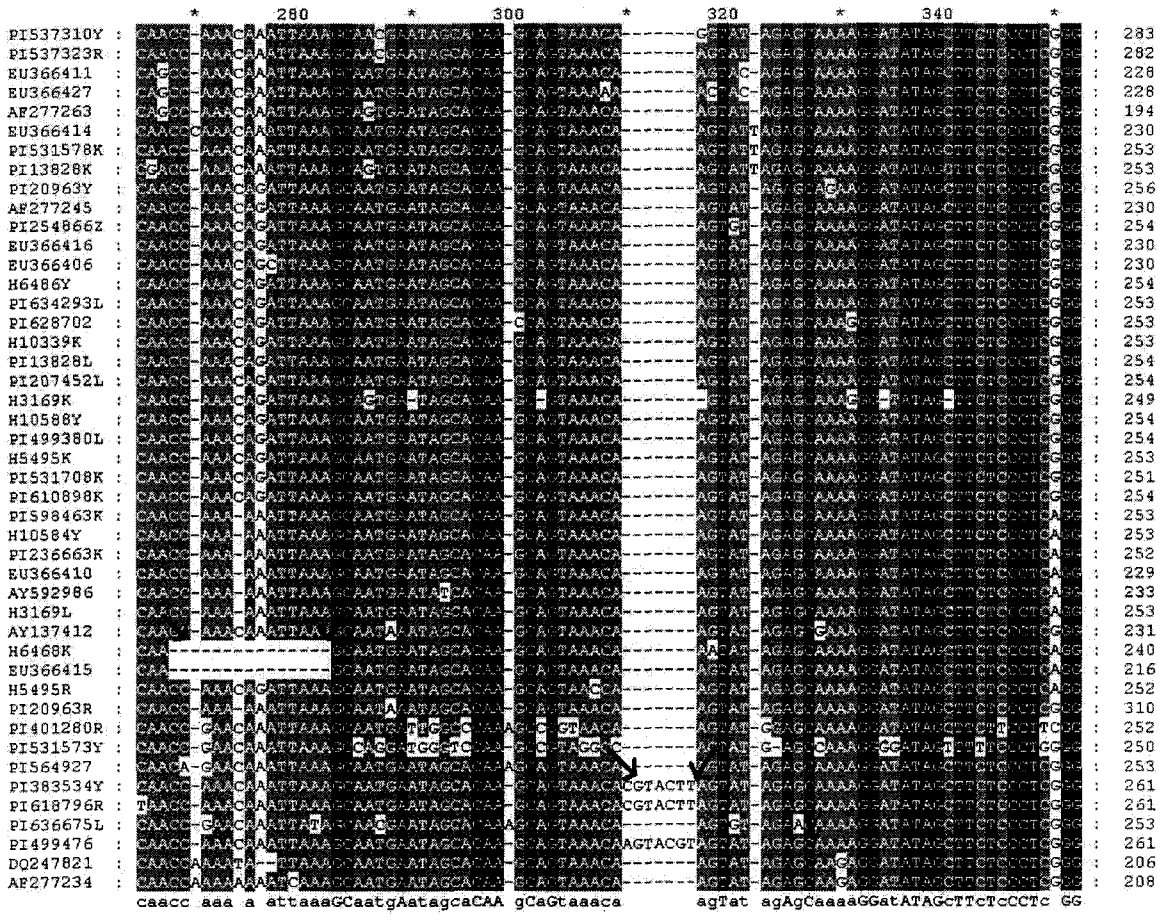
The H clade, a well-supported clade with a bootstrap value of 96%, included *Hordeum* species and nine *Elymus* species. Within the H genome, five *Elymus* species are clustered with *H. capense* (AY592986) in a subclade, which includes *E. confuses* (PI 598463), *E. virescens* (H10584), *E. lanceolatus* (PI 236663), *E. wawawaiensis* (EU366410) and *E. caninus* (H3169). Species including two accessions of *E. hystrix* (EU366415 and H5495) and *E. cordilleranus* (H6468) formed a sister group to the above with 78% support; one of two distinct gene copies of *E. virescens*, *E. caninus*, *E. cordilleranus* and *E. hystrix* were found in this clade. In the P clade, four species are present including *E. macrochaetus* (PI 618796), *E. gayanus* (636675), *E. strictus* (PI 499476) and the P genome type diploid *A. cristatum* (PI 383534).

Neighbor Joining method was employed as well. The result was quite similar as shown in Figure 5.

Figure 3: Partial complete alignment using software Clustal X. A larger indel occurred in two StH species, *E. trachycaulus* (PI 537323) and *E. scabrifolius* (PI 537310), at position of 225-257 (3A). Two smaller indels were also found. One was at position 268 to 282 between *E. hystrix* (EU366415) and *E. cordilleranus* (H6468). The other is at position 309-317 among *A. cristatum* (PI 383534), *E. macrochaetus* (PI 618796) and *E. strictus* (PI 499476) (3B).


180 * 200 * 220 * 240 * 260
PI537310Y : TAAAATCATAAACTTTCCTGTGCAACTTGAGCAAAACAATTTATCAAGAGTTTTATTTTGCACATAAATTTTATCAAGAGTTTAA : 205
PI537323R : TAAAATCATAAACTTTCCTGTGCAACTTGAGCAAAACAATTTATCGAAGTTTTATTTTGCACATAAATTT-ATCAAGAGTTTAA : 204
EU366411 : TAAAATCATAAACTTTCCTGTGCAAGTTGAGCAAAACAATTTATCAAG----------AGTTTAA : 150
EU366427 : TAAAATCATAAACTTTCCTGTGCAAGTTGAGCAAAACAATTTATCAAG----------AGTTTAA : 150
AF277263 : TAAAATCATAAACTTTCTGTGCAAGTTGACAAAACAATTTATCAAG----------AGTTTAA : 116
EU366414 : TCAAATCATGAACTTTCCTGTGCAAGTTGAGCAAAACAATTTATCAAG----------AGTTTAA : 150
PI531578K : TCAAATCATGAACTTTCCTGTGCAAGTTGAGCAAAACAATTTATCAAG----------AGTTTAA : 174
PI13828K : TCAAATCATGAACTTTCCTGTGCAAGTTGAGCAAAACAATTTATCAAG----------AGTTTAA : 174
PI20963Y : TAAAATCATAAACTTTCCTGTGCAAGTTGAGCAAAACAATTTATTAAG----------AGTTTAA : 178
AF277245 : TAAAATCATAAACTTTCCTGTGCAAGTTGAGCAAAACAATTTATTAAG----------AGTTTAA : 152
PI254866Z : TAAGATCATAAACTTTCCTGTGCAAGTTGAGCAAAACAATTTATTAAG----------AGTTTAA : 176
EU366416 : TAAGATCATAAACTTTCCTGTGCAAGTTGAGCAAAACAATTTATTAAG----------AGTTTAA : 152
EU366406 : TAAGATCATAAACTTTCCTGTGCAAGTTGAGCAAAACAATTTATTAAG----------AGTTTAA : 152
H6486Y : TAAAATCATAAACTTTCCTGTGCAAGTTGAGCAAAACAATTTATTAAG----------AGTTTAA : 176
PI634293L : TAAGATCATAAACTTTCCTGTGCAAGTTGAGCAAAACAATTTATTAAG----------AGTTTAA : 175
PI628702 : TAAAATCATAAACGTTCCTGTGCAAGTTGAGCAAAACAATTTATCAAG----------AGTTTAA : 175
H10339K : TAAAATCATAAACTTTCCTGTGCAACTTGAGCAAAACAATTTATTAAG----------AGTTTAA : 175
PI13828L : TAAAATCATAAACTTTCCTGTGCAAGTTGAGCAAAACAATTTATTAAG----------AGTTTAA : 176
PI207452L : TAAAATCATAAACTTTCCTGTGCAAGTTGAGCAAAACAATTTATTAAG----------AGTTTAA : 176
H3169K : TAAAATCATAAACTTTCCTGTGCAAGTTGAGCAAAACAATTTATTAAG----------AGTTTAA : 176
H10588Y : TAAAATCATAAACTTTCCTGTGCAAGTTGAGCAAAACAATTTATTAAG----------AGTTTAA : 176
PI499380L : TAAAATCATAAACTTTCCTGTGCAAGTTGAGCAAAACAATTTATTAAG----------GGTTTAA : 176
H5495K : TAAGATCATAAACTTTCCTGTGCAAGTTGAGCAAAACAATTTATTAAG----------AGTTTAA : 175
PI531708K : TAAAATCATAAACTTTCCTGTGCAAGTTGAGCAAAACAATTTATTAAG----------AGTTTAA : 173
PI610898K : TAAAATCATAAACTTTCCTGTGCAAGTTGAGCAAAACAATTTATTAAG----------AGTTTAA : 176
PI598463K : CAAAATCAGAAACTTTCCTGTGCAAATTGAGCAGAATAATTTATCAAG----------AGTTTAG : 176
H10584Y : CAAAATCATAAACTTTCCTGTGCAAATTGAGCAGAATAATTTATCAAG----------AGTTTAG : 176
PI236663K : CAAAATCATAAACTTTCCTGTGCAAATTGAGCATATTAATTTATCAAG----------AGTTTAG : 175
EU366410 : CAAAATCATAAACTTTCCTGTGCAAATTGAGCAGAATAATTTATCAAG----------AGTTTAG : 152
AY592986 : CAAAATCATAAACTTTCCTGTGCAAATTGAGCAGAATAATTTATCAAG----------AGTTTAG : 156
H3169L : CAAAATCATAAACTTTCCTGTGCAAATTGAGCAGAATAATTTATCAAG----------AGTTTAG : 176
AY137412 : CAAAATCATAAACTTTCCTGTGCAAATTGAGCAGAATAATTTATCAAG----------AGTTTAA : 153
H6468K : CAAAATCATAAACTTTCCTGTGCAAATTGAGCAGAATAATTTATCAAG----------AGTTTAA : 176
EU366415 : CAAAATCATAAACTTTCCTGTGCAAATTGAGCAGAATAATTTATCAAG----------AGTTTAA : 152
H5495R : TAAGATCATAAACTTTCCTGTGCAAGTTGAGCAAAACAATTTATTAAG----------AGTTTAA : 174
PI20963R : CAAAATCATAAACTTTCCTGTGCAAATTGAGCAGAATAATTTATCAAG----------AGTTTAG : 233
PI401280R : TAAAATGTTTAACTTTCCTGTGCAAGTTGAGCAAAACAATTTATCAAG----------AGTTTAA : 173
PI531573Y : TAAAATGTTTAACTTTCCTGTGCAAGTTGAGCAAAACAATTTATCAAG----------AGTTTAA : 173
PI564927 : TAAAATGTTTAACTTTCCTGTGCAAGTTGAGCAAAACAATTTATCAAG----------AGTTTAA : 174
PI383534Y : TAAAATCATAAACTTTCCTGTGCAAGTTGAGCAAAACAATTTATCGAG----------TGTTCAA : 176
PI618796R : TAAAATCATAAACTTTCCTGTGCAAGTTGAGCAAAACAATTTATCAAG----------TGTTTAA : 176
PI636675L : TAAAATGTTTAACTTTCCTGTGCAAGTTGAGCAAAACAATTTATCAAG----------AGTTTAA : 174
PI499476 : TAAAATCATAAACTTTCCAGTGCAAGTTGAGCAAAACAATTTATCAAG----------TGTTTAA : 176
DQ247821 : TAAAAT-ATAAACTTTCCAGTGCAGGT---------CAATTTATCAAG----------AGTTTAA : 129
AF277234 : TAAAAT-ATAAACATTCCAGTGCAGGT---------CAATTTATCAAG----------AGTTTAA : 129
tAaATcataAACtTTGctGTGCaa TCtgagca aa AATTTTaT aAG aGTTtaa

3A


280
300
320
340
PI537310Y
PI537323R
EU366411
EU366427
AF277263
EU366414
PI531578K
PI13828K
PI20963Y
AF277245
PI254866Z
EU366416
EU366406
H6486Y
PI634293L
PI628702
H10339K
PI13828L
PI207452L
H3169K
H10588Y
PI499380L
H5495K
PI531708K
PI610898K
PI598463K
H10584Y
PI236663K
EU366410
AY592986
H3169L
AY137412
H6468K
EU366415
H5495R
PI20963R
PI401280R
PI531573Y
PI564927
PI383534Y
PI618796R
PI636675L
PI499476
DQ247821
AF277234
caacc aaa a attaaaGCaatgAatagcaCAA gCaGtaaaca agTat agAgCaaaaGGatATAGcTTcTcCCTc GG

3B

Figure 4: A strict consensus tree derived from DMC1 sequence data of 45 sequences from 39 accessions was generated using a heuristic search with TBR branch swapping. Numbers on branches are bootstrap values. *Bromus arvensis* and *Bromus sterilis* are used as outgroups. Species whose names followed by signal + are StY *Elymus* species; species whose names followed by signal * are *Hystrix* species. The genome types for clades were determined and indicated by the presence of diploid accessions with commonly recognized genome content as shown in parentheses after species' names.

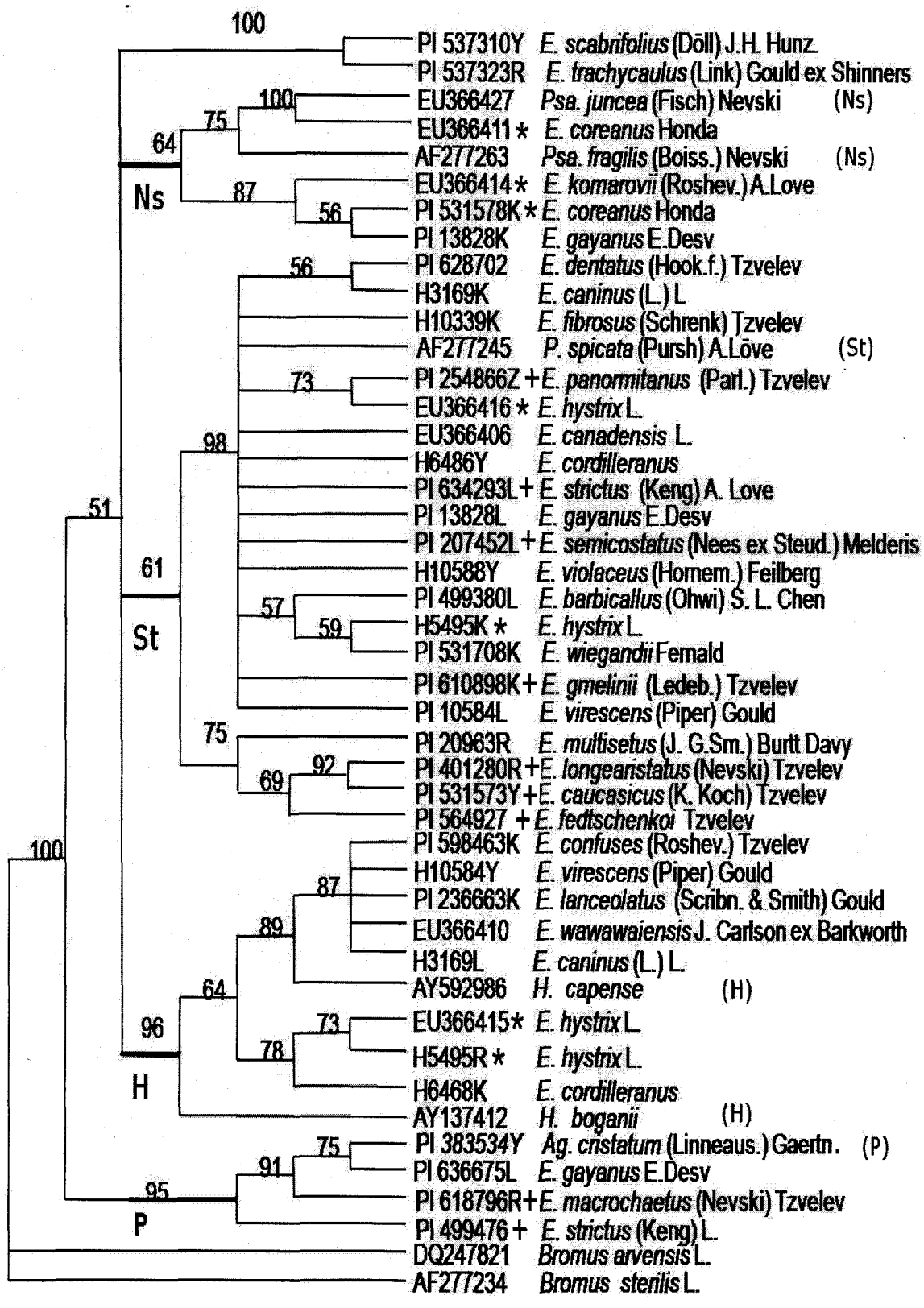
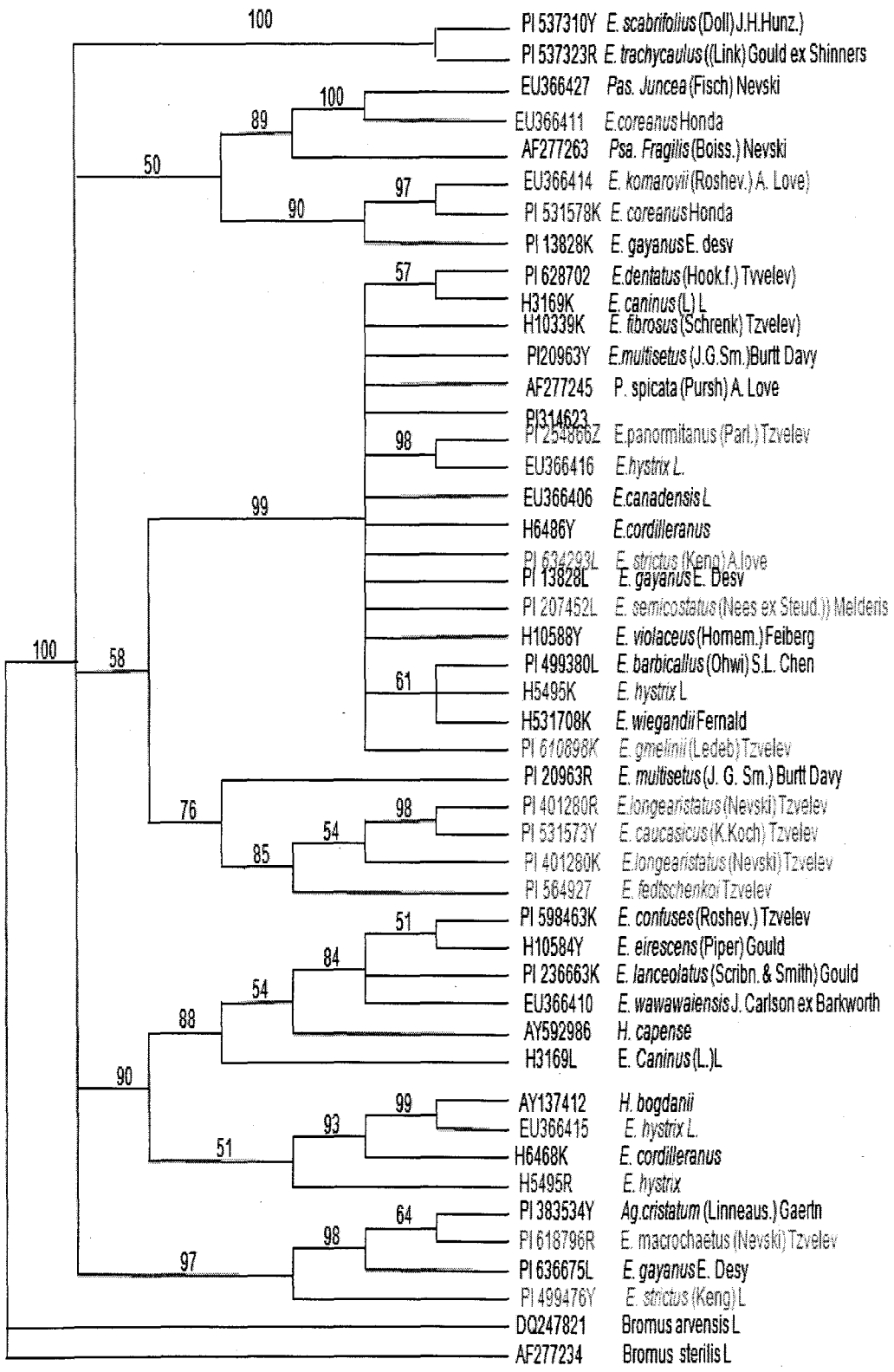

100
PI 537310Y E. scabrifolius (Döll) J.H. Hunz.
PI 537323R E. trachycaulus (Link) Gould ex Shinners
100
EU366427 Psa. juncea (Fisch) Nevski (Ns)
75
EU366411 * E. coreanus Honda
64
AF277263 Psa. fragilis (Boiss.) Nevski (Ns)
Ns
87
EU366414 * E. komarovii (Roshev.) A.Love
56
PI 531578K * E. coreanus Honda
PI 13828K E. gayanus E.Desv
56
PI 628702 E. dentatus (Hook.f.) Tzvelev
H3169K E. caninus (L.) L
H10339K E. fibrosus (Schrenk) Tzvelev
AF277245 P. spicata (Pursh) A.Löve (St)
73
PI 254866Z + E. panormitanus (Parl.) Tzvelev
EU366416 * E. hystrix L.
98
EU366406 E. canadensis L.
H6486Y E. cordilleranus
PI 634293L + E. strictus (Keng) A. Love
51
PI 13828L E. gayanus E.Desv
PI 207452L + E. semicostatus (Nees ex Steud.) Melderis
61
H10588Y E. violaceus (Homem.) Feilberg
PI 499380L E. barbicallus (Ohwi) S. L. Chen
57
59
H5495K * E. hystrix L.
St
PI 531708K E. wiegandii Fernald
PI 610898K + E. gmelinii (Ledeb.) Tzvelev
PI 10584L E. virescens (Piper) Gould
75
PI 20963R E. multisetus (J. G.Sm.) Burtt Davy
92
PI 401280R + E. longearistatus (Nevski) Tzvelev
69
PI 531573Y + E. caucasicus (K. Koch) Tzvelev
PI 564927 + E. fedtschenkoi Tzvelev
100
PI 598463K E. confusus (Roshev.) Tzvelev
H10584Y E. virescens (Piper) Gould
87
PI 236663K E. lanceolatus (Scribn. & Smith) Gould
89
EU366410 E. wawawaiensis J. Carlson ex Barkworth
H3169L E. caninus (L.) L
64
AY592986 H. capense (H)
96
73
EU366415* E. hystrix L.
78
H5495R * E. hystrix L.
H
H6468K E. cordilleranus
AY137412 H. boganii (H)
75
PI 383534Y Ag. cristatum (Linneaus.) Gaertn. (P)
91
PI 636675L E. gayanus E.Desv
95
PI 618796R + E. macrochaetus (Nevski) Tzvelev
P
PI 499476 + E. strictus (Keng) L.
DQ247821 Bromus arvensis L.
AF277234 Bromus sterilis L.

Figure 5. A strict consensus tree was constructed using Neighbor Joining Method. The result was quite similar to the one from Maximum parsimony method.


PI 537310Y E. scabrifolius (Doll) J.H.Hunz.)
PI 537323R E. trachycaulus ((Link) Gould ex Shinners
EU366427 Pas. Juncea (Fisch) Nevski
EU366411 E.coreanus Honda
AF277263 Psa. Fragilis (Boiss.) Nevski
EU366414 E. komarovii (Roshev.) A. Love)
PI 531578K E. coreanus Honda
PI 13828K E. gayanus E. desv
PI 628702 E.dentatus (Hook.f.) Tvvelev)
H3169K E. caninus (L) L
H10339K E. fibrosus (Schrenk) Tzvelev)
PI20963Y E.multisetus (J.G.Sm.)Burtt Davy
AF277245 P. spicata (Pursh) A. Love
PI314623
PI 254866Z E.panormitanus (Parl.) Tzvelev
EU366416 E.hystrix L.
EU366406 E.canadensis L
H6486Y E.cordilleranus
PI 634293L E. strictus (Keng) A.love
PI 13828L E. gayanus E. Desv
PI 207452L E. semicostatus (Nees ex Steud.)) Melderis
H10588Y E. violaceus (Hornem.) Feiberg
PI 499380L E. barbicallus (Ohwi) S.L. Chen
H5495K E. hystrix L
H531708K E. wiegandii Fernald
PI 610898K E. gmelinii (Ledeb) Tzvelev
PI 20963R E. multisetus (J. G. Sm.) Burtt Davy
PI 401280R E.longearistatus (Nevski) Tzvelev
PI 531573Y E. caucasicus (K.Koch) Tzvelev
PI 401280K E.longearistatus (Nevski) Tzvelev
PI 564927 E. fedtschenkoi Tzvelev
PI 598463K E. confuses (Roshev.) Tzvelev
H10584Y E. eirescens (Piper) Gould
PI 236663K E. lanceolatus (Scribn. & Smith) Gould
EU366410 E. wawawaiensis J. Carlson ex Barkworth
AY592986 H. capense
H3169L E. Caninus (L.)L
AY137412 H. bogdanii
EU366415 E. hystrix L.
H6468K E. cordilleranus
H5495R E. hystrix
PI 383534Y Ag.cristatum (Linneaus.) Gaertn
PI 618796R E. macrochaetus (Nevski) Tzvelev
PI 636675L E. gayanus E. Desy
PI 499476Y E. strictus (Keng) L
DQ247821 Bromus arvensis L
AF277234 Bromus sterilis L

4. Discussion

4.1. Origin of StH species

It is a popular hypothesis supported by cytological analyses that polyploid *Elymus* species are derived from polyploidization through hybridization between different ancestral genera. For example, the StH genome *Elymus* species are suggested to be derived from hybridization between *Pseudoroegneria* (St) and *Hordeum* (H) (Dewey, 1984; Lu, 1994). In the present study, most sequences from StH tetraploid species are clustered within St and H clades and these molecular results provide yet more evidence for this hypothesis.

The St genome is present in all *Elymus* species. A few recent molecular studies agree that *Pseudoroegneria* was the maternal genome donor during the polyploid speciation of *Elymus* species. Meiotic pairing data has indicated that the St genome in the *Elymus* and *Pseudoroegneria* have a very high homology (Dewey, 1982; Torabinejad *et al.*, 1993; Jensen *et al.*, 1995). Comparisons of the Giemsa C-banded and N-banded karyotypes of *Elymus* species also suggest that the St genome in *Elymus* resembles the St genome in *Pseudoroegneria* (Linde-Laursen *et al.*, 1994). The chloroplast gene trnL-F sequences revealed an especially close relationship of *Pseudoroegneria* to all *Elymus* species (Liu *et al.*, 2006). Another study based on analysis of the single copy gene RPB2 suggested that *P. spicata* may be the St genome donor for *Elymus* species (Sun *et al.*, 2008). In the present study, *P. spicata* is closely clustered with fifteen tetrapolyploid *Elymus* species, all of which include the St genome as defined by cytological data. On this point, the results are in agreement with previous studies.

In the St clade, two subclades are observed, one of which includes 17 sequences from *Elymus* species and *Pseudoroegneria* with 98% support. The other one contains one

sequence from StH species and three StY species. *E. panormitanus* (PI 254866Z) and *E. hystrix* (EU366416) with 73% bootstrap value; *E. dentatus* (PI 628702) and *E. caninus* (H3169) with 56% bootstrap value and *E. Hystrix* (H5495K), *E. wiegandii* (PI 531708K) and *E. barbicallus* (PI 499380L) with 57% bootstrap value. The most convincing hypothesis from these results is that the St genome in tetrapolyploid *Elymus* originated from more than one *Pseudoroegneria* species/population through hybridization. According to previous literature, the St haplome donor genus, *Pseudoroegneria*, contains approximately 15 diploid or tetraploid species distributed worldwide (Dewey, 1984; Löve, 1984) and cytological data suggest that genome differentiation exists among the *Pseudoroegneria* species (Wang *et al.*, 1986). Analysis form RPB2 gene has revealed that both *Pseudoroegneria spicata* and *Pseudoroegneria stipifolia* may be the St genome donor of *Elymus* species (Sun *et al.*, 2008). Another study on North American *Elymus* using the starch synthase gene also suggests the possibility of multiple origins of the St genome in *Elymus* from a polymorphic *Pseudoroegneria* gene pool (Mason-Gamer *et al.*, 2005).

Hordeum has been generally hypothesized as another genome donor to the StH *Elymus* species based on genome-pairing data (Dewey 1984; Jensen *et* al., 1990a, 1990b, 1995), starch synthase sequence analyses (Mason-Gamer *et al.*, 2001) and *RPB2* data (Sun *et al.*, 2008). In the present study, a well-supported clade is drawn with a bootstrap value of 96% including *Hordeum* and StH *Elymus* species. This result is consistent with previous findings that *Hordeum* may be one of the genome donors of *Elymus*; the existence of subclades within the H clade indicates that the H genome of StH species originates either from different diploid *Hordeum* species donors. Evidence for multiple

origins from a polymorphic *Hordeum* gene pool has also been found in other studies using different molecular tools such as the RAPD method (Sun *et al.*, 1997) and other studies using data from the single copy gene RPB2 (Sun *et al.*, 2007). The above molecular studies do not conflict with C- and N-banded karyotype research suggesting that the H genomes of different *Elymus* species have different karyotypes, indicating their polymorphic origins (Linde-Laursen *et al.*, 1994).

4.2. Genome content and classification of *Hystrix*

Hystrix Moench is a small perennial genus of the tribe Triticeae. The genus *Hystrix* revised includes six species and one of them is divided into three subspecies: *H. patula*, *H. californica* (Bol.) Kuntze, *H. duthiei* (Stapf) Bor subsp. *duthiei*, *H. duthiei* subsp. *longearistata* (Hack.) Baden, Fred. & Seberg, *H. duthiei* subsp. *japonica* (Hack.) Baden, Fred. & Seberg, *H. komarovii* (Roshev.) Ohwi, *H. coreana* (Honda) Ohwi and *H. sibirica* (Trautv.) Kuntze. (Baden *et al.*, 1997; Fan *et al.*, 2007).

The definition of genus *Hystrix* and its precise taxonomic treatment are still under discussion. Some scientists have suggested placing these species in *Hystrix* (Osada 1993; Baden *et al.*, 1997; Zhou *et al.*, 2000) while others have suggested they be placed within the *Elymus* genus (Dewey 1982; Löve 1984) or *Leymus* Hochst (Jensen *et al.*, 1997).

Originally, the only known species was described as *H. patula*, which was believed to possess a close affinity to species of the *Elymus canadensis* complex, and was treated as *E. hystrix* (Jensen *et al.*, 1997). Since *E. canadensis* is supposed to contain the basic genome of StH, then *E. hystrix* should also contain the same basic genomes (Dewey 1982; Claw *et al.*, 1997).

Recently, based on the results of GISH and Southern genomic hybridization, it was supported that *E. hystrix* does possess the StH genome (Ellenskog-Staam *et al.*, 2007). Conclusions from nuclear rDNA ITS and single copy plastid ACCase data are also consistent with this hypothesis and high intraspecific genetic diversity is suggested to exist in *E. hystrix*. Two copies of the Acc1 gene from *E. hystrix* appear to be rather divergent (Fan *et al.*, 2007; Zhang *et al.*, 2008). However, disagreement arises from analysis of genome-specific repetitive DNA and RAPD markers, which implies that the St genome is either not present in *E. hystrix* or that the St genome has been considerably modified during the evolution of the species since no trace of the St genome was found in this study (Svitashev *et al.*, 1998). In single copy gene rpb2 phylogeny analysis, only one gene copy was found and was clustered in the H group; no St copy was detected according to these results (Sun *et al.*, 2008). In the present analysis, two accessions from *E. hystrix* are included: H5495 and PI531616. For both accessions, one genetic copy is well clustered in the St clade and the other in the H clade, as shown in Figure 2. The present results show distinct gene copies from two *E. hystrix* accessions (EU366416 and H5495) are present in the St and H clades, respectively. DMC1 gene data confirm strongly that *E. hystrix* possesses the StH genome and could be placed in *Elymus*.

On the other hand, in the genus *Hystrix,* other species were supposed to carry entirely different genome contents. It has been suggested that *H. duthiei* subsp. *duthiei, H. duthiei* subsp. *longearistata*s, *H. coreana* and *H. komarovii* share a close affinity with *Psathyrostachys* and *Leymus* species, and that *H. komarovii* may contain the NsXm genome of *Leymus* (Jenson *et al.*, 1997; Svitashev *et* al., 1998; Dubcovsky *et al.*, 1997; Fan *et al.*, 2007; Zhang *et al.*, 2008). In the present study, *E. coreanus* (*H. coreana*) and

E. komarovii (*H. komarovii*) are grouped with *P. juncea and P. fragilis.* The data show that the species *E. komarovii and E. coreanus* do have a close relationship to the genus *Leymus*.

Therefore, as mentioned, the definition of the genus *Hystrix* is under discussion and its precise taxonomic treatment needs revision. Base on present study and previous suggestions, it is reasonable to treat *Hystrix patula* as *Elymus hystrix L* and to treat *E. coreanus* (*H. coreana*) and *E. komarovii* (*H. komarovii*) as species of a section of *Leymus.*

4.3. The relationship of the Y genome to other genomes

The origins of the Y genome remain unknown, although it is a common *Elymus* genome in central and eastern Asia (Dewey, 1984). No diploid species containing the Y genome have been identified (Wang *et al.*, 1994). Chromosome pairing analysis has reported low affinities between the ''St" and ''Y" genomes (Lu *et al.*, 1989). A few datasets used in previous work could not clearly separate the StYgenome species from StH genome species (Liu *et al.*, 2006). ITS sequence data suggest that all StY species have only one genomic type of ITS sequence and all sequences are included in the St clade. The preferred explanation is that the St and Y genomes may have the same origin since the St genome clade in this study did not show any St- or Y-genome associated subclades. However, it is worth pointing out the possibility that homogenization of ITS sequences has occurred between the St and Y genomes in polyploid *Elymus* species (Liu *et al.*, 2006; Rauscher *et al.*, 2004). In another, recent, study using the single copy gene RPB2, two genomically distinct sequences were found for each StY species. One sequence from each StY species was grouped in the St clade. Other sequences that should belong to the

Y genome formed a distinct clade from that of the St genome and all the Y-genome copies showed a different insertion/deletion pattern from St copies. This study suggests that the St and Y genomes have distinct origins (Sun *et al.*, 2008), which is in agreement with analyses based on β-amylase gene sequence data (Mason- Gamer *et al.*, 2005).

In the present study, for all nine StY species, only one genomic type of DMC1 sequence was found. Four accessions of StY species were placed in the St clade, which included *E. panormitanus*, *E. strictus*, *E. semicostatus* and *E. gmelinii*. Three accessions grouped together have a relationship with the St clade with a moderate bootstrap value of 61%, including *E. longearistatus*, *E. caucasicus* and *E. fedtschenkoi*. The remaining two accessions were placed distinctly in the P clade.

There is more than one scenario to explain the result based on lack of two distinguishable genome copies: 1) The St and Y genomes may be considered to have a close relationship on this locus. This hypothesis agrees with another study using ITS sequences (Liu *et al.*, 2006). Based on cytological research, relatively close affinities between the St and Y genomes was reported (Lu *et al.*, 1992). It is also possible that homogenization of DMC1 sequences has occurred between the St and Y genomes in polyploid *Elymus* species. But because DMC1, as a single copy gene, was supposed to exhibit low potentiality of concerted evolution (Small *et al.*, 2004), this hypothesis was not supported. 2) The number of single clone sequences for each accession is not enough to detect all the two different copies. In present study, on average five to seven random clones were sequenced for each StY accession due to limitation in the number of recombinants confirmed after cloning effort (low cloning efficiency). The possibility might be that the limited that sequences obtained in present study cover all the two

distinct gene copies. If so, the relationship between St and Y genome could not be determined. 3) There are mutations on present Y genome on the primer binding sites so that its gene information could be not revealed and amplified by PCR.

The other issue, which should be paid attention to, is the relationship between the P and Y genome. In present result, two other StY accessions, *Elymus macrochaetus* and *E. strictus*, are included in the P genome clade with 95% support. This clade is distantly separated from all the others. It suggests the P genome may share a relationship with the StY species and this is supported by RPB2 gene data (Sun *et al.*, 2008). A direct contribution from the P genome to the genomes involved in StY species through introgression remains a possibility.

Another possibility is that the original donor of the Y genome has become extinct. One piece of evidence for this hypothesis is that no diploid Y genome species have been so far detected (Lu, 1993b). According to present results and previous studies, it is also possible that the St and P genomes both have had an influence on the current Y genome in StY species.

4.4. Utility of DMC1 in phylogenetic studies of tetraploid *Elymus*

In the present analysis, the DMC1 data matrix obtained contained 1119 characters, of which 654 were constant, 262 were parsimony uninformative and 203 were parsimony informative. The number of phylogenetically informative characters is due to a rather economic character sampling compared to the sequencing efforts in other studies of *Elymus* (Mason-Gamer, 2001). So far, only a few single copy nuclear genes have been used in the phylogeny of *Elymus,* including granule-bound starch synthase (Mason-

Gamer, 2001), phosphoenolpyruvate carboxylase (Helfgott and Mason-Gamer, 2004), and the RPB2 gene (Sun *et al.*, 2007, 2008).

DMC1 has been applied to a limited number of phylogenetic analyses. Phylogenetic analyses of the diploid species of *Hordeum* were carried out using DMC1 combined with other gene datasets. A revised classification of the genus was suggested (Peterson and Seberg, 2004a). In the study of *Psathyrostachys*, a direct PCR sequencing method was used. *Psathyrostachys* consists mainly of diploids, but tetraploid cytotypes are known in *P. lanuginose* and *P. fragilis* subsp. *secaliformis*. The direct sequencing of a nuclear gene could be problematic if there are two different copies of the gene (Peterson and Seberg, 2004b). Differing from former studies dealing with diploid species, another study based on DMC1 and EF-G focused on the tetraploid *Hordelymus* species (Peterson and Seberg, 2008), where seven accessions of *Hordelymus* were used. Two copies of DMC1 were found for these accessions. However, the originally distinct alleles were more similar than expected and overall sequence variability was reduced as it was apparent that these two DMC1 groups in the phylogenetic tree were much more closely related than the two EF-G sequence groups. The reason for the high homology of these two different copies is still under discussion (Peterson and Seberg, 2008).

The present study is the first attempt to investigate a large number of tetraploid species in Triticeae using the DMC1 gene. Among the thirty one tetrapolyploid accessions involved, only six accessions were discovered to have two distinguishable gene copies and all of these show StH genome contents. A possible explanation is that on this locus, the gene variability level was originally limited between the two genome donors. An alternative hypothesis is that one of the DMC1 sequence copies originally

inherited was later deleted, whereas the other was duplicated wholly or partially (Peterson and Seberg, 2008). As a result, these two paralogous sequences could not be detected or become more difficult to detect.

Studies of the evolution of genomes in polyploid species always encounter a difficulty in distinguishing the orthologous (or homoeologous) copy of the gene in polyploid species (Sun *et al.*, 2007). On this point, the DMC1 gene might not be the most efficient choice for analysis of tetraploid species in *Elymus.*

Stowaway elements were found in intron 14 of the DMC1 gene in some species of Triticeae, including *Taeniatherum caput-medusae, Australopyrum retrofractum*, *Australopyrum pectinatum* and *Australopyrum velutinum*; these species share the presence of an element of lengths 76, 59, 59 and 75 bp, respectively. The remaining Triticeae species, including *Hordeum* and *Psathyrostachys* species, do not exhibit this element, although footprints (remains of an element left after excision) seem to be present in species such as *Henrardia persica*, *Eremopyrum distans*, *Sitopsis speltoides* and *Amblyopyrum muticum* (Peterson and Seberg, 2000). In the present study, it is consistent with the expectation that no similar element or footprint was detected in *Elymus*, *Hordeum* or the other species investigated.

4.5. Summary

This molecular phylogeny study using DMC1 gene data verifies that tetraploid StH members of *Elymus* have an St and H genomic content, and are derived from *Hordeum* and *Pseudoroegneria.* The current data revealed within StY species is limited. It is possible that St and P have a relationship to Y genome in StY species. The detailed

relationships among *Elymus* StY species, and their relationships to any specific genome donor, await a better-resolved phylogeny including more StY accessions and distinct gene copies for each. Additionally, for four tetraploid accessions which were formerly classified into genus *Hystrix* or *Elymus*, their precise taxonomic treatment was revised: *Hystrix patula* should be *Elymus hystrix L*, *E. coreanus* (*H. coreana*) and *E. komarovii* (*H. komarovii*) are species of a section of *Leymus.* The DMC1-based analysis of tetraploid *Elymus* species is a small step forward to better understand the historical relationships within *Elymus*.

PART 2. Phylogenetic relationship between tetraploid *Hordelymus* and *Hordeum*

1. Introduction

Hordelymus europaeus has been referred to other genera in Triticeae, e.g. *Hordeum* or *Elymus*, which is an indication that its position and relationship within the Triticeae have been unclear (Bothmer and Jacobsen, 1989). Very few studies have been conducted on the origin of *Hordelymus* and its relationship to other genera in the tribe. Cytogenetic studies suggested that *Hordelymus* was originated from *Taeniatherum* (Ta genome) and *Hordeum* (H genome) (Bothmer and Jacobsen, 1989). Another cytogenetic study confirmed the presence of Ta, but refuted the presence of H genome, and proposed that Ns genome from *Psathyrostachys* was present in *Hordelymus* (Bothmer *et al.,* 1994). A site-specific protein research further confirmed the absence of H genome (Pelger, 1993). The presence of Ns is indicated in an RAPD analysis, but the presence of Ta is suspicious since no species containing Ta genome was analyzed by RAPD directly (Svitashev *et al.*, 1998). Genomic hybridization suggested *Hordelymus* is autopolyploid with only Ns genome (Ellneskog-Staam *et al.*, 2006). Recently, Petersen and Seberg (2008) reported *Psathyrostachys* as female genome donor of *Hordelymus*. The data from the nuclear *DMC*1 gene is consistent with this placing one of the two sequence copies as the sister to *Psathyrostachys*. However, data from the nuclear *EF-G* gene do not place any of the two *Hordelymus* sequence copies as sister to *Psathyrostachys* (Petersen and Seberg, 2008).

Additional data from nuclear sequences are need in order to strengthen or reject the sister group relationships suggested by Petersen and Seberg (2008).

RPB2 is a single nuclear gene encoding for the second largest subunit of RNA polymerase, which is analogous to the "subunit" in the RNA polymerase enzyme in prokaryotes (Sweetser *et al.*, 1987; Kolodziej *et al.*,, 1990; Denton *et al.*, 1998). Recently, Sun *et al.*, (2007) have characterized molecular evolution of *RPB2* in tetrploid StH genomic species of *Elymus*, and used *RPB2* gene to explore the origin, geographic differentiation of H genome, and the relationship of the Y genome to other genomes in *Elymus* species (Sun *et al.*, 2008), and phylogeny of *Hordeum* (Sun *et al.*, 2009). Higher level of polymorphism and genome-specific amplicons generated by this gene indicated that *RPB2* is an excellent candidate gene to test the relationships among *Hordelymus*, *Psathyrostachys*, *Hordeum*, and *Taeniatherum*.

In this study, we sequenced and analyzed the *RPB2* gene from *Hordelymus europaeus* (L.) Harz, and its potential diploid ancestor species that were suggested in previous studies (Bothmer and Jacobsen 1989; Bothmer *et al.*, 1994; Petersen and Seberg 2008). The focus of this study was to examine the phylogenetic relationship of *Hordelymus* genomes with *Hordeum, Psathyrostachys*, *Hordeum*, and *Taeniatherum.* Several others genomes in Triticeae species were also included for analysis. Here we are the first to include all diploid *Hordeum* and *Taeniatherum* in the analysis, and show that a close phylogenetic relationship between tetraploid *Hordelymus* and diploid *Hordeum* species.

2. Materials and methods

2.1. Plant materials

Four accessions of *Taeniatherum* species and two accessions of *Hordelymus europaeus* were used. They were analyzed together with 35 taxa (42 accessions) of the Triticeae previously used for phylogenetic analysis of *Hordeum* (Sun *et al.*, 2009) and *Elymus* (Sun *et al.*, 2008). *Bromus catharticus* and *B. inermis* were used as an outgroup based on previous phylogenetic analysis of Poaceae (Hsiao *et al.*, 1995). The genomic constitutions, accession numbers, and geographical origins were given in Table 3.

Table 3. Taxa used in this study (Part 2)

Species	Accession No.	Genome*
Bromus catharticus		CN32048
Bromus inermis		CN 30955
Hordelymus europaeus		NGB9217
Hordelymus europaeus		NGB9218
Taeniatherum caput-medusae subsp. *caput-medusae*		PI 208075
T. caput-medusae subsp. *caput-medusae*		PI 220591
T. caput-medusae subsp. *caput-medusae*		PI 222048
T. caput-medusae subsp. *asperum*		PI 561091
Agropyron cristatum (L.) Gaertn.		PI 383534
Australopyrum retrofractum (Vickery) Á. Löve		PI 533014
		PI 547363
Dasypyrum villosum (L.) P. Candargy		PI 368886
Pseudoroegneria. libanotica (Hack.) D. R. Dewey		PI 330688
		PI 330687
		PI 401274

	Origin	GenBank No
	Plant Gene Resource of Canada	
	Plant Gene Resource of Canada	
??	KVL 3090, Denmark	
??	KVL 3091, Denmark	
Ta	Kars, Turkey	
Ta	Afghanistan	
Ta	Afghanistan	
Ta	Siirt, Turkey	
P	Kars, Turkey	EU187438
W	New South Wales, Australia	EU187482
W	New South Wales, Australia	EU187470
V	Gaziemir, Turkey	EU187471
St	Sirak-Sar, Iran	EF596751
St	Kandavan Pass, Iran	EF596753
St	Saqqez, Iran	EF596752

Table 3 (Continue)

Species	Accession No.	Genome*	Origin	GenBank No
P. spicata (Pursh) Á. Löve	PI 506274	St	Washington, United States	EF596746
	PI 610986	St	Utah, United States	EF596747
P. stipifolia (Czern. ex Nevski) Á. Löve	PI 325181	St	Stavropol, Russian Federation	EF596748
H. stenostachys Godr.	H1780	H	Argentina	
H. cordobense Bothmer, Jacobsen & Nicora	H6460	H	Argentina	
H. flexuosum Steud.	H2127	H	Uruguay	
Psathyrostachys juncea	PI 406469	Ns	Former Soviet Union	
Lophopyrum elongatum (Host) Á. Löve	PI 142012	E^e	Odessa, Russian Federation	EU187439
Thinopyrum bessarabicum (Savul. & Rayss) Á. Löve	PI 531712	E^b	Estonia	EU187474
Hordeum bogdanii Wilensky	PI 499498	H	Inner Mongolia, China	EF596768
	PI 499645	H	Xinjiang, China	EU187472
	H4014	H	Pakistan	
H. stenostachys Godr.	H6439	H	Argentina	EU187473
H. vulgare subsp. *vulgare*	H7514A	I	China	
H. vulgare ssp. *spontaneous* (K.Koch) Thell	H3140A	I	Cyprus	
H. bulbosum L.	H3878	I	Italy	
H. murinum L. subsp. *glaucum* (Steud.)Tzvel.	H74	Xu	Egypt	
	H52	Xu	Jordan	

Table 3 (Continue)

Species	Accession No.
H. intercedens Nevski	H1941
H. pubiflorum Hook. f.	H1236
H. comosum Presl.	H1181
H. marinum Huds. subsp. *marinum*	H121
H. marinum Huds. subsp. *gussoneanum* (Parl.)Thell	H581
H. roshevitzii Bowden	H9152
H. patagonicum (Haumann) Covas subsp. *patagonicum*	H6052
H. patagonicum (Haumann) Covas subsp. *setifolium* (Parodi & Nicora) Bothm. *et al.*,	H1352
H. patagonicum (Haumann) Covas subsp. *santacrucense* (Parodi & Nicora) Bothm. *et al.*,	H1353
H. patagonicum (Haumann) Covas subsp. *magellanicum* (Parodi & Nicora) Bothm. *et al.*,	H1342
H. patagonicum (Haumann) Covas subsp. *mustersii* (Nicora) Bothm. *et al.*,	H1358
H. chilense Roem. & Schult.	H1816

Genome*	Origin	GenBank No
H	U.S.A.	
H	Argentina	
H	Argentina	
Xa	Greece	
Xa	Greece	
H	China	
H	Argentina	
H	Argentina	
H	Argentina	
H	Argentina	
H	Argentina	
H	Chile	

Table 3 (Continue)

Species	Accession No.	Genome*	Origin	GenBank No
H. brachyantherum Nevski subsp. *californicum*	H3317	H	U.S.A.	
(Covas & Stebbins) Bothm. *et al.*,*H. pusillum* Nutt.	H2024	H	U.S.A.	
H. euclaston Steud.	H2148	H	Uruguay	
H. muticum J. Presl.	H6479	H	Argentina	
H. erectifolium Bothmer, Jacobsen & Jørgensen	H1150	H	Argentina	

*Note: The genome designations are according to Wang *et al.*, (1994).

2.2. DNA amplification and sequencing

The *RPB2* sequences were amplified by polymerase chain reaction (PCR) using the primers of P6F and P6FR. Protocols are given in Sun *et al.*, (2007). The amplified PCR products were cloned into pGEM-easy T vector (Promega Corporation, Madison, WI, USA) according to the manufacturer's instruction. The detailed methods afterwards were the same as Part 1.

2.3. Data analysis

Automated sequence outputs were visually inspected with chromatographs. The sequences from both strands of each clone were compared using BLASTN to find the overlap region, and was joined together to generate full sequence of each clone. Multiple sequence alignments were performed using ClustalX with default parameters (Thompson *et al.*, 1997). Phylogenetic analysis using the maximum parsimony (MP) method was performed with the computer program PAUP* version 4 beta 10 Win (Swofford, 2003). Detailed parameters and options were the same as Part 1.

3. Results

3.1. Sequence variation

Forty-nine sequences and two outgroup sequences were analyzed. Two distinguishable copies of sequences were obtained from *Hord. europaeus* (NGB9218) and were designed as NGB9218 and NGB9218U. Only one copy of sequence was obtained from accession

NGB9217 of *Hord. europaeus*. Obvious deletion was found between the two copies of sequences from NGB218 (*Hord. europaeus*). The sequence NGB9218U showed 24-bp (ACTGGTAAATGACACGAATCTTTA) deletion compared to another copy sequence from NGB9218 and the sequence from *Hord. europaeus* accession NGB9217 (Figure 6). This deletion was observed in the two sequences from *Bromus* species, but not observed in the sequences from any Triticeae species examined here. The identical residues between the NGB9128U and two *Bromus* sequences are 89%, which is the highest score among the pairwise comparison between NGB9128U and any other sequences. The both NGB9218 and NGB9217 sequences showed highest identical residues (94%) with H7514A (*H. vulgare* subsp. *vulgare*), followed by 93% with H3140A (*H. vulgare* subsp. *spontaneous*), H3878 (*H. bulbosum*), H1352 (*H. patagonicum* subsp. *setifolium*), H1342 (*H. patagonicum* subsp. *magellanicum*), H1353 (*H. patagonicum* subsp. *santacrucense*), and H1236 (*H. pubiflorum*).

The sequences from *Hord. europaeus* were compared to the sequences from other Triticeae species examined here. A 28 bp indel was observed at position 139-167, where the sequences from *Hord. europaeus* showed the same indel length with all the sequences from *Hordeum* species, but are different from the sequences in St, V, E, W, Y, P, Ta, and Ns genomes (Fig.6b). The second indel with about 100 bp was found among the sequences at position 186-282. All sequences from St genome except for *P. libanotica* have 38-39 bp insertion in this region compared to the sequences from *Hord. europaeus*, *Hordeum*, Y, W, P, Ns, Ta and E genome. The longest insertion (100 bp) in this region was found for the V genome of *D. villosum*. The sequence from *Hord. europaeus* showed

the same pattern as sequences from E, W, Y, P, Ta, Ns genomes and the sequences from *Hordeum* species. At position 346, a 30 bp deletion was observed in the sequences from most species in *Hordeum* except four H genome species: *H. patagonicum* subsp. *setifolium* (H1352), subsp. *santacrucense* (H1353), subsp. *magellanicum* (H1342), *H. pubiflorum* (H1236), and three I genome species: *H. vulgare* subsp. *vulagre* (H7514A), subsp. *spontaneous* (H3140A), and *H. bulbosum* (H3878). The 30 bp deletion was not detected in the sequences from *Hord. europaeus* and other Triticeae species examined here (Fig. 6c).

Figure 6. Example of partial alignment of the amplified sequences of *RPB2* from *Hordelymus* and other genomes in Triticeae.

a): The boxed region showed the indel between two distinct copies of sequences from *Hordelymus*. The indel was detected in two sequences from *Bromus* species. All other sequences from Triticeae do not have the indel.

b): The sequences from *Hord. europaeus* showed the same indel length with all the sequences from *Hordeum* genome except Xu genome in *Hordeum* which shows a indel with different length. This region in *Hordelymus* sequences is different from the sequences in St, V, E, W, Y, P, Ta, and Ns genomes (boxed region).

c): A 30 bp deletion was observed in the sequences from most species in Hordeum except four H genome species: *H. patagonicum* subsp. *setifolium* (H1352), subsp. *santacrucense* (H1353), subsp. *magellanicum* (H1342), H. *pubiflorum* (H1236), and three I genome species: *H. vulgare* subsp. *vulagre* (H7514A), subsp. *spontaneous* (H3140A), and H. *bulbosum* (H3878). The 30 bp deletion was not detected in the sequences from *Hord. europaeus* and other Triticeae species (boxed).

Figure 1

```
                  *       600         *       620         *       640         *       660
H2024     : CTGCCGGTGATATCTCCTATTAAAGATGAATCCAGAACGAAATCAGTGACTGTTGTAACTGGTAAAT-GACACGAAA-CTTTAACCTT  H  H. pusillum
NGB7217   : CCGCCGGTGATATCTCCTGTTAAAAACGAATCCAGAACGAAATCAGTGACTGTTGTAACTGGTAAAT-GACACGAAT-CTTTAACCTT  ?  Hord. europaeus
NGB9218   : CCGCCGGTGATATCTCCTGTTAAAAACGAATCCAGAACGAAATCAGTGACTGTTGTAACTGGTAAAT-GACACGAAT-CTTTAACCTT  ?  Hord. europaeus
NGB9218U  : CTGTAGGTGATATCTCCTGTTAAAAACAAATCCAGAACCAAATCAGTGACTGTTGTA----------ACCTT  ?  Hord. europaeus
CN32048   : CTGCCGGTGATATCTCCTGTTAAAAACAAATCCAGAACCAAATCAGTGACTGTTGTA----------ACCTT  Broums catharticus
CN30955   : CTGCCGGTGATATCTCCTGTTAAAAACAAATCCAGAACCAAAACAGTGACTGTTGTA----------ACCTT  Broums inermis
```

a

```
                100         *       120         *       140         *       160         *
H2024     : AATGACTTTCTTA--A-CATGTT---CACCATACATCTTGAAAAAG----------ATCTTTACTT  H  H. pusillum
H6479     : AATGCCTTTCTTA--A-CATGTT---CACCATACATCTTGAAAAAG----------ATCTTTACTT  H  H. muticum
NGB7217   : AA-GACTTTCTTA--A-CATGTT---CACCATACATCTTGAAAATG----------ATCTTTACTT  ?  Hord. europaeus
H7514A    : AATGACTTTCTTA--A-CATGTT---CACCATACATCTTGAAAATG----------ATCTTTACTT  I  H. vulgare subsp. vulgare
PI142012  : AATGGCTTTCTTCTTAAAATGTTCTATGCTAAAGAAATTCAAGATGTAC----------TATCCAAAATGATCTTTACTT  Ec  L. elongatum
PI531712  : AATGGCTTTCTTCTTAACATGTTCTATGCTAAAGAAATTCAAGATGTACATGTTTACCATACATCTTGAAAATGATCTTTACTT  Eb  T. bessarabicum
PI208075  : AATGACTTTCTTA---ACATGTTCTACGCTGAA-AAATTCAAGATGTACATGTTTACCATACATCTTGAAAATGATCTTTACTT  Ta  T. caput-medusae
PI533014  : AATGACTTTCTTA---GCATGTTCTATGCTAAAGAAATTCAAGATGTACATGTTTACCATACATCTTGAAAATGATCTTTACTT  W  Aust. retrofractum
PI383534  : AATGACTTTCTTA---ACATATTCTATGCTAAAGAAATTCAAG----------TCTTGAAAATGATCTTTACTT  P  Ag. cristatum
PI401274  : AATGACTTTCTTA--A-CATGTTCTATGCTAAAGAA-TTCAAGATGTATATGTTTACCATACATCTTGAAAATGATCTTTACTT  St  P. libanotica
NGB9218   : AATGACTTTCTTA--A-CATGTT---CACCATACATCTTGAAAATG----------ATCTTTACTT  ?  Hord. europaeus
NGB9218U  : AATGACTTTCTTA--A-CATGTT---CACCATACATCTTGAAAATG----------ATCTTTACTT  ?  Hord. europaeus
PI406496  : AA-GACTTTCTTA--A-CATGTTATATAGTAAAGAAATTCAAGATGTACAT----------GTTTACCGTACATCTTTACTT  Ns Psa. juncea
H74       : AA--ACATGTCTA--GATACATCCGTATCTAGACAAATCCAAGAC----------AGTAATATGAACCG  Xu  H. murinum subsp. glaucum
```

b

```
                *       340         *       360         *       380         *       400
H1236     : GGGCAACAAAATCACATCAGTTTCATT--AATGGAGTTTTTCCAATCTGTTAAATATACAAAATATGCAGATAATA  H  H. pubiflorum
H4014     : GGGCAACAAAATCACATCAGTT----------AAATATACAAAATATGCAGATAATA  H  H. bogdanii
H121      : GGGCAACAAAATCACAACAGTT----------AAATATACAAAATATGCAGATAATA  Xa  H. marinum subsp. marinum
NGB7217   : GGGCAACAAAATCACATCAGTTTTATT--AATGGAGTGCTCCCAATCTGTTAAATATACAAATTATGCAGATAATA  ?  Hord. europaeus
H7514A    : GGGCAACAAA-CCACATCAGTTTTATT--AATGGAGTGCTCCCAATCTGTTAAATATACAAAATATGCAGATAATA  I  H. vulgare subsp. vulgare
PI142012  : GGGCAACAAAATCACATCAGTTTTATT--AATGGAGTGCT-CCAATCTGTTAAATATACAAATTATGCAGACAATA  Ec  L. elongatum
PI531712  : GGGCAACAAAATCACATCAGTTTTATT--AATGGAGTGCT-CCAATCTGTTAAATATACAAATTATGCAGCTAATA  Eb  T. bessarabicum
PI208075  : GGGCAACAAAATCACATTAGTTTTTATTAAATGGAGTAATCCCAATCTGTTAAATATACAAATTATGCAGATAATA  Ta  T. caput-medusae
NGB9218   : GGGCAACAAAATCACATCAGTTTTATT--AATGGAGTGCTCCCAATCTGTCAAATATACAAATTATGCAGATAATA  ?  Hord. europaeus
NGB9218U  : GGGCAACAAAATCACATCAGTTTCATT--AATGGAGTACTCCCAATCCGTTACATATACAAATCATGCAGATGACA  ?  Hord. europaeus
```

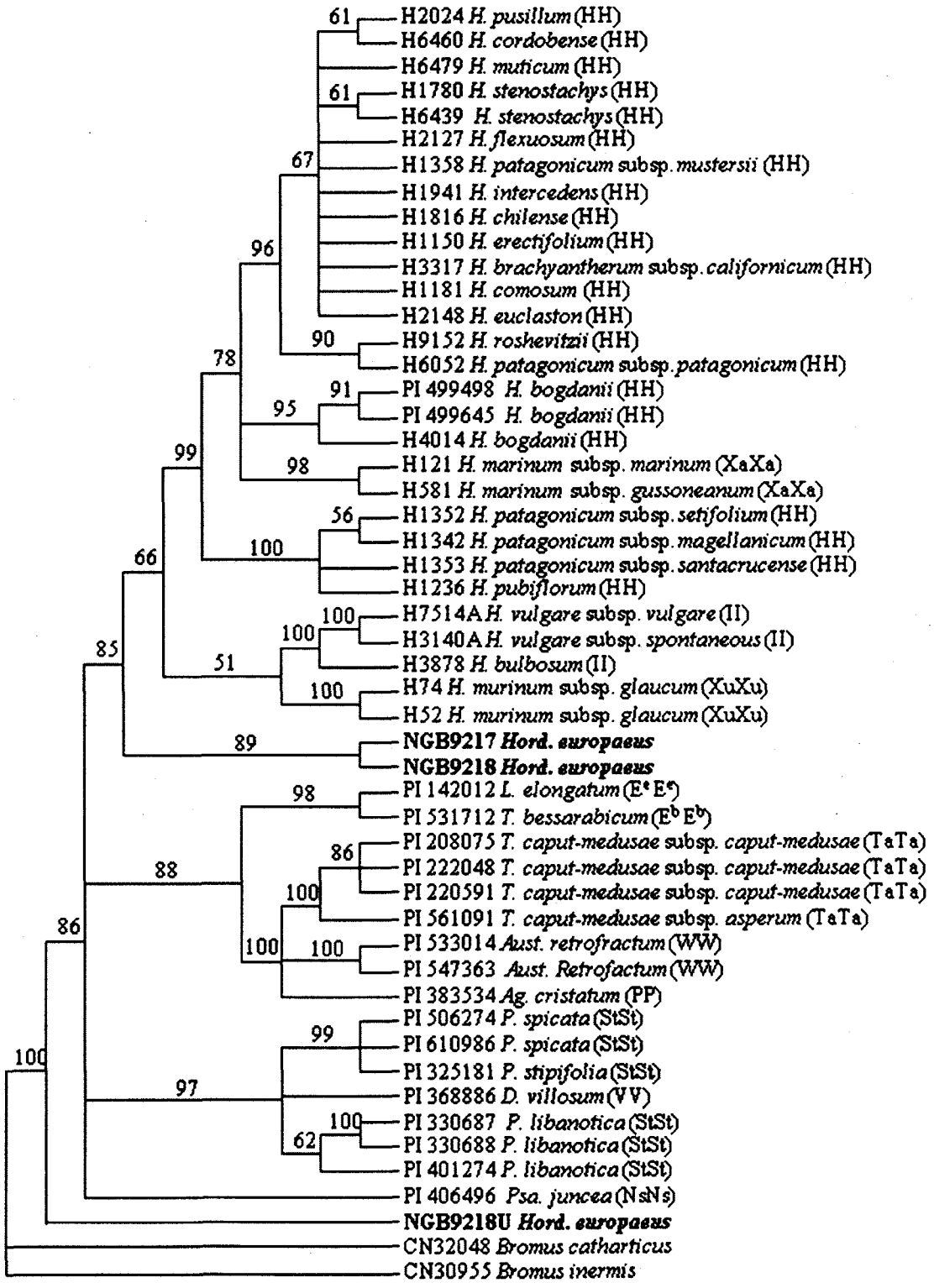
c

3.2. Phylogenetic analysis of the *RPB2* sequences

Maximum parsimony analysis was conducted using *B. catharticus* and *B. inermis* as outgroup. The data matrix contained 778 characters, of which 487were constant, 81 were parsimony uninformative, and 210 were parsimony informative. As a result, 456 trees were generated with the tree length of 456 steps, a consistency index (CI) of 0.798, retention index (RI) of 0.923. Strict consensus tree was constructed from these trees as shown in Figure 2 with bootstrap value. The number of bootstrap replicate is 1000.

Parsimony analysis grouped the two sequences from two accessions of *Hord. europaeus* (NGB9217 and NGB9218) together with 89% bootstrap support. The two sequences are strongly supported as sister to the sequences from *Hordeum* species (85% bootstrap support). Whereas another copy of sequence from NGB9218 is the sister to all Triticeae with very strongly support (bootstrap value 100%). As expected, all sequences from *Hordeum* species were grouped together. The sequences from Ta genome was put into a strongly support group (100% bootstrap value). Whereas Ns genome form it own clade. Results were as shown is Figure 7.

Figure 7. Strict consensus tree of the 456 most parsimonious trees derived from *RPB2* sequence data was conducted using heuristic search with TBR branch swapping. Two *Bromus* species were used as outgroup. Consistency Index (CI) = 0.798, retention index (RI) = 0.923.


H2024 H. pusillum (HH)
H6460 H. cordobense (HH)
H6479 H. muticum (HH)
H1780 H. stenostachys (HH)
H6439 H. stenostachys (HH)
H2127 H. flexuosum (HH)
H1358 H. patagonicum subsp. mustersii (HH)
H1941 H. intercedens (HH)
H1816 H. chilense (HH)
H1150 H. erectifolium (HH)
H3317 H. brachyantherum subsp. californicum (HH)
H1181 H. comosum (HH)
H2148 H. euclaston (HH)
H9152 H. roshevitzii (HH)
H6052 H. patagonicum subsp. patagonicum (HH)
PI 499498 H. bogdanii (HH)
PI 499645 H. bogdanii (HH)
H4014 H. bogdanii (HH)
H121 H. marinum subsp. marinum (XaXa)
H581 H. marinum subsp. gussoneanum (XaXa)
H1352 H. patagonicum subsp. setifolium (HH)
H1342 H. patagonicum subsp. magellanicum (HH)
H1353 H. patagonicum subsp. santacrucense (HH)
H1236 H. pubiflorum (HH)
H7514A H. vulgare subsp. vulgare (II)
H3140A H. vulgare subsp. spontaneous (II)
H3878 H. bulbosum (II)
H74 H. murinum subsp. glaucum (XuXu)
H52 H. murinum subsp. glaucum (XuXu)
NGB9217 Hord. europaeus
NGB9218 Hord. europaeus
PI 142012 L. elongatum (E^e E^e)
PI 531712 T. bessarabicum (E^b E^b)
PI 208075 T. caput-medusae subsp. caput-medusae (TaTa)
PI 222048 T. caput-medusae subsp. caput-medusae (TaTa)
PI 220591 T. caput-medusae subsp. caput-medusae (TaTa)
PI 561091 T. caput-medusae subsp. asperum (TaTa)
PI 533014 Aust. retrofractum (WW)
PI 547363 Aust. Retrofactum (WW)
PI 383534 Ag. cristatum (PP)
PI 506274 P. spicata (StSt)
PI 610986 P. spicata (StSt)
PI 325181 P. stipifolia (StSt)
PI 368886 D. villosum (VV)
PI 330687 P. libanotica (StSt)
PI 330688 P. libanotica (StSt)
PI 401274 P. libanotica (StSt)
PI 406496 Psa. juncea (NsNs)
NGB9218U Hord. europaeus
CN32048 Bromus catharticus
CN30955 Bromus inermis

4. Discussion

Cytogenetic evidence suggested that *Hor. europaeus* is an alloploid species with two distantly related genomes (Bothmer *et al.*, 1994). But very recently it was suggested that *Hord. europaeus* is autopolyploid with only Ns genome (Ellneskog-Staam *et al.*, 2006). The nuclear *DMC*1 and *EF-G* gene data supported that it is an alloploid origin (Petersen and Seberg, 2008). The origin of *Hordelymus* genome has been debated academically for years, and no consensus conclusion was reached. The present phylogenetic study of single copy nuclear gene *RPB*2 shed some light on this issue, and strongly supports that *Hord. europaeus* is allopolyploid. A close relationship between H genome and one copy of *Hordelymus* genome was observed. No obvious relationship was found between *Hordelymus* and either Ta genome donor, *T. caput-medusae* species or Ns genome donor, *Psa. juncea*.

4.1. The relationship of *Hordeum* and *Hordelymus* genome

Cytogenetic data have shown that the diploid *Hordeum* species can be divided into four monogenomic groups: the I-genome group (*H. vulgare* and *H. bulbosum*), the Xa genome group (*H. marinum*, formerly X), the Xu genome group (*H. murinum*; formerly Y), and the H genome group (the remaining diploid species) (Bothmer *et al.*, 1986, 1987). The genetic relationship between H genome species in *Hordeum* and *Hordelymus* was suggested in previous studies. Löve (1984) suggested that *Hord. europaeus* contains the H genome from *Hordeum* and the Ta genome from *Taeniatherum*. Chromosome pairing data of hybrid meiosis suggested one genome in common between *Hord. europaeus* and *H. depressum* (Bothmer and Jacobsen, 1989). However, this suggestion had been later

proved to be based on a wrong interpretation of the meiotic data since later it was found that *H. depressum* is autoploid species (Petersen, 1991), and no homolog between *Hord. europaeus* and *Hordeum* species were proposed (Bothmer *et al.*, 1994). Our *RPB2* data showed the strong affinity between the genomes in *Hordeum* species and *Hordelymus* in the phylogeny tree with 85% bootstrap support. These results suggested that one genome in tetraploid *Hord. europaeus* is closely related to the genome in *Hordeum*. A close relationship between *Hord. europaeus* populations and some populations of *E. caninus* (StH genome) was reported in a AFLP analysis (Minianty *et al.*, 2006). The disagreement between all the previous studies and this one could be explained if the *Hordeum* copy genome in *Hordelymus* has been experiencing a different evolution rate between the polyploid species and its donor diploid species, which led to the failure of intergenetic crosses and chromosome pairing. Variable rates of evolution indeed were found between the loci in diploid and allotetraploid species (Small and Wendel, 2000; Barrier *et al.*, 2001; Caldwell *et al.*, 2004; Sun *et al.*, 2007).

4.2. *Hordelymus* and Ta (*Taeniatherum*) , Ns (*Psathyrostachys*), the unknown genome

Karyotype and Giemsa C-banding pattern suggested that one of the *Hord. europaeus* genomes may be a Ta genome derived from *Taeniatherum*, and the other possibly an Ns genome derived from a *Psathyrostachys* species (Bothmer *et al.*, 1994). Unfortunately, no information about hybridizations of *H. europaeus* and *Psathyrostachys* ssp., and *Taeniatherum* has been published. One major study among few previous literatures for

the hypothesis of the presence of Ta genome in fact did not include Ta genome species directly in their intergeneric materials and did not confirm the presence of Ta genome chromosomes (Bothmer *et al.*, 1994). The relationship of *Hord. europaeus* and Ns genome was suggested on the base of RAPD method research (Svitashev *et al.*, 1998). It is well known, mutations at the priming site could result in the gain or loss of RAPD bands. And the concerns of the suitability of RAPD assay for determining parentage or origin has been raised in past few years (Mullis, 1991; Scott *et al.*, 1992; Riedy *et al.*, 1992; Gwakisa, 1994). Another concern on the RAPD data of Svitashev *et al.*, (1998) is that the genome specific RAPD marker did amplify the DNA from other genome species, for example, the Y genome specific marker amplified the DNA from cytogenetically confirmed H genome (*H. bogdanii*). Recently, genomic southern hybridization data indicated that the presence of an Ns genome highly homologous to that of *Psathyrostachys* and *Leymus* and rejected the presence of H, Ta, E or St genomes in *Hord. europaeus* (Ellneskog-Staam *et al.*, 2006). The data from the nuclear *DMC*1 gene showed that one copy of *Hordelymus* sequences as the sister to *Psathyrostachys*, but data from the nuclear *EF-G* gene do not place any of the two *Hordelymus* sequence copies as sister to *Psathyrostachys* (Petersen and Seberg, 2008). Our *RPB2* sequence data confirmed the absence of Ta genome in *Hord. europaeus*, and did not support the presence of a Ns genome in *Hordelymus* species suggested by previous studies (Löve 1984; Bothmer *et al.*, 1994; Ellneskog-Staam *et al.*, 2006). Two distinct copies of sequences were discovered from NGB9218 accession of *Hord. europaeus*. Phylogenetic analysis suggested that one genome in *Hord. europaeus* is sister to the Triticeae genomes studied here, and

corresponds well to the *EF-G* result of Petersen and Seberg (2008). The sequence NGB9128U shows 89% identical with two *Bromus* sequences which is the highest score among the pairwise comparison between NGB9128U and any other sequences in our study. Previous RFLP analysis of tandemly repeated DNA sequences in Triticeae also showed that *Hord. europaeus* and the all species outside tribes Triticeae did not hybridize to dpTa1 even at low stringency (Vershinin *et al.*, 1994). These results suggested that a direct contribution from another unknown genome donor outside the tribe Triticeae to *Hordelymus* remains a possibility. Another possibility could be that the origin of the donor might have extinct already, just like the Y genome in *Elymus* species, which the diploid Y donor was not discovered so far (Wang *et al.*, 1994). Our *RPB2* data does not support the hypothesis that *Hordelymus europaeus* is an autotetraploid (Ellneskog-Staam *et al.*, 2006), but confirms that it is alloploid origin.

Actually it is no wonder a complex phylogenetic history of genus in Triticeae, therefore *Hordelymus* could be another example to confirm its complexity. A well studied example is hexaploid *Elymus repens* (L) Gould, a widespread, morphologically variable species in the Triticeae tribe, in which three distinct genome donors including one unknown that was apparently derived from outside of the tribe, and introgression are discovered. At least three level of reticulate evolution have shaped the genome of it (Mason-Gamer 2008). It is possible that *Hord. europaeus* might have a similar evolutionary phenomenon discovered in *E. repens*. So far, molecular phylogenetic data fail to identify the second genome in *Hord. europaeus*. Its clarification awaits a more

detailed analysis including more diploid species within and outside Triticeae with more sequence data.

4.3. Summary

The molecular phylogenetic analyses of *Hordelymus* by *RPB2* gene uncover a close relationship between *Hordeum* and one genome in *Hord. europeaus*. Absence of Ta and Ns is convincing here. Phylogenetic analysis suggested that one genome in *Hord. europaeus* is sister to the Triticeae genomes studied here, and corresponds well to the *EF-G* result of Petersen and Seberg (2008). A direct contribution from another unknown genome donor within or outside the tribe Triticeae to *Hordelymus* remains a possibility. Our *RPB2* data does not support the hypothesis that *Hordelymus europaeus* is an autotetraploid (Ellneskog-Staam *et al.*, 2006), but confirms that it is alloploid origin

PART 3. References

Agafonov, A.V., 2004. Intraspecific structure and reproductive relationships between *Elymus mutabilis* and *E. transbaicalensis* (Poaceae) in southern Siberia from the viewpoint of taxonomical genetics. Russian Journal of Genetics 11, pp. 1229–1238.

Baden, C., Frederiksen, S., Seberg, O., 1997. A taxonomic revision of the genus *Hystrix* (Triticeae, Poaceae). Nordic Journal of Botany 17, pp. 449–467.

Bailey, C.D., Doyle, J.J., 1999. Potential phylogenetic utility of the low-copy nuclear gene pistillata in dicotyledonous plants: comparison to nrDNA ITS and trnL intron in Sphaerocardamum and other Brassicaceae. Molecular Phylogenetics and Evolution. 13, pp. 20-30.

Barrier, M., Robichaux, R.H., Purugganan, M.D., 2001. Accelerated regulatory gene evolution in an adaptive radiation. Proceeding of the National Academy of Science of the United States 98, pp. 10208-10213.

Bothmer, V.R., Fink, J., Landström, T. ,1986.Meiosis in interspecific *Hordeum* hybrids. I. Diploid combinations. Canadian Journal of Genetics and Cytology 28, pp.525–535.

Bothmer, V. R., Fink, J., Landström, T., 1987. Meiosis in *Hordeum* interspecific hybrids. II. Triploid hybrids. Evolution Trends of Plants 1, pp. 41–50.

Bothmer, V.R., Jacobsen, N., 1989. Intergeneric hybirdization between *Hordeum* and *Hordelymus* (Poaceae). Nordic Journal of Botany 9, pp.113-117.

Bothmer, V. R, Lu, B.R., Linde-Laursen, I.B., 1994. Intergeneric hybridation and C-banding pattern in *Hordelymus* (Triticeae, Poaceae). Plant Systematics and Evolution 189, pp.259-266.

Bothmer, V.R., Jacobsen, N., Baden, C., Jørgensen, R.B., Linde-Laursen, I., 1995. An ecogeographical study of the genus *Hordeum*, 2nd ed. International Plant Genetic Resources Institute, Rome.

Caldwell, K.S., Dvorak, J., Lagudah, E.S., Akhunov, E., Luo, M.C., Wolters, P., Powell,W., 2004. Sequence polymorphism in polyploid wheat and their D-genome diploid ancestor. Genetics 167, pp. 941-947.

Claw, B., Signe, F., Ole, S., 1997. A taxonomic revision of the genus *Hystrix* (Triticeae, Poaceae). Nordic Journal of Botany 17, pp. 449-467.

Cronn, R.C., Small, R.L., Haselkon, T., Wendel, J.F., 2003. Cryptic reeated genomic recombination during speciation in Gossyptum *gossypioides*. Evolution 57, pp. 2475-2489.

Denton, A.L., McConaughy, B.L., Hall, B.D., 1998. Usefulness of RNA Polymerase II coding sequences for estimation of green plant phylogeny. Molecular Biology of Evolution 15, pp. 1082-1085.

Dewey, D.R., 1982. Genomic and phylogenetic relationships among North American perennial Triticeae. In Grasses and grasslands. Edited by J.R. Estes. Oklahoma University Press, Norman, Okla. pp. 51–88.

Dewey, D.R., 1983*a*. New nomenclatural combinations in the North American perennial Triticeae. Brittonia 35, pp. 30–33.

Dewey, D.R., 1984. The genomic system of classification as a guide to intergeneric hybridization within the perennial Triticeae. In: Gene Manipulation in Plant Improvement, Proceeding of the 16th Stadler Genetics Symposium. Edited: J.P. Gustafsen. New York: Plenum Press. pp, 209-279.

DiRuggiero, J., Brown, J.R., Bogert, A.P., Robb, F.T., 1999. DNA repair systems in Archaea: Mementos from the last universal common ancestor? Journal of Molecular Evolution 49, pp. 474–484.

Doutriaux, M.P., Couteau, F., Bergounioux, C., White, C., 1998. Isolation and characterisation of the RAD51 and DMC1 homologs from *Arabidopsis thaliana*. Molecular Genetics 257, pp. 283–291.

Doyle, J.J., Doyle, J. L., Brown, H.D., 1999 Incongruence in the diploid B- genome species complex of *Glycine* (Leguminosae) revised: histone H3-D alleles va. Chloroplast haplotypes. Molecular of Biology and Evolution 16, pp. 354-362.

Dubcovsky, J., Schlatter, A.R., Echaide, M., 1997. Genome analysis of South American *Elymus* (Triticeae) and *Leymus* (Triticeae) species based on variation in repeated nucleotide sequences. Genome 40, pp. 505-520.

Ellneskog-Staam, P., Taketa, S., Salomon, B., Anamthawat-Jónsson, K., Bothmer, V. R.,2006. Identifying the genome of wood barley *Hordelymus europaeus* (Poaceae: Triticeae). Hereditas 143, pp.103-112.

Ellenskog-Staam, P., von Bothmer R., Anamthawat-Jo′ nsson K., Salomom, B., 2007. Genome analysis of species in the genus *Hystrix* (Triticeae; Poaceae). Plant Systematics and Evolution 265, pp. 241–249.

Fan, X., Zhang, H.Q., Sha, L.N., Zhang, L., Yang, R.W., Ding, C.B., Zhou, Y.H., 2007 Phylogenetic analysis among *Hystrix*, *Leymus* and its affinitive genera (Poaceae: Triticeae) based on the sequences of a gene encoding plastid acetyl-CoA carboxylase Plant Science 172, pp. 701–707.

Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, pp. 783–791.

Ford, V. S. Gottlieb, L. D., 1999. Molecular characterization of *PgiC* in a tetraploid plant and its diploid relatives. Evolution 53, pp. 1060–1067.

Gaut, B.S., 1998. Molecular clocks and nucleotide substitution rates in higher plants. Evolutionary Biology 30, pp. 93–120.

Ge, S., Sang, T., Lu, B.R., Hong, D.Y., 1999. Phylogeny of rice genomes with emphasis on origins of allotetraploid species. Proceedings of National Academy of Science 96, pp. 14400–14405.

Gwakisa, P.S., Kemp, S.J., Teale, A.J., 1994.Characterization of zebu cattle breeds in Tanzania using random amplified polymorphic DNA markers. Animal Genetics 25, pp. 89–94.

Grant, V., 1981. Plant Speciation (Columbia Univ. Press, New York), 2nd Ed.

Harris, S.A., 1999. RAPDs in systematic- a useful methodology Molecular systematic and Plant Evolution, the systematics (Edited by Peter M. H., Richard M. B, Richard J.G.) Association Special Volome Series 57 Taylor & Francis Limited.pp. 211-228.

Hartmann, S., Nason, J.D., Bhattacharya, D., 2001. Extensive ribosomal DNA genic variation in the columnar cactus *Lophocereus*. Journal of Molecular Evolution 53, pp. 124-134.

Hedrick, P. W., 1987. Genetic load and the mating system in homosporous ferns. Evolution 41, pp. 1282–1289.

Helfgott, D.M., Mason-Gamer, R.J., 2004. The evolution of North American *Elymus* (Triticeae, Poaceae) allotetraploids: evidence from phosphoenolpyruvate carboxylase gene sequences. Systematic Botany 29, pp. 850–861.

Hitchcock, A.S., 1951. Tribe 3. Hordeae. In: Manual of the Grasses of the United States. U.S. Government Printing Office, Washington, D.C. pp. 230–280.

Hsiao, C., Chatterton, N. J., Asay, K. H. Jensen, K. B., 1995. Molecular phylogeny of the Pooideae (Poaceae) based on nuclear rDNA (ITS) sequences. Theoretical and Applied Genetics 90, pp. 389–398.

Hsiao, C., Chatterton, N.J., Asay, K.H., Jensen, K.B. 1995. Phylogenetic relationships of the monogenomic species of the wheat tribe, Triticeae (Poaceae), inferred from nuclear rDNA (Internal transcribed spacer) sequences. Genome 38, pp.221-223.

Jensen, K.B.1990a. Cytology, fertility, and morphology of *Elymus kengii* (Keng) Tzvelev and *E. grandiglumis* (Keng) A. Löve (Triticeae: Poaceae). Genome 33, pp. 563-570.

Jensen, K.B.1990b. Cytology and taxonomy of *Elymus kengii, E. grandiglumis, E. alatavicus*, and *E. batalinii* (Poaceae: Triticeae). Genome 33, pp. 668-673.

Jensen, K.B., 1993. Cytology of *Elymus magellanicus* and its interspecific hybrids with *P. spicata, H. violaceum, E. subsecundus, E. lanceolatus,* and *E. glaucus* (Poaceae: Triticeae). Genome 36, pp. 72–76.

Jensen, K.B., Salomon, B., 1995. Cytogenetics and morphology of *Elymus panormitanus* var. heterophyllus (Keng) Á.Löve and its relationship to *Elymus panormitanus* (Poaceae: Triticeae). International Journal of Plant Science 156, pp. 731–739.

Jensen, K.B., Wang R.R.C., 1997. Cytological and molecular evidence for transferring *Elymus coreanus* from the genus *Elymus* to *Leymus* and molecular evidence for *Elymus californicus* (Poaceae: Triticeae). International Journal of Plant Science 158, pp. 872–877.

Junghans, H., Metzlaff, M., 1990. A simple and rapid method for the preparation of total plant DNA. Biotechniques 8, pp. 176.

Kellogg, E. A., Appels, R., 1995. Intraspecific and interspecific variation in 5S RNA genes are decoupled in diploid wheat relatives. Genetics 140, pp. 325–343.

Klimyuk, V. I., Jones, J. D. G., 1997. A tDMC1, the *Arabidopsis* homologue of the yeast DMC1 gene: Characterization, transposon-induced allelic variation and meiosis-associated expression. Plant Journal 11, pp. 1–14.

Kolodziej, P.A., Woychik, N., Liao, S., Young, R.A., 1990. RNA Polymerase II subunit composition, Stoichiometry, and Phosphorylation. Molecular and Cellular Biology 10, pp. 1915-1920.

Linde-Laursen, I., Seberg, O., Salomon, B., 1994. Comparison of the Giemsa C-banded and N-banded karyotypes of two *Elymus* species, *E. dentatus* and *E. glaucescens* (Poaceae: Triticeae). Plant Systematics and Evolution 192, pp. 165–176.

Liu, Q., Ge, S., Tang, H,, Zhang, X.L., Zhu, G.F., Lu, B.R., 2006. Phylogenetic relationships in *Elymus* (Poaceae: Triticeae) based on the nuclear ribosomal internal transcribed spacer and chloroplast trnL-F sequences. New phytologist 170, pp. 411–420.

Liu, Q., Brubaker, C.L., Green, A.G., Marshall, D.R., Sharp, P.J., Singh, S.P., 2001. Evolution of the FAD2-1 fatty acid desaturase 5' UTR intron and the molecular systematics of *Gossypium* (Malvaceae). American Journal of Botany 88, pp. 92-102.

Löve, Á. 1984. Conspectus of the Triticeae. Feddes Reppert. 95, 425–521.

Lu, B. R., Bothmer, R. von., 1989. Cytological studies of a dihaploid and hybrid from intergeneric cross *Elymus* shandongensis. Triticum aestivum. Heredity 111, pp. 231-238.

Lu, B. R., Björn, S., 1992. Differentiation of the SY genomes in Asiatic *Elymus.* Heredity. 116, pp. 121-126.

Lu. B. R., 1993a. Genomic relationships within the *Elymus* parviglumis group (Triticeae: Poaceae). Plant Systematics and Evolution 187, pp. 191–211.

Lu, B. R., 1993b. Biosystematic Investigations of Asiatic Wheatgrass-*Elymus* L. (Triticeae: Poaceae) The Swedish University of Agricultural Sciences. Department of Plant Breeding Research, Dissertation, Svalöv. Printed by Förvaltning savdelningens repro. Alnarp, Sweden.

Lu, B.R., 1994. The genus *Elymus* L. in Asia. Taxonomy and biosystematics with special reference to genomic relationships. In: Wang RRC, Jensen KB, Jaussi C, eds. Proceedings of the 2nd International Triticeae Symposium, Logan, UT, USA. Logan, UT, USA: The Utah State University Press. pp. 219–233.

Magnus, P., Bengt, O., 2001. Inferring the history of the polyploid *Silene aegaea* (Caryophyllaceae) using plastid and homoeologous nuclear DNA sequences. Molecular Phylogenetics and Evolution 20, pp. 474–481.

Magnus, P., Bengt, O., 2007. Origin and Evolution of North American Polyploid *Silene* (*Carryophyllaceae*). American Journal of Botany 94, pp. 330–349.

Masterson, J., 1994. Stomatal size in fossil plants: evidence for polyploidy in majority of angiosperms. Science 264, pp. 421–424.

Mason-Gamer, R.J., 2001. Origin of North America *Elymus* (Poaceae:Triticeae) allotetraploids based of granule-bound starch synthase gene sequences. Systematic Botany 26, pp. 757–758.

Mason-Gamer, R.J., Orme, N.L., Anderson, C.M., 2002. Phylogenetic analysis of North American *Elymus* and the monogenomic Triticeae (Poaceae) using three chloroplast DNA data sets. Genome 45, pp. 991-1002.

Mason-Gamer, R.J., Burns, M.M., Naum, M., 2005. Polyploidy, introgression, and complex phylogenetic patterns within *Elymus*. Czech Journal of Genetics and Plant Breeding 41, pp. 21–26.

Mason-Gamer, R.J. 2008. Allohexaploidy, introgression, and the complex phylogenetic history of *Elymus repens* (Poaceae). Molecular Phylogenetics and Evolution 47, pp.

598-611.

McMillan, E., Sun, G.L., 2004. Genetic relationships of tetraploid *Elymus* species and their genomic donor species inferred from polymerase chain reaction-restriction length polymorphism analysis of chloroplast gene regions Theoretical and Applied Genetics 108, pp. 535-542.

Mizianty, M., Bieniek, W., Czech, A., Strzalka, W., Szklarczyk, M., 2006. Variability and structure of natural populations of *Elymus caninus* (L.) L. and their possible relationship with *Hordelymus europaeus* (L.) Jess. ex Harz as revealed by AFLP analysis. Plant Systematics and Evolution 256, pp. 193-200.

Mullis, K.B.,1991. The polymerase chain reaction in an anemic mode: how to avoid cold oligodeoxyribonuclear fusion. PCR Method Application 1, pp. 1–4

Naohiro, I. 1998. All-In-One Seq-Analyzer Version 1.36 for Netscape Communicator ver 4.05, MS Explorer 5 or later. http://www-personal.umich.edu/~ino/blast.htmL. Copy right Mar/18/1998.

Olmstead R.G., Pamer J.D., 1994. Chloroplast DNA systematics; a review of methods and data analysis. American Journal of Botany 81, pp. 1205-1224.

Osada, T., 1993. Illustrated Grasses of Japan, Enlarged edition. Heibonsha Ltd, Press, Tokyo.

Pelger, S., 1993. Prolamin variation and evolution in Triticeae as recognized by monoclonal antibodies. Genome 36, pp. 1042-1048.

Peter, M. H., Richard M. B, Richard J.G., 1999. Molecular systematic and Plant Evolution , the systematics Association Special Volome Series 57 Taylor & Francis Limited.

Petersen, G., Seberg, O., 2000. Phylogenetic evidence for excision of Stowaway miniature inverted-repeat transposable elements in Triticeae (Poaceae). Molecular Biology and Evolution 17, pp. 1589–1596.

Petersen, G., Seberg, O., 2002. Molecular evolution and phylogenetic application of DMC1. Molecular Phylogenetics and Evolution 22, pp. 43–50.

Petersen, G., Seberg, O., 2003. Phylogenetic analyses of the diploid species of *Hordeum* (Poaceae) and a revised classification of the genus. Systematic Botany 28, pp. 293–306.

Petersen, G., Seberg, O., 2004a. On the Origin of the Tetraploid species *Hordeum capense* and *H. secalinum* (Poaceae). Systematic Botany 29, pp. 862-873.

Petersen, G., Seberg, O., 2004b. A phylogenetic analysis of the genus *Psathyrostachys* (Poaceae) based on one nuclear gene, three plastid genes, and morphology. Plant Systematics and evolution 249, pp. 99-110

Petersen, G., Seberg, O., 2008. Phylogenetic relationships of allotetraploid *Hordelymus* europaeus (L.) Harz (Poaceae: Triticeae). Plant Systematic Evolution 273, pp. 87–95

Petersen, S., 1991. Intergeneric hybridization between *Hordeum* and *Secale*. II. Analysis of meiosis in hybrids. Hereditas 114, pp. 141-159.

Rauscher, J.T., Doyle, J.J., Brown, A.H., 2004. Multiple origins and nrDNA internal transcribed spacer homoeologue evolution in the *Glycine tomentella* (Leguminosae) allopolyploid complex. Genetics 166, pp. 987–998.

Riedy, M.F., Hamilton, W.J., Aquadro, C.F., 1992. Excess of nonparental bands in offspring from known primate pedigrees assayed using RAPD PCR. Nucleic Acids Research 20, pp. 918-925.

Saitou, N., Nei, M., 1987. The neighbour-joining method: a new method for reconstructing phylogenetic trees. Journal of Molecular Evolution 4, pp. 406–425.

Sang, T., 2002. Utility of Low-Copy Nuclear Gene Sequences in Plant Phylogenetics. Critical Reviews in Biochemistry and Molecular Biology 37, pp. 121–147.

Scott, M.P., Haymes, K.M., Williams, S.M., 1992. Parentage analysis using RAPD PCR. Nucleic Acids Research 20, pp.5493-5501.

Segraves, K.A., Thompson, J.N., Soltis, P.S., and Soltis, D.E., 1999. Multiple origins of polyploidy and the geographic structure of *Heuchera grossulariifolia.* Molecualr Ecology 8, pp. 253–262.

Small, R.L., J.A. Ryburn, R.C. Cronn, T. Seelanan, J.F. Wendel. 1998. The tortoise and the hare: choosing between noncoding plastome and nuclear Adh sequences for phylogeny reconstruction in a recently diverged plant group. American Journal of Botany 85, pp. 1301-1315.

Small, R.L., Wendel, J.F., 2000. Phylogeny, duplication, and intraspecific variation of *Adh* sequences in New World diploid cottons (*Gossypium* L., Malvaceae). Molecular Phylogenetics and Evolution 16, pp. 73-84.

Small, R.L., Cronn, R. C., Wendal, J. F., 2004 Use of nuclear genes for phylogeny reconstruction in plants. Australian Systematic Botany 17, pp. 145-170.

Soltis, D.E., Soltis, P.S., 1993. Molecular data and the dynamic nature of polyploidy. Critical Reviews in Plant Science 12, pp. 243–273.

Soltis, D.E., Soltis, P.S., 1998. Choosing an approach and an appropriate gene for phylogenetic analysis. In "Molecular systematic of plants II : DNA sequencing". Kluwer Academic Publisher, Boston, pp.1-42.

Soltis, D.E., Soltis, P.S., 1999.Polyploidy: recurrent formation and genome evolution. Trends in Ecology and Evolution 14, pp. 348–392.

Soltis, D.E., Soltis, P.S., 2000a. Contributions of plant molecular systematics to studies of molecular evolution. Plant Molecular Biology 42, pp. 45–75.

Soltis, D.E., Soltis, P.S., 2000b. The role of genetic and genomic attributes in the success of polyploids. Proceedings of National Academy of Science 97, pp. 7051–7057.

Stassen, N. Y., Logsdon, J. M., J., Vora, G. J., Offenberg, H. H., Palmer, J. D., Zolan, M. E., 1997. Isolation and characterization of rad51 orthologs from Coprinus cinereus and Lycopersicon esculentum, and phylogenetic analysis of eukaryotic recA homologs. Current Genetics 31, pp. 144–157.

Stebbins, G. L., 1950. Variation and Evolution in Plants (Columbia Univ. Press, New York).

Suh,Y., Thien, L.B., Reeve, H.E., Zimmer, E.A., 1993. Molecular evolution and phylogenetic implications of internal transcribed spacer sequences of ribosomal DNA in Winteraceae. American Journal of Botany 80, pp. 1042-1055.

Sun, G.L., 1997. Analysis of tetraploid *Elymus* species using wheat microsatellite markers and RAPD markers. Genome 40, pp. 806-814.

Sun, G.L., Tracy D., Ni, Y., 2007. Molecular evolution and genome divergence at RPB2 gene of the St and H genome in *Elymus* species. Plant Molecular Biology 64, pp. 645–655.

Sun, G.L., Ni, Y., Tracy, Daley, 2008. Molecular phylogeny of RPB2 gene reveals multiple origin, geographic differentiation of H genome, and the relationship of the Y genome to other genomes in *Elymus* species. Molecular Phylogenetics and Evolution 46, pp. 897–907.

Sun, G.L., Pourkheirandish, M., Komatsuda, T., 2009. Molecular evolution and phylogeny of *RPB*2 gene in the genus *Hordeum.* Annals of Botany 103, pp. 975-983.

Svitashev, S., Bryngelsson, T., Li, X. M., Wang, R., 1998. Genome-specific repetitive DNA and RAPD markers for genome identification in *Elymus* and *Hordelymus* Genome, 41, 120-128.

Sweetser, D., Nonet, M., Young, R.A., 1987. Prokaryotic and eukaryotic RNA polymerases have homologous core subunits. Proceeding of the National Acadamy of Science of the United States of America 84, pp. 1192-1196.

Swofford, 2003. PAUP*: Phylogenetic Analyses using Parsimony (* and other methods), Version 4.0b10, Sinauer, Sunderland.

Thompson J.D., Gibson, T.J., Plewniak,b F., Jeanmougin, F., Higgins, D.G.. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research 25, pp. 4876–4882.

Torabinejad, J., Mueller, R. J., 1993 Genome constitution of the Australian hexaploid grass *Elymus* scabrus (Poaceae: Triticeae) Tzvelev, N.N.,1976. Poaceae USSR. Editio Nauka, Leningrad.

Vershinin, A., Svitashev, S., Gummesson, P.O., Salomon, B., Bothmer, R., Bryngelsson, T., 1994 Characterization of a family of tandemly repeated DNA sequences in Triticeae. Theoretical and Applied Genetics 89, pp. 217-225.

Wang, R.R.C., Dewey, D.R., Hsiao, C. 1986. Genome analysis of the tetraploid *Pseudoroegneria tauri*. Crop Science 26, pp. 723–727.

Wang, R.R.C., von Bothmer R., Dvorak, J., Fedak, G., Linde-Laursen, I., Muramatsu, M., 1994. Genome symbols in the Triticeae. In: Wang RR-C, Jensen KB, Jaussi C (eds) Proceedings of the 2nd international Triticeae symposium, June 20–24, 1994, Logan, Utah. Utah State University Publication Design and Production, Logan, pp. 29–34.

Watson, L., Dallwitz, M.J., 1992, The grass genera of the world. CAB International Wallingford

Wendel, J.F., 2000. Genome evolution in polyploids. Plant Molecular Biology. 42, pp. 225–249.

Wolf, P.G., Soltis, P.S., Soltis, D.E., 1989. Tetrasomic inheritance and chromosome pairing behaviour in the naturally occurring autotetraploid *Heuchera grossulariifolia* (Saxifragaceae). Genome 32, pp. 655–659.

Wolfe, K.H., Li, W. H., Sharp, P. M., 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proceedings of National Academy of Science U.S.A. 84, pp. 9054-9058.

Xu, D.H., Ban, T., 2004. Phylogenetic and evolutionary relationships between *Elymus humidus* and other *Elymus* species based on sequencing of non-coding regions of cpDNA and AFLP of nuclear DNA. Theoretical and Applied Genetics 108, pp. 1443–1448.

Zhang, H.Q., Fan, X., Sha, L.N, Zhang, C., Yang, R.W., Zhou, Y.H., 2008. Phylogeny of *Hystrix* and related genera (Poaceae: Triticeae) based on nuclear rDNA ITS sequences. Plant Biology 10, pp. 635–642.

Zhang, X.Q., Salomon, B., von Bothmer, R., Díaz, O., 2000. Patterns and levels of genetic differentiation in North American populations of the Alaskan wheatgrass complex. Heredity 133, pp. 123-132.