Christopher A. Cullis   *Editor*

Genetics and Genomics of Linum

Springer

Plant Genetics and Genomics: Crops and Models

Volume 23

Series Editor

Richard A. Jorgensen
University of Arizona
Tucson, AZ, USA

Plant Genetics and Genomics: Crops and Models book series provides current overviews and summaries of the state of the art in genetics and genomics for each of the important crop plants and genetic models for which such a volume does not now exist or is out of date.

Most volumes will focus on a single crop, species, or group of close relatives, including especially those plants that already have advanced genomic resources developed and preferably complete or advancing genome sequences. Potential targets include sorghum, maize, poplar, the Brassicaceae, rice, Medicago and Mimulus. Nearly all of these genomes have been sequenced or are currently being sequenced and have substantial communities of biologists and biotechnologists working on them. Priorities are set according to importance of the crop or model species, size of community, and stage of development of the field.

Other volumes take a disciplinary or technological focus, including the currently planned volume on plant cytogenetics and potential volumes on topics such as comparative genomics, epigenetics, and functional genomics.

More information about this series at http://www.springer.com/series/7397

Christopher A. Cullis

Editor

Genetics and Genomics of Linum

Editor
Christopher A. Cullis
Department of Biology
Case Western Reserve University
Cleveland, OH, USA

Preface

Flax (*Linum usitatissimum* L.) is one of the founding agricultural crops that was domesticated from the wild progenitor, commonly known as pale flax, by hunter-gatherers some 30,000 years ago. It has been selected either for the production of oil or for fiber. Fiber flax has been bred for its long, mainly unbranched stem containing long fibers. Linseed (or oil-seed flax) has been selected for short and highly branched plants to increase the number of flowers to maximize seed production. The health-related properties of flax in human and animal nutrition have opened an additional avenue for flax improvement. Modern flax breeding has the overall objectives to develop cultivars with increased fiber or seed yields, or improved health-related properties. Among the important characteristics are an improved adaptation and disease resistance to augment yields of the commercially important commodities. More than 40,000 accessions representing 54 *Linum* species are conserved in the gene banks around the world. These accessions provide an essential genetic resource for flax breeding and research, and this genetic resource has been partially characterized from cytogenetic, genetic, and genomic aspects. The new molecular genetic tools will make the flax germplasm more accessible to flax. The Canadian breeding program is presented as a model for how variety improvement has proceeded and the choice of which important characteristics need to be considered.

The initial molecular characterization of the flax germplasm was through molecular markers including isozymes, random amplified polymorphic DNAs (RAPDs), and restriction fragment length polymorphisms. The TUFGEN (Total Utilization of Flax Genomics) project has been a turning point for flax genomics resources with the development of the whole-genome sequence of CDC Bethune and the underpinning resources such as the datasets for BAC-end sequences and large-scale ESTs. The initial genome assembly has been refined by the inclusion of a range of techniques, including flax BNG optical map, the BAC physical map, and consensus genetic maps, resulting in a high-quality integrated genome. The assembly has resulted in chromosomes ranging from 15.6 to 29.4 Mb in size. However, the 15 pairs of chromosomes appear to be indistinguishable cytologically in size in spite of a twofold difference in DNA content.

The assembled genome sequence has been augmented by other molecular resources including the characterization of both coding and small RNA molecules. The transcriptomic analysis of the flax seed development will further the improvement of oil production and health-modulating phytochemicals of flax, which, along with the bast fiber, are arguably the most economically important traits in flax. The assembled genome sequence has also facilitated the theoretical identification of small RNA molecules as well as their physical identification.

Although there is more variation in the wild flax progenitor germplasm than in the cultivated varieties, chemical mutagenesis-derived mutant populations have been successfully used to identify important genes involved in cell wall formation among others. Flax is also amenable to transformation, and the development of the floral dip procedure for this species may facilitate other approaches to mutant production and increasing available variation. However, the issues with the transgenic flax, Triffid, and its unregulated escape into the commercial seed provide a cautionary tale for the development of genetically modified flax.

Some flax varieties have a particular characteristic, which is not known in other plant species, namely, the rapid modulation of its genome under certain growth conditions. The variation, with the generation of altered stable lines termed genotrophs, has been shown to occur within a specific defined subset of the genome. Understanding the mechanisms and characteristics of this genome compartment that appears to have the function of modifying the phenotype without the deleterious effects associated with a random mutagenesis makes it an interesting evolutionary mechanism. The plethora of molecular tools available allows this phenomenon to be understood at the molecular level, and potentially capable of manipulation.

The molecular resources available, the commercially important, and the health-related characteristics make flax an accessible interesting model system for understanding complex pathways while also providing the basis for improving the commercial attractiveness of the crop.

The development of the molecular resources for flax is directly attributable to the investment in the TUFGEN program and to the collaborations that were fostered, and continue to thrive, through the program.

Cleveland, OH, USA   Christopher A. Cullis

Contents

Chapter 1
A Taxonomic View on Genetic Resources in the Genus *Linum* L. for Flax Breeding

Axel Diederichsen

1.1 Introduction

The most comprehensive taxonomic review of the genus *Linum* mentioned that there are about 200 botanical species (Winkler 1931). This conspectus is, however, outdated. The lack of a recent review has caused the treatments used for *Linum* taxa in local, national or regional floras to apply scientific names and classifications that are not harmonized. Inconsistencies among floras in assigning a taxonomic rank to a given taxon occur. Many synonyms exist, and they are often used without cross-referencing. A conspectus is needed to compare *Linum* species described from different areas with different names that in fact may be closely related or even identical. A coherent taxonomic review is required to understand and communicate information on the species and infraspecific diversity in the genus *Linum*. Such taxonomic inconsistencies are quite common in many genera and have great implications on managing crop genepools in genebanks. For example, similar scenarios exists in sunflower (*Helianthus annuus* L.; Atalgić and Teryić 2016) and chickpea (*Cicer arietinum* L.; Diederichsen et al. 2009).

Since 1998, the Canadian national genebank for plant genetic resources for food and agriculture, Plant Gene Resources of Canada, has put major efforts into the conservation and research of flax diversity. Classical taxonomic approaches were

A. Diederichsen (✉)
Plant Gene Resources of Canada, Saskatoon Research and Development Centre, Agriculture and Agri-Food Canada, Saskatoon, SK, Canada
e-mail: axel.diederichsen@agr.gc.ca

part of these efforts to group and assess the diversity of the Canadian genebank holdings of 3500 accessions of cultivated flax and more than 100 accessions of 18 other *Linum* species (Diederichsen and Fu 2008; PGRC 2016). This book chapter uses flax as a case study to describe the role of taxonomy and systematics in the conservation and utilization of plant genetic resources. The objective is to review contributions made by taxonomists and crop plant researchers to find a broadly accepted way to describe and communicate the diversity in the genus *Linum* with emphasis on the conservation and utilization of this diversity in the breeding of cultivated flax.

Rational utilization and conservation of genetic diversity, as well as assessing this diversity, requires a tool to communicate unambiguously about such diversity. The taxonomical unit of the species is generally used without much thought. The loss of a species is of great concern and makes the news in public media. Species extinction mobilizes policymakers to take action. However, when we speak of genetic resources, we deal mostly with the diversity within a species, which is the genetic diversity. Loss of such infraspecific diversity is much less spectacular, but to lose such diversity in crop plants and their wild relatives may be of critical importance due to the implications on food security.

It is very much disputed whether systematics and taxonomy, the categorizing and naming of groups of similar elements using scientific methods, are appropriate tools for distinguishing genetic diversity within a species, i.e. on the infraspecific level. Moreover, the relevance of taxonomy for utilization of plant genetic resources is sometimes doubted, as are the concepts of taxonomy in general. Some taxonomists have questioned the Linnaean species concept as such (Bachmann 1998). The main argument is that taxonomy is not considered an objective science, but a subjective view on diversity. Following the path initiated by Linnaeus and well-described in its unfolding over time by Stearn (1986), taxonomy has in particular shied away from applying established taxonomic principles to genepools of crop plants. Consequently, the rules of the International Code for Botanical Nomenclature of Cultivated Plants (Brickell et al. 2009) are a useful tool for naming modern cultivars, but are not useful for orientation in the wider genepool of a cultivated species, which is urgently required for genebank management (Diederichsen 2004).

Classical taxonomy was strictly based on morphological features distinguishable with the naked eye. Later, micro-morphological features, chemical features and, more recently, molecular technologies culminating in sequencing the nuclear DNA have become available and are today dominating the debate on biodiversity. Interestingly, research in flax that started in the 1950s has shown that the amount of nuclear DNA within the species varies and can vary from generation to generation (Durrant 1962; Cullis 2005). We may have reached a point in time where we need to acknowledge that the molecular descriptions only represent another approach that contributes to the understanding of a species but do not allow us to capture fully a species as it unfolds in real life. As previously articulated (see

review by Small 1989), the author also feels that each approach to a species will contribute to the understanding of that species. However, we must acknowledge that any definition of a botanical species or other taxon is only an approximation of our thinking to the whole and complex truth and is based on the tools for perception we use for making our observations. Without tools we return to what classical taxonomists relied on: the bare eye, the senses given to us as humans to perceive the world. The highly sophisticated tools used in phenomics applied today are nothing but a major refinement of what our eyes can observe. In addition, molecular methods, including all kind of markers and sequencing of plastome and nuclear DNA, are nothing but observations using other tools and are essentially of a descriptive nature. Slight modifications happen constantly in living organisms. The recently emerging discipline of epigenetics points at genomic plasticity, a phenomenon observed in flax 50 years ago (Durrant 1962; Cullis 2005). It is remarkable that flax has directed research towards epigenetics but at the same time has been the species that guided research to the gene-for-gene interaction between a plant and a pathogen based on very static genetic principles (Flor 1955).

It may be impossible to capture the idea of a biological entity with terms that suggest a static and unchangeable state because a living being will always change while alive and even more so over generations. Mansfeld (1962) clearly described this dilemma or pseudo-problem (he used the term *Scheinproblem*) of systematics. We need to find a convention for communication that comes closest to this ever-changing biological entity and, at the same time, enhances our ability to work with it in plant breeding, in agriculture, in trade, in conservation efforts or in research.

Classical concepts may be currently falling out of favour, but to overview trends over time, it is important to understand the earlier publications that were built on the classical concepts for assessing the diversity of crop genepools. This method was essential to the Vavilov school in crop plant research, which shaped genetic resources exploration in the former Soviet Union and influenced in particular the eastern European countries and Japan (Vavilov 1931; Loskutov 1999). The phenotype of a plant is immediately accessible to our senses, and many phenotypic characters are very relevant for the utilization of that plant. The functionality of many morphological features allows interpretations in the larger ecological or evolutionary context of a species. For efficient communication, the species name is relevant. It is for good reason that common names for plant groups, which are congruent with what later was recognized by cartesian science as a botanical species, have been developed in all cultures. In addition, many infraspecific groups for particular usages based on distinct differences in phenotype have received common names for distinguishing such types within a cultivated plant species (Diederichsen 2004). For farmers and trade, and for interaction among humans, such naming and language have been instrumental for thousands of years. Scientific conventions, such as the taxonomy used in the genebank database

system GRIN-Global, maintained by the United States Department of Agriculture to communicate the content of the National Plant Germplasm System, show how relevant this still is (Wiersema and Leon 2016). The International Seed Testing Association (ISTA) also relies on consistent taxonomy for supporting trade and standardization of testing methods (Wiersema 2013). Small (1993) emphasized the great economic impact plant systematics have on agriculture. Other crop plant researchers such as Hanelt (1988) stressed the relevance of taxonomy for managing genebanks. Unfortunately, most modern genebanks lack a close association with taxonomists. As a result, material is sometimes stored and passed on to genebank clients without stringent protocols for botanical identification. This is of great concern and negatively affects our ability to conserve and utilize this diversity. Concerns regarding the lack of taxonomic expertise for effective conservation of biodiversity have been expressed elsewhere (Small et al. 1995), but how this impacts the operations of genebanks preserving plant genetic resources for food and agriculture is rarely articulated.

1.2   The Primary Genepool of Genetic Resources for Flax Breeding

The only agricultural crop in the genus *Linum* is the species *Linum usitatissimum* L., and it has two distinct usages: linseed is used for seed oil and fibre flax is used for extracting the long stem fibre for textile production. In many languages, distinct names exist for flax plant types used for seed oil extraction and for types used for fibre (Diederichsen and Richards 2003). Obviously, common sense guided people in various language families in creating these distinct names to facilitate efficient communication of genetic diversity below the species rank. Taxonomists, concerned about the correct scientific naming of plants, have become reluctant over the years to be as precise as common language in assigning scientific names to these distinct usage groups in flax. Diederichsen and Richards (2003) supported the grouping of cultivated flax into four major groups and using the scientific names for them proposed by Kulpa and Danert (1962). A wide range of diversity in agronomically important characters exists in cultivated flax (Diederichsen et al. 2013). Research has relied on identifying infraspecific groups in flax to interpret archaeological findings and make conclusions on its evolution under domestication (Herbig and Maier 2011). Even flax workers reluctant to apply a detailed infraspecific classification may see the practical evidence in the grouping of cultivated flax into these four principal groups (convarieties).

1.3   Key for Determination of the Four Convarieties of Cultivated Flax, *L. usitatissimum*

A	Capsules open completely septicidaly and loculicidaly during ripening (Fig. 1.1). Seeds are shattered. Later, empty capsules are dropped.	**1.** convar. *crepitans*
A∗	Capsules not opening during ripening or only slightly septicidaly separating from each other (Fig. 1.1). Seeds are not easily shattered. Empty capsules are not dropped.	**B**
B	Plant height more than 70 cm and only the upper 1/3 or less of the entire stem length with side branches; if less than 70 cm, then stem branches only in the upper 1/5 of the entire stem length.	**2.** convar. *elongatum*
B∗	Plant height usually less than 70 cm; more than 1/5 of the entire stem length with side branches.	**C**
C	Weight of 1000 seeds more than 9 g; plants usually without basal branches.	**3**. convar. *mediterraneum*
C∗	Weight of 1000 seeds less than 9 g; plants often with basal branches.	**4.** convar. *usitatissimum*

1. **L. usitatissimum** convar. **crepitans** (Boeninningh.) Kulpa et Danert, Kulturpflanze, Beih. 3, (1962) 374.

The convar. *crepitans* (Fig. 1.2) refers to dehiscent flax. It has been used in Central and Southeast Europe as a fibre plant (Hegi 1925). Seed shattering makes it difficult to harvest the seeds. This type of flax is no longer cultivated, and only the germplasm collections conducted early in the twentieth century facilitated conservation of this type in genebanks. Its range of variation is limited. With the exception of the dehiscence of the capsules, the plants are phenotypically similar to those of the convar. *usitatissimum*.

Fig. 1.1 Different degrees of dehiscence in mature flax capsules. The complete dehiscence characterizes the primitive type of cultivated flax, L. *usitatissimum* convar. *crepitans*. (Photo: R. Underwood, Agriculture and Agri-Food Canada)

Fig. 1.2 Flowering/maturing
plant of *L. usitatissimum*
convar. *crepitans*. This type
of flax has only been
conserved in genebanks.
PGRC accession CN 100852,
landrace from Portugal.
(Photograph: Z. Bainas,
Agriculture and Agri-Food
Canada)

2. ***L. usitatissimum*** convar. ***elongatum*** Vav. et Ell. in Kul't. Fl. SSSR 5, 1 (1940)
   153, pro prole sub *L. indehiscens* Vav. et Ell. subsp. *eurasiaticum* Vav. ex Ell.

The convar. *elongatum* (Fig. 1.3) refers to typical fibre flax. It has long stems,
which are only branched at the top. This type of flax has been of great importance
in the temperate and northern areas of Europe and in particular Eastern Europe
(Vavilov 1926). China is also a centre of diversity for fibre flax. Fibre flax has a
shorter vegetative period than the large-seeded flax. This group is identical with the
fibre flax as defined by Dillman (1953).

3. ***L. usitatissimum*** convar. ***mediterraneum*** (Vav. ex Ell.) Kulpa et Danert,
   Kulturpflanze, Beih. 3, (1962) 376.

The convar. *mediterraneum* (Fig. 1.4) refers to large-seeded flax with large flowers
and capsules and branched stems. It is only used for seed production. Flax of this
type originates from the Mediterranean area and has a long vegetative period. This
group is identical with the Mediterranean seed flax defined by Dillman (1953).

4. ***L. usitatissimum*** convar. ***usitatissimum***.

The convar. *usitatissimum* (Fig. 1.5) refers to the intermediate flax, or dual purpose
flax. This is the most common type of flax in the world. Within this convariety, further
segregation into several different morphotypes is possible. This group covers the spring-
type seed flax, winter-type seed flax and Indian and Ethiopian (Abyssinian) flax, as

Fig. 1.3 Flowering/maturing plant of *L. usitatissimum* convar. *elongatum*, the typical plant type for using the long stem fibre that was common in Europe. PGRC accession CN 18991, cultivar Nike from Poland. (Photograph: Z. Bainas, Agriculture and Agri-Food Canada)

Fig. 1.4 Flowering/maturing plant of *L. usitatissimum* convar. *mediterraneum.* This type is only used for seed production and most flax from India and the Mediterranean area belongs to this group. PGRC accession CN 98566, landrace Neelum (3/2) from India. (Photograph: Z. Bainas, Agriculture and Agri-Food Canada)

Fig. 1.5 Flowering/maturing
plant of *L. usitatissimum*
convar. *usitatissimum.* This
intermediate flax covers the
majority of flax types and is
common in all regions.
PGRC accession CN 19017,
cultivar CDC Normandy
from Canada. (Photograph:
Z. Bainas, Agriculture and
Agri-Food Canada)

defined by Dillman (1953). All Canadian oilseed cultivars belong to this group because
their weight of 1000 seeds does not reach the size of the large-seeded flax types.

On the next lower taxonomic level, Kulpa and Danert (1962) proposed the dis-
tinction of 28 botanical varieties within these 4 convarieties of cultivated flax. This
classification system was used to compare the genebank holdings of the national
genebanks of Germany (1606 accessions) and Canada (2748 accessions)
(Diederichsen 2009). The results showed a similar concentration towards certain
phenotypes in both genebanks. This grouping was also instrumental when assem-
bling a core collection of the Canadian flax collection (Diederichsen et al. 2012).
However, the utilization of formal taxonomic names for describing genetic diversity
within a species is not widely accepted, and it will probably remain a tool only used
by genebank curators specializing in particular crops.

Other formal infraspecific groupings of cultivated flax into several botanical vari-
eties exist. Alefeld (1866) described 11 botanical varieties. Howard (1924) suggested
a grouping of Indian flax into 26 botanical varieties based on flower and seed colour.
Elladi (1940) expanded the classification to 119 botanical varieties. The most recent
formal classification of cultivated flax was proposed by Černomorskaja and Stankevič
(1987) and distinguished five subspecies (Diederichsen and Richards 2003). None of
these classifications has been applied to a genebank collection. In this context, it is
also important to note Dillman's (1953) very comprehensive description and cate-
gorization of diversity of cultivated flax. This is not a formal taxonomic grouping,
but a grouping into cultivar groups. In contrast to the formal taxonomic groupings,

it considered agronomically relevant traits such as the need for vernalization as a distinguishing feature. Dillman's system was applied to the world flax collection of the United States Department of Agriculture (USDA) by grouping the genepool into six plant types: (1) fibre flax; (2) spring-type seed flax; (3) winter-type seed flax; (4) short, large-seeded, Indian seed flax; (5) Ethiopian forage-type flax; and (6) Mediterranean or Argentine seed flax.

1.4 The Wild Progenitor of Cultivated Flax Belonging to the Primary Genepool

Linum bienne Mill., Gard. dict. ed. 8 (1768) n. 8 – *L. usitatissimum subsp. angustifolium* (Huds.) Thell., Fl. adv. Montp. (1912) 361; other synonyms are *L. ambiguum* Jord., *L. hohenhackeri* Boiss., *L. usitatissimum* subsp. *hispanicum* Thell., *L. dehiscens* Vav. et Ell. subsp. *angustifolium* (Huds.) Vav. et Ell. and *L. angustifolium* Huds. (Hammer 2001). The English common name of the wild progenitor is pale flax.

Pale flax (Fig. 1.6) has a biennial or perennial growth habit, i.e. it needs a vernalization to induce flowering. Heer (1872) was the first to identify pale flax as the wild

Fig. 1.6 The middle row are plants of *L. bienne* Mill., the wild progenitor of cultivated flax. The many basal branches, bushy growth habit and the dehiscence and shattering of capsules are typical for this species. (Photograph: A. Diederichsen, Agriculture and Agri-Food Canada)

progenitor of cultivated flax. The flowers are homostylous and self-pollinated. The capsules open spontaneously and the seeds shatter. The species occurs in the Mediterranean area and in Western Europe. Tammes (1928) demonstrated that this species is interfertile with cultivated flax. However, no reports exist about its usage in flax breeding.

A significant contribution towards the conservation and understanding of *L. bienne* was made by scientists collecting germplasm of this species in Turkey and depositing it in a genebank (Uysal et al. 2012). This germplasm originated from areas of flax domestication that were until recently not represented in world gene-banks, and it is in particular useful for domestication research in flax. Investigation of the morphology of pale flax showed that some characters have a wider range of variation in this taxon than in cultivated flax (Diederichsen and Hammer 1995). Recent molecular studies have confirmed the relationship with the cultivated species (Fu and Allaby 2010; Soto-Cerda et al. 2014). Additional germplasm of this species from the Balkans has recently been deposited in genebanks (Gutaker 2014).

From the biological point of view, pale flax should be placed in the same botanical species as cultivated flax since the two types can freely intercross. Harlan and de Wet (1971) proposed the genepool concept as a basis for rational classification of crop genepools and strongly suggested that formal taxonomy reflected such biological and evolutionary relationships. Based on this principle, the correct name for pale flax is *L. usistatissimum* subsp. *angustifolium* (Huds.) Thell., and all cultivated flax would fall in the other subspecies, L. *usitatissimum* subsp. *usitatissimum*. Hammer (2001) followed this principle. However, lengthy names are often cumbersome to use. Even worse, if someone is lax and omits the subspecies name, the cultivated form and wild progenitor can be easily confused. For that reason, and for convenience, it seems the species rank is mostly still applied to distinguish pale flax, and the correct name for pale flax at the botanical species rank is *L. bienne* Mill. All other *Linum* species will not produce fertile offspring when crossed with cultivated flax and, therefore, belong to the tertiary genepool from the perspective of plant breeding. Various reports about crossing cultivated flax with taxa other than its wild progenitor are very questionable, as the taxonomic identification of the material was not presented convincingly (Diederichsen 2007). This exemplifies how important it is to verify the botanical identity of material based on solid taxonomy. When scientists specializing in physiology or genomics use misidentified material, errors in communication occur.

1.5   The Secondary and Tertiary Genepools of Species in the Genus *Linum*

Studies of the relationships among *Linum* taxa have not changed the views that were based on the earlier morphological and cytological results (Fu and Allaby 2010). The review of Winkler (1931) is still the baseline. In recent years, descriptions of new wild species in the genus *Linum* have been provided for Southern Italy (Peruzzi 2011), Turkey (Tugay et al. 2010; Yılmaz 2010; Yılmaz and Kaynak 2008), Greece (Iatrou, 1989; Christodoulakis 1999) and Mexico (Rogers 1982). The species composition in the genus *Linum* has not been subject to many studies.

Diederichsen (2007) reported that 33 world genebanks preserved about 600 accessions of 52 *Linum* species belonging to the secondary and tertiary gene-pools for breeding of cultivated flax. A recent inspection of the database at the World Information and Early Warning System (WIEWS) on Plant Genetic Resources for Food and Agriculture of the Food and Agricultural Organization of the United Nations (FAO 2016) showed that 82 institutions around the world maintain a total of 685 *Linum* accessions that are not listed as *L. usitatissimum*, *L. bienne*, *L. angustifolium* or *L. usitatissimum* subsp. *angustifolium*. Assuming all these botanical identifications are correct, these 685 accessions would belong to *Linum* taxa that are neither cultivated flax nor its wild progenitor, pale flax. At the same time, the WIEWS database lists 3479 accessions as *Linum* sp., i.e. the botanical identification has not been conducted or at least not been reported to the database. Thus, the situation of *Linum* germplasm conservation has not changed much since Diederichsen's (2007) review. It is obvious, however, that the botanical identification of material even to the species level is a bottleneck in many genebank collections.

The described situation highlights the wide margin of error we must accept when considering the global status of both in situ and ex situ conservation of plant genetic resources in the genus *Linum*, especially when decisions are made that directly affect the conservation of these genetic resources. Without a consistent taxonomy and lacking reliable species identification, we can only make very approximate estimates of the diversity and conservation status of *Linum* germ-plasm. Databases that are meant to support decision making for conservation and utilization of genetic resources suffer tremendously due to taxonomic insecurity. For example, the Global Biodiversity Information Facility lists only 26 accepted species names in the genus *Linum* (GBIF 2016). This list does not include cultivated flax or the wild progenitor species of the primary genepool. However, older, well-reputed floras for Turkey (Davis 1967) and Europe (Ockendon and Walters 1968) list more species than that. The database "Plant List" (Royal Botanic Gardens Kew and Missouri Botanical Garden 2016) lists scientific names of species and infraspecific taxa and includes 169 accepted names and 220 names that are unresolved in the genus *Linum*. The species delimitations remain vague for many species, in particular in the *L. flavum* and *L. perenne* groups (Ockendon and Walters 1968). Ornamental plants have been selected in these two groups. These examples show that a great deal of taxonomic confusion continues to exist in the genus *Linum* and that even on the species level we have not improved our ability to communicate *Linum* diversity using taxonomy since Winkler (1931) despite enormous advances in technology. For field botanists, it is essential to have botanical keys that allow identification of plants at least to the species level in the field. In Fig. 1.7, a situation encountered by the author on the Crimean Peninsula south of Sevastopol with three *Linum* species growing on the same meadow in close proximity to each other illustrates this need (Diederichsen et al. 2012). This illustrates the usefulness of species delimitations based on morphological characters that can be recognized by the field botanist, germplasm collector, genebank curator or plant breeder.

Fig. 1.7 Three *Linum* species encountered at the same location in close proximity on the Crimean peninsula south of Sevastopol in 2009. Left to right: *Linum tenuifolium* L., *L. corymbulosum* Reichenb. and *L. austriacum* L. (Photograph: A. Diederichsen, see also Diederichsen et al. 2012).

1.6 Conclusions

A taxonomic revision of the genus *Linum* is urgently required to allow progress in understanding the diversity in this genus and provide an essential tool for conservation and utilization of this diversity. In flax, classical taxonomic concepts can be used to communicate distinct intraspecific diversity. The classical approaches of systematics and taxonomy remain useful tools for those encountering diversity during collecting missions or when dealing with phenotypic diversity during regeneration of diverse ex situ collections of cultivated plants. Results from additional observations using modern molecular methods can be interpreted based on the classical phenotypic groupings and add considerably to our understanding of crops and crop wild relatives.

References

Alefeld F (1866) Landwirthschaftliche Flora. [Agricultural flora]. Wiegandt and Hempel, Berlin

Atalgić J, Teryić S (2016) The challenges of maintaining a collection of wild sunflower (*Helianthus*) species. Genet Resour Crop Evol 63:1219–1236

Bachmann K (1998) Species as units of diversity: an outdated concept. Theor Biosci 117:213–230

Brickell CD, Alexander C, David JC, Hetterscheid WLA, Leslie AC, Malecot V, Jin X, Cubey JJ (2009) International code of botanical nomenclature for cultivated plants (ICNP or cultivated plant code) 8th ed. Scripta Horticulturae, International Society of Horticultural Science 10:1–184

Černomorskaja NM, Stankevič AK (1987) K voprosu o vnutryvidodvoj klassifikacii l'na obyknovennogo (*Linum usitatissimum* L.). [On the problem of intraspecific classification of common flax (*Linum usitatissimum* L)]. Selekcija I genetika techničeskych kul'tur 113:53–63

Christodoulakis (1999) *Linum phitosianum* spec. nova (linaceae) aus Griechenland. Phyton 33:289–294

Cullis CA (2005) Mechanisms and control of rapid genomic changes in flax. Ann Bot 95:201–206

Davis PH (1967) *Linum* L. In: Davis PH, Cullen J, Coode MJE (eds) Flora of Turkey, vol 2. Edinburgh University Press, Edinburgh, pp 425–450

Diederichsen A (2004) Case studies for the use of infraspecific classifications in managing germplasm collections of cultivated plants. Acta Hortic (634):127–139

Diederichsen A (2007) *Ex-situ* collections of cultivated flax (*Linum usitatissimum* L.) and other species of the genus *Linum* L. Genet Resour Crop Evol 54:661–678

Diederichsen A (2009) Taxonomical and morphological assessments of infraspecific diversity in cultivated flax (*Linum usitatissimum* L.). J Agr Rural Dev Trop Subtrop Supplement 92:33–51

Diederichsen A, Fu YB (2008) Flax genetic diversity as the raw material for future success. Proceedings, the 2008 international conference on flax and other Bast plants, Saskatoon, Canada, July 21–23, 2008, pp 270–280

Diederichsen A, Hammer K (1995) Variation of cultivated flax (*Linum usitatissimum* L. subsp. *usitatissimum*) and its wild progenitor pale flax (subsp. *angustifolium* [Huds.] Thell.). Genet Resour Crop Evol 42:263–272

Diederichsen A, Richards KW (2003) Cultivated flax and the genus *Linum* L. – taxonomy and germplasm conservation. In: Muir A, Westcott N (eds) Flax, the genus *Linum*. Taylor & Francis, London, UK, pp 22–54

Diederichsen A, Seferova IV, van der Maesen LJG, Lülsdorf M (2009) Challenges in preserving, communicating and utilizing the diversity of wild perennial species of the genus *Cicer* L. proceedings international conference ПРОБЛЕМЫ ЭВОЛЮЦИИ И СИСТЕМАТИКИ КУЛЬТУРНЫХ РАСТЕНИЙ St. Peterburg, Russia, Dec 9–11, 2009, pp 273–278

Diederichsen A, Rozhkov RV, Korzhenevsky VV, Boguslavsky RL (2012) Collecting genetic resources of crop wild relatives in Crimea, Ukraine, in 2009. Crop Wild Relative 8:34–38

Diederichsen A, Kusters PM, Kessler D, Bainas Z, Gugel RK (2013) Assembling a core collection from the flax world collection maintained by Plant Gene Resources of Canada. Plant Genetic Res Crop Evol 60:1479–1485

Dillman AC (1953) Classification of flax varieties 1946. United States Department of Agriculture Technical Bulletin no. 1054. United States Department of Agriculture, Washington, 56 p

Durrant A (1962) The environmental induction of heritable change in *Linum*. Heredity 17:27–61

Elladi VN (1940) *Linum usitatissimum* (L.) Vav. Consp. Nov. – Len. (Russ.). In: Vul'f EV, Vavilov NI (eds) *Kul'turnaja flora SSSR, fibre plants*, Sel'chozgiz, Moskva, Leningrad, vol. 5, Part 1, pp 109–207

FAO (2016) World Information and Early Warning System (WIEWS) on Plant Genetic Resources for Food and Agriculture (PGRFA). http://www.fao.org/wiews-archive/wiews.jsp

Flor HH (1955) Host-parasite interaction in flax rust – its genetics and other implications. Phytopathology 45:680–685

Fu YB, Allaby RG (2010) Phylogenetic network of *Linum* species as revealed by non-coding chloroplast DNA sequences. Genet Resour Crop Evol 57:667–677

GBIF (2016) Global biodiversity information facility. http://www.gbif.org/

Gutaker R (2014) The genetic variation of cultivated flax (*Linum usitatissimum*) and the role of its wild ancestor (*Linum bienne* Mill.) in its evolution. Ph.D. Thesis, University of Warwick, 186 p

Hammer K (2001) Linaceae. In: Hanelt P. and Institute of Plant Genetics and Crop Plant Research (eds.). In: Mansfeld's encyclopedia of agricultural and horticultural crops (except ornamentals), vol 2. Springer, Berlin, Heidelberg, pp 1106–1108

Hanelt P (1988) Taxonomy as a tool for studying plant genetic resources. Kulturpflanze 36:169–187

Harlan JR, de Wet JMJ (1971) Toward a rational classification of cultivated plants. Taxon 20:509–517

Heer O (1872) Über den Flachs und die Flachskultur im Altertum. Neujahrsblatt der Naturforschenden Gesellschaft Zurich 74:1–26

Hegi G (1925) Ilustrierte Flora von Mitteleuropa. Lehmanns Verlag, München, vol. 5, Part 1, pp 3–38

Herbig C, Maier U (2011) Flax for oil or fibre? Morphometric analysis of flax seeds and new aspects of flax cultivation in late Neolithic wetland settlements in Southwest Germany. Veg Hist Archaeobotany 20:527–533

Howard GLC (1924) Studies in Indian oil seeds no. 2 linseed. India Dept. Agriculture, Memoirs. Botanical Series 12:135–183

Iatrou GA. (1989) *Linum hellenicum* (Linaceae), a new species from Peloponnesos, Greece. Wildenovia 19: 69–73

Kulpa W, Danert S (1962) Zur Systematik von *Linum usitatissimum* L. [On the systematics of *Linum usitatissimum* L.]. (in German). Kulturpflanze, Beiheft 3:341–388

Loskutov I (1999) Vavilov and his institute, a history of the world collection of plant genetic resources in Russia. International Plant Genetic Resources Institute, Rome, p 188p

Mansfeld R (1962) Über "alte" und "neue" Systematik der Pflanzen. Kulturpflanze Beiheft 3:26–46

Ockendon DJ, Walters SM (1968) *Linum* L. In: Tutin TG, Heywood VH (eds) Flora Europaea, vol 2. Cambridge University Press, Cambridge, pp 206–211

Peruzzi L (2011) A new species of *Linum perenne* group (Linaceae) from Calabria (S Italy). Plant Biosys 145:938–944

PGRC (2016) Plant Gene Resources of Canada – Ressources phytogénétiques du Canada. http://pgrc.agr.gc.ca/

Rogers CM (1982) *Linum mgvaughii*, a new species from Mexico. Contr. Univ. Mich. Herb. 15:205–207

Royal Botanic Gardens Kew and Missouri Botanical Garden (2016). The Plant List, a working list of all plant species. http://www.theplantlist.org/

Small E (1989) Systematics of biological systematics (or, taxonomy of taxonomy). Taxon 38:335–356

Small E (1993) The economic value of plant systematics in Canadian agriculture. Can J Bot 71:1537–1551

Small E, Cayouette J, Catling PM, Brookes B (1995) An opinion survey of priorities for plant systematic and phytogeography in Canada. Bull Canadian Botanical Association 28:19–22

Soto-Cerda B, Diederichsen A, Duguid S, Booker H, Rowland G, Cloutier S (2014) The potential of pale flax as a source of useful genetic variation for cultivated flax revealed through molecular diversity and association analyses. Mol Breed 34:2091–2107

Stearn WT (1986) Historical survey of the naming of cultivated plants. Acta Hortic (182):19–28

Tammes T (1928) The genetics of the genus *Linum*. Bibliographica Genetica 4:1–36

Tugay O, Bağci Y, Uysal T (2010) *Linum ertugrulii* (Linaceae), a new species from Central Anatolia, Turkey. Ann Bot Fenn 47(2):135–138

Uysal H, Kurt O, Fu Y-B, Diederichsen A, Kusters P (2012) Variation in morphological and phenotypic characters of pale flax (*Linum bienne* Mill) from Turkey. Genetic Res Crop Evol 59:19–30

Vavilov NI (1926) Studies on the origin of cultivated plants. Bull Appl Botany 16(2):3–248

Vavilov NI (1931) Linneevskij vid kak sistema. [The Linnaean species as concept]. (in Russian). Tr po prikl bot, gen i sel 26(3):109–134

Wiersema JH (2013) ISTA list of stabilized plant names, ed 6. http://www.ars-grin.gov/~sbmljw/istaintrod.html

Wiersema JH, Leon B (2016) The GRIN-taxonomy crop wild relative inventory. In Maxted N, Mulloo ME, Ford-Lloyd BV (eds) Enhancing crop genepool use: capturing wild relative and landrace diversity for crop improvement. Council of Applied Biology International, Oxford, UK, pp 453–457

Winkler H (1931) Linaceae, Trib. r. 3. Linoideae-Eulineae. In Engler A. (ed) Die natürlichen Pflanzenfamilien nebst ihren Gattungen und wichtigeren Arten, insbesondere den Nutzpflanzen, 2nd edn. W. Engelmann, Leipzig, vol. 19a, pp 111–120.

Yılmaz Ö, Kaynak G (2008) A new species of *Linum* (Linaceae) from west Anatolia, Turkey. Bot J Linn Soc 156:459–462

Yilmaz Ö (2010) *Linum kaynakiae* sp. nov. (sect. Syllinum, Linaceae) from Turkey. Nordic Journal of Botany 28: 605–612

Chapter 2
A Molecular View of Flax Gene Pool

Yong-Bi Fu

2.1 Introduction

The concept of the crop gene pool was developed by Harlan and de Wet (1971) to assist the access of crop genetic resources for plant breeding and to assess the extent of gene flow among populations of a crop and related taxa. They classified germ-plasm resources into primary, secondary, and tertiary gene pools according to their genetic accessibility to the target cultigen. The primary gene pool is comprised of related species that can be directly mated with the crop to produce fertile progeny. The secondary gene pool consists of those closely related species which can be hybridized with the crop to give a partially fertile hybrid. The tertiary gene pool includes those distantly related species which produce sterile or inviable hybrids with the crop only through exploitation with advanced techniques such as those of genetic engineering. Accordingly, the classification of these gene pools has been made for many crops such as barley, wheat, and common bean (Spillane and Gepts 2001). However, practical issues associated with the use of this gene pool classification are not lacking (Smartt 1984; Hammer 1998; Spillane and Gepts 2001; Gladis and Hammer 2002), and a modification of the original gene pool model was also made to include the quaternary gene pool for genetically engineered crops (Gepts 2000; Gepts and Papa 2003). Traditionally, plant breeders have emphasized and/or utilized closely related, well-adapted domesticated materials within the primary gene pool as sources of genetic diversity (Kannenberg and Falk 1995; Kelly et al. 1998), but technical advances such as plant transformation and genomics have made

Y.-B. Fu (✉)
Plant Gene Resources of Canada, Saskatoon Research and Development Centre, Agriculture and Agri-Food Canada, Saskatoon, SK, Canada
e-mail: yong-bi.fu@canada.ca

the utilization of the secondary and tertiary gene pools more feasible. Efforts have been increasingly made to expand existing crop gene pools for plant breeding through introgressing traits by pre-breeding (e.g., see Sharma et al. 2013) and collecting crop wild relatives as genetic resources (e.g., Maxted et al. 1997, 2012; Castañeda-Álvarez et al. 2016). Thus, it is important to classify the gene pools for a crop species as genetic resources to guide germplasm exploration and conservation (Ford-Lloyd et al. 2011) and to evaluate them for plant breeding (Henry 2014). However, little research is done to characterize the classified gene pools (Spillane and Gepts 2001; Brozynska et al. 2016).

Flax (*Linum usitatissimum* L.) is one of the founding agricultural crops in the Near East (Zohary and Hopf 2000) and was domesticated for production of either oil or for fiber by hunter-gatherers some 30,000 years ago (Kvavadze et al. 2009). Fiber flax is bred for its long stem containing long fibers and is mainly grown in Russia, China, Egypt, and near the Northwestern European coast, whereas linseed was cultivated for short and highly branched plants to increase the number of flowers for seed production in Canada, China, USA, India, and Russia (Vromans 2006). Modern flax breeding has faced many challenges with the overall objectives to develop flax cultivars with increased fiber or seed yields, better adaptation, and disease resistance for changing market needs (Soto-Cerda et al. 2014a). Climate change is also expected to impact flax breeding and production (Hall et al. 2016). The health-related properties of flax in human and animal nutrition will stimulate the search for new traits in flax germplasm collections and their incorporation into breeding schemes (Muir and Wescott 2003). Meeting these challenges requires an accelerated access to diverse flax gene pools and an effective search for useful genetic variability (Brozynska et al. 2016), as natural variation is the raw material for any crop improvement and constitutes a critical part of any long-term strategy to enhance the productivity, sustainability, and resilience of crop varieties and agricultural systems (Godfray et al. 2010; Henry and Nevo 2014).

Currently, more than 40,000 flax accessions representing 54 *Linum* species are conserved in 81 genebanks around the world (Diederichsen 2007; FAO 2010). These accessions provide an essential genetic resource for flax breeding and research (Diederichsen and Fu 2008). Efforts have been made around the world to evaluate and characterize this genetic resource (e.g., Fu 2005; Smýkal et al. 2011; Soto-Cerda et al. 2013). Molecular genetic tools have been developed and applied to enhance germplasm characterization, making flax germplasm more accessible to flax breeding (e.g., Soto-Cerda et al. 2013, 2014b). These applications also provide an opportunity to characterize the poorly understood flax gene pool. In this chapter, we attempt to classify the flax gene pool mainly using the concept of the crop gene pool developed by Harlan and de Wet (1971) and to identify its molecular features (or patterns of genetic variation derived from molecular markers or other related studies) through a literature review. It is our hope that this review will help to paint a picture of the extant flax gene pool with molecular features for plant breeding, germplasm conservation, and evolutionary research.

2.2 Flax Gene Pool

Many taxonomic, cytological, and evolutionary studies have been conducted for the genus *Linum* (e.g., see Winkler 1931; Ockendon and Walters 1968; Gill and Yermanos 1967a, b; Harris 1968; Chennaveeraiah and Joshi 1983; McDill et al. 2009; McDill and Simpson 2011; Fu et al. 2016), but no reports have been found specifically on the classification of flax genetic resources into various flax gene pools, following the concept of the crop gene pool developed by Harlan and de Wet (1971) and/or its modification by Gepts and Papa (2003) (Diederichsen 2007). Here, we have made the first attempt, based on the knowledge acquired from literature, to draw a picture of the flax gene pool for further research and flax improvement on the boundaries of the conceptualized gene pools.

2.2.1 Primary Gene Pool

The flax primary gene pool should consist of cultivated flax (*L. usitatissimum*) and its progenitor pale flax (*L. bienne*) (Diederichsen 2007). Both species are diploid, inbreeding species with 2n = 30 chromosomes and homostylous flowers. They can easily hybridize with each other in both directions to yield fertile hybrids (Gill 1966). Cytologically, cultivated flax is differentiated from pale flax by one chromosome translocation (Gill 1966). Previous studies confirmed that pale flax is the progenitor of cultivated flax (Tammes 1928; Diederichsen and Hammer 1995; Fu et al. 2002a; Allaby et al. 2005; Uysal et al. 2010). The domestication of pale flax occurred in the Near East several thousand years ago (Heer 1872; Zohary and Hopf 2000), although there is also evidence of flax use in Neolithic cultures as a source of fiber. Flax fiber has been identified in prehistoric sites in Israel, Syria, and Georgia (van Zeist and Bakker-Heeres 1975; Kvavadze et al. 2009). Flax was grown in Egypt between 4500 and 4000 BC and reached Switzerland around 3000 BC (Helbaek 1959).

Pale flax has several synonyms, including *L. usitatissimum* subsp. *angustifolium* (Huds.) Thell, *L. ambiguum* Jord., *L. hohenhackeri* Boiss., *L. usitatissimum* ssp. *hispanicum* Thell., *L. dehiscens* Vav. et Ell. subsp. *angustifolium* (Huds.) Vav. et Ell., and *L. angustifolium* Huds. (Hammer 2001). It is indigenous to the geographical territory bordering the Mediterranean Sea, Iran, and the Canary Islands (Diederichsen and Hammer 1995). Currently, there are 350–400 accessions of pale flax conserved in several major genebanks. However, the largest original collection of 120 accessions is conserved at Leibniz-Institut für Pflanzengenetik und Kulturpflanzenforschung (IPK), Gatersleben, Germany, including collections from botanical gardens and from natural habitats in Europe. Two recent collections were made: 36 accessions from Turkey (Uysal et al. 2010) and 16 accessions from the Balkan countries (Gutaker 2014). Overall, the existing germplasm collections have large geographical gaps while attempting to represent the species distribution, and

efforts are needed to collect germplasm from across the complete natural geographical range of species distribution.

There were more than 40,000 accessions of cultivated flax currently conserved in 81 genebanks around the world (FAO 2010), and possibly only 10,000–15,000 accessions are genetically unique (Diederichsen 2007). These accessions represent primitive flax, flax landraces, released cultivars, and breeding lines that were collected in different countries over time. Specifically, they can be divided into six major groups based on morphological and qualitative traits: (1) fiber flax; (2) oil flax; (3) dual-purpose flax that is an intermediate form between the first two types cultivated for fiber and oil; (4) large seeded flax with a set of specific morphological features and cultivated for oil in the Mediterranean region and North Africa; (5) winter flax cultivated for fiber and oil in the Caucasus, Turkey, Balkans, and some other south regions of Europe; and (6) dehiscent flax, the primitive flax form with dehiscent capsules (Dillman 1953; Kulpa and Danert 1962; Diederichsen 2007; Melnikova et al. 2014). Detailed analyses of ex situ flax germplasm distributions were made by Maggioni et al. (2002) and Diederichsen (2007). Four infraspecific groups, proposed by Kulpa and Danert (1962), have been characterized using phenotypic and molecular characteristics (Diederichsen and Fu 2006).

Large efforts have been made to characterize the germplasm of cultivated flax, as flax breeding has focused largely on either well-adapted materials or genetic diversity within cultivated flax (Vromans 2006; Diederichsen and Fu 2008; Hall et al. 2016). Characterization includes the application of various molecular markers, such as RAPD, AFLP, SSR, and SNP to assess genetic diversity, structure, and relationship among flax germplasm, and the genetic and genomic analyses of breeding materials. The recent sequencing of the flax genome (Wang et al. 2012) provides valuable tools for genomics-assisted breeding of this crop. The potential of pale flax as a source of useful genetic variation for cultivated flax was also explored through genomic analysis of 125 accessions (Soto-Cerda et al. 2014c).

2.2.2　Secondary Gene Pool

This gene pool consists of wild flax species which can be hybridized with cultivated flax to give a partially fertile hybrid. Several studies on the crossability of cultivated flax with wild flax species revealed successful crosses with the following *Linum* species: *L. nervosum*, *L. pallescens*, *L. africanum*, *L. corymbiferum*, *L. decumbens*, *L. hirsutum*, *L. floccosum*, and *L. tenue* (Gill and Yermanos 1967a; Bari and Godward 1970; Seetharam 1972). Specifically, Gill and Yermanos (1967a) found the following successful crosses:

L. usitatissimum x L. africanum
L. corymbiferum x L. usitatissimum
L. usitatissimum x L. decumbens
L. nervosum x L. usitatissimum
L. pallescens x L. usitatissimum,