School of Natural Sciences,

Trinity College Dublin

Ireland

Characterisation of *Miscanthus* genetic resources: a combined analysis of plastid and nuclear microsatellites, nrDNA sequences, flow cytometry and morphology

A thesis submitted for the degree of Doctor of Philosophy

August 2012

By

Mariateresa de Cesare

Supervisors: Dr. Trevor R. Hodkinson (TCD)

and Dr. Susanne Barth (Teagasc)

Declaration

I, the undersigned, hereby declare that I am the sole author of this dissertation and that the work presented in it, unless otherwise referenced, is my own.

I also declare that the work has not been submitted, in whole or in part, to any other university or college for a degree or other qualification.

I authorize the library of Trinity College Dublin to lend or copy this dissertation on request.

II

Dedicated to my mother and sister for their love and support.

And to Gabriele, for the joy he is going to bring in our lives!

ACKNOWLEDGEMENTS

I am very grateful to:

Summary

Miscanthus is a highly important forage and horticultural genus of perennial grasses (Poaceae) primarily native to South East Asia. *Miscanthus* is under intense global investigation as a biomass source for renewable energy production and several breeding initiatives are underway to develop new genotypes optimized for improved biomass and tolerance to a range of environmental stress conditions. A collection of 128 accessions belonging to the genus *Miscanthus* was established in Oak Park, Teagasc, Carlow, in 2008 and was investigated for morphological and molecular variation. Morphological traits were measured at the end of the second growing season and were compared with herbarium specimens of *Miscanthus*. Vegetative and inflorescence traits were scored and analysed using basic summary statistics, tests of normality and Principal Components Analysis (PCA). A large degree of morphological variation was recorded in the collections. The PCA of herbarium specimens was able to separate some species from others but there was also considerable overlap among species in the ordination, especially *M. sacchariflorus*, *M. sinensis*, *M. condensatus* and *M. floridulus*. These are known to be closely related and can interbreed. The PCA of the specimens from the Oak Park collection was less informative because of missing data due to lack of inflorescences (accessions did not flower). It was clear that morphology alone is often insufficient to distinguish taxa especially when inflorescence characters and ploidy information is lacking.

The ploidy level of the accessions in the collection was evaluated through flow cytometry. The ploidy included di-, tri- and tetraploids. All individuals labelled as *M. ×giganteus* showed a triploid status, together with the newly bred *M. sacchariflorus×M. sinensis* hybrids. Most *M. sinensis* were diploids. *Miscanthus sinensis* Tea-62 was triploid and comparable to the value of the *M. ×giganteus*. A different situation was found for other non-diploid *M. sinensis*, in particular four *M. sinensis* 'Goliath' and the *M. sinensis* 'Zebrinus' Tea-33. In these the ratio measured by the flowcytometer was in between the values of the triploid *M. giganteus* and tetraploid *M. sacchariflorus* standards. The 'Goliath-like' hybrid is likely an autotriploid with three *M. sinensis* haploid sets, whereas *M. ×giganteus* is an allotriploid that is supposed to have two genomes from *M. sinensis* and one from *M. sacchariflorus*, which has a lower amount of DNA per haploid genome.

DNA sequences of the internal transcribed spacer of the nrDNA were obtained for 76 genotypes in the collection and compared for polymorphism. The SNPs were particularly

useful for differentiating *M. sinensis*, *M. sacchariflorus* and *M. ×giganteus* accessions and in combination with ploidy and morphology offer high potential for taxon identification.

To gather more markers for population level diversity and differentiation studies, new microsatellite markers for both plastid and nuclear genomes were developed. For the development of plastid markers the chloroplast genome information of *Saccharum officinarum* was used. The nuclear SSRs (nSSRs) were developed from the sequences of 192 clones obtained from microsatellite enriched library. New primer pairs for the amplification of nineteen nuclear loci and six chloroplast loci were developed. Both chloroplast (cpSSR) and nSSR primers were used to characterise DNA variation, to help establish gene pools and to better understand hybridization and introgression. Huge genotypic variation was found within the genus, mostly in the species *M. sinensis*. The markers showed wide utility across a large number of *Miscanthus* species and also some closely related genera. The analysis of the cpSSRs showed a high number of different haplotypes but with a clear bias in allele composition between *M. sinensis* and the two species *M. sacchariflorus* and *M. ×giganteus*, thus confirming *M. sacchariflorus* as the maternal lineage of the hybrid *M. ×giganteus*. The nSSRs were found to be highly polymorphic across the collection and transferable to closely related genera such as *Saccharum*. The new markers were also used in UPGMA clustering and Bayesian structuring analysis to group individuals according to their similarity. Three major clusters of individuals were defined using the Bayesian STRUCTURE analysis with nuclear markers (nSSRs) and two with plastid markers (cpSSRs).

In conclusion, the morphological, ploidy, sequence and microsatellite results highlighted the high level of diversity still unexplored in the genus and have clarified taxon identity of many accessions in the collection. A large set of new markers have been developed for the plant breeding and systematics community. The newly developed markers will be useful to further explore this diversity and to select useful traits for breeding of new and improved genotypes for biomass production.

Table of contents

List of Figures

List of Tables

<div align="center">

Chapter 1

General introduction to the characterisation of genetic and morphological diversity of a collection of *Miscanthus*

</div>

1.1 Introduction

1.1.1 *Miscanthus*

Miscanthus is a perennial rhizomatous C_4 grass genus native to East Asia, where it is found in a wide range of climatic conditions. The genus belongs to the '*Saccharum* complex' together with *Erianthus, Narenga, Saccharum* and *Sclerostachya* due to the ability of the five genera to produce fertile offspring. Some authors include the southern African species known as *Miscanthidium* in the genus *Miscanthus* on the basis of morphology but it is likely that their close similarity is due to convergence (Hodkinson et al. 2002a). *Miscanthus sensu stricto* (s.s) is well defined and has a basic chromosome number of 19 compared to most of its close relatives that have x=10 (Linde-Laursen 1993).

Miscanthus is native to eastern or south-eastern Asia. Its natural range extends from northeastern Siberia, 50°N, in the temperate zone to Polynesia 22°S, in the tropical zone, westward to central India and eastward to Polynesia (Clifton-Brown et al. 2008). *Miscanthus* species are therefore adapted to a wide range of climatic zones, from the coast up to high mountain, and to different habitats. Some species such as *M. floridulus* (Labill.) Warb. generally grow best at sea level in tropical climates, whereas other species such as *M. paniculatus* (B. S. Sun) Renvoize & S. L. Chen can tolerate high elevation conditions (Chen and Renvoize 2006).

Miscanthus was introduced in Europe in the 19th century as an ornamental plant. Later, in 1935, the hybrid *M. ×giganteus* Greef et Deuter ex Hodkinson & Renvoize, was collected in Yokohama, Japan, by the Danish botanist Aksel Olsen and was distributed throughout Europe. This hybrid has raised interest in the last decades as a potential bioenergy crop due to its ability to produce high yields of biomass (Clifton-Brown et al. 2008).

1

1.1.2 *Miscanthus* for energy production

Currently, *Miscanthus* is ranked among the top nine potential perennial energy crops (Głowacka et al. 2010). Some characteristics of *Miscanthus* make these plants particularly attractive over other crops for bioenergy. *Miscanthus* is a C_4 genus belonging to the tribe Andropogoneae which all share C_4 photosynthesis (C4-NADP type). In fact nearly half (ca. 4500 spp.) of all grass species (ca. 11,000 spp.) are C_4 (Grass Phylogeny Working Group II 2012). C4 grasses dominate in climates with sporadic rainfall, as they can photosynthesise better than C3 plants under warm temperature water stress. In such conditions they show a higher growth rate compared to C_3 plants (Monteith 1978), thanks to a better water utilization during CO_2 fixation.

In addition, unlike other C_4 grasses, *Miscanthus* maintains high levels of photosynthesis at low temperature, with two key enzymes, pyruvate orthophosphate dikinase (PPDK) and Rubisco less affected by cold than its C_4 relatives (Naidu et al. 2003), thus allowing *Miscanthus* species to adapt to a broad range of climatic conditions including temperate regions where C_3 grasses usually predominate (Chen and Renvoize 2006).

Miscanthus is a perennial outbreeding grass genus. At the end of the growing season, minerals are translocated to the rhizomes, allowing the plant to re-use these nutrients in the following growing season (Jones and Walsh 2001). For agriculture this has the benefit of reducing the amount of fertilizer needed in the following season. It is also resistant to pests and diseases, and it is often hybridized with *Saccharum* in sugarcane breeding programmes to transfer such genes (James 2004).

Since 1983 field trials of *Miscanthus* ×*giganteus* have been carried out in Northern Europe, followed in 1993 by trials in Southern Europe, proving its potential in biomass production, as well as its limits. Field trials of *Miscanthus* ×*giganteus* in the 1990's did not survive the first winter after establishment in cold regions of Northern Europe (Pude et al. 1997), where *M. sinensis* hybrids were found to perform better (Clifton-Brown and Lewandowski 2000). It is an active area of research for Teagasc Research Station, Carlow Ireland (Figure 1.1.1). Recent field trials in China have also demonstrated the adaptability of the species and shown

2

it to tolerate a wide range of habitats. For example, Yan et al. (2012) used field trials on *M. sinensis*, *M. sacchariflorus* and *M. lutarioriparius* (B. S. Sun) Renvoize & S. L. Chen across a range of sites in China and were able to select genotypes best suited to each of the contrasting habitats.

(a)

Figure 1.1.1 (a) *Miscanthus ×giganteus* trial at Oak Park Research Centre, Carlow . Photo by John Finnan, with permission.

(b)

Figure 1.1.1 (b) Harvesting of dry culms in a *Miscanthus ×giganteus* trial at Oak Park Research Centre, Carlow. Photo by John Finnan, with permission.

As a sterile hybrid with 2n=3x=57 chromosomes, *Miscanthus ×giganteus* could be propagated only vegetatively through tissue culture or rhizome division. In nature *Miscanthus* reproduces through seeds, and the possibility of using direct sowing would reduce the costs of field establishment. Furthermore, clonality increases the susceptibility of *Miscanthus* fields to pests and diseases (Clifton-Brown et al. 2008).

For all these reasons, attention has recently turned towards the putative parents of *Miscanthus ×giganteus,* i.e. *M. sinensis* and *M. sacchariflorus* to obtain new hybrids (Jones and Walsh 2001). Among all *Miscanthus* species, *M. sinensis* has the widest geographical distribution in Asia, reflected in a considerable phenotypic variation for crucial traits, whereas *M. sacchariflorus* has a more limited distribution (Clifton-Brown et al. 2008). *Miscanthus floridulus* also has a wide distribution but is more tropical in its distribution, extending out in to Indonesia and the pacific from SE Asia but it is not generally considered a cold tolerant

4

genus (Hodkinson et al. (2002b); Chapter 2 this thesis). Therefore the species with most potential for biomass and bioenergy production are considered to be *M. sacchariflorus*, *M. sinensis* and *×giganteus*. *Miscanthus sinensis* is distinguished from *M. sacchariflorus* and *M. ×giganteus* by its awned spikelets and shorter callus hairs (Figure 1.1.2). It also generally does not have culm buds that are characteristic of the other two species. Distinguishing *M. ×giganteus* from *M. sacchariflorus* requires detailed genetic analysis and ploidy determination (see Chapter 4)

(a)   (b)
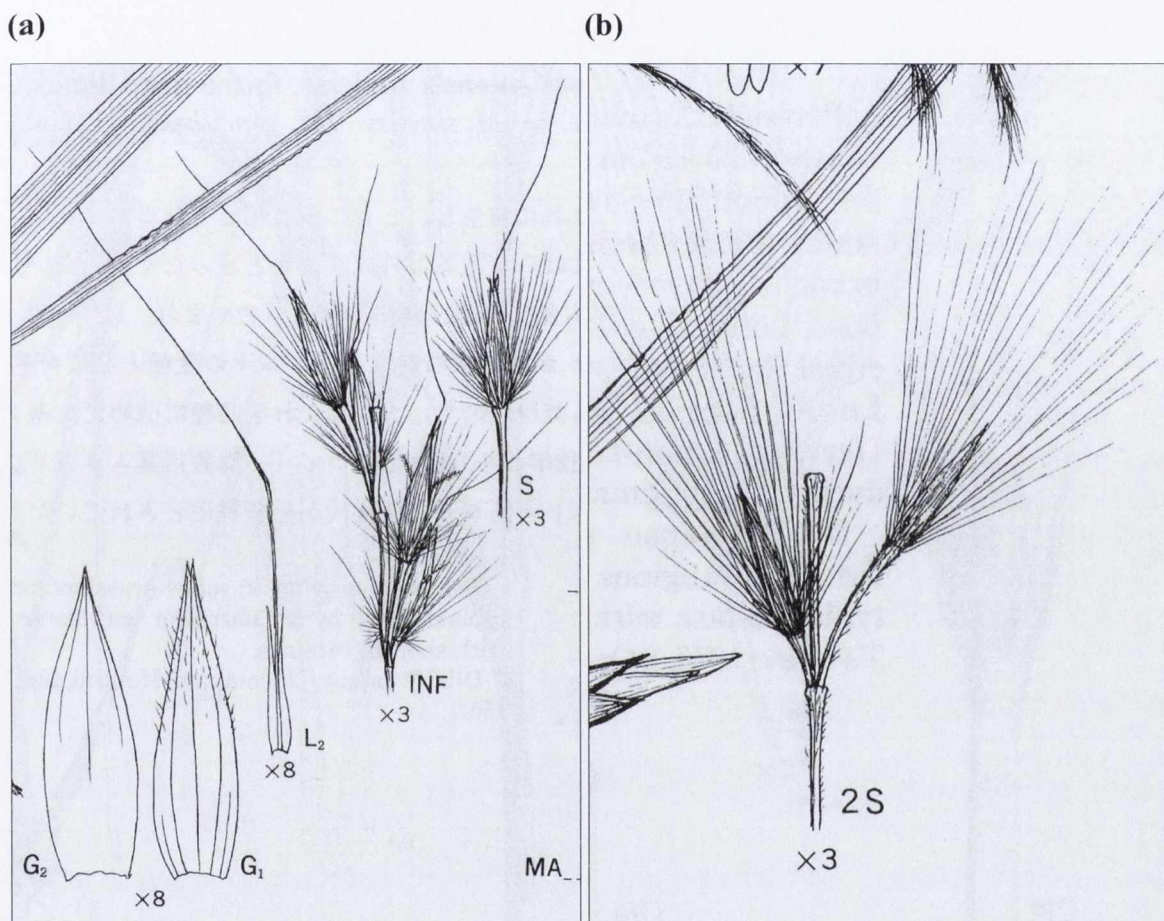
Figure 1.1.2 (a) *Miscanthus sinensis* and (b) *M. sacchariflorus* and *M. xgiganteus* spikelets. G=glumes, L=lemma, S=spikelets, INF=inflorescence (Modified from Osada (1989).

In order to make crosses, the first problem to overcome is the delay in flowering time between the two species. Tests carried out at five different locations in Europe showed that *M. sinensis* is day neutral while in *M. sacchariflorus* some genotypes requires similar conditions to *M. sinensis* for flowering and some others are day sensitive (Lewandowski and Clifton Brown 2000).

Flowering time is an important trait that could also affect yield quality and quantity in *Miscanthus* (Jensen et al. 2011). In trials of *Miscanthus* genotypes, Jensen et al. (2011) showed that *M. sinensis* genotypes were the earliest to flower and differences in flowering time across the entire collection ranged from 160 to 334 days (June to November), and photoperiods between 7.8 and 16.6 h, in Wales, UK. Early flowering shortens the growing season, but when plants do not flower before the autumn frost in northern regions, the reuptake of nutrient by rhizomes is less effective, resulting in the loss of important elements for growth as well as a higher ash content (Clifton-Brown et al. 2008).

The composition of the biomass is also influenced by the amount of fertilizer used, the genotype and the harvest time, that should follow the ripening, because leaves contribute most to ash, and allow for the translocation of nutrients (Clifton-Brown et al. 2008). In Europe, most of the *Miscanthus* is used in combustion, in both straw-burning power station and in co-combustion with coal. Combustion in pure biomass-burning power stations and production second generation fuels such as ethanol are going to be the future utilizations for *Miscanthus* (Vermerris 2008).

1.1.3 Other uses of *Miscanthus*

Paper pulp production

The European deficit in new cellulose fibre has raised interest in the use of non-woody crops for paper pulp production. The raw materials which are most widely used are straw, bagasse and bamboo. *Miscanthus sacchariflorus* is one of the most used raw materials in China. Investigation had been carried out both in China and in different European countries in order to improve the yield and quality of the paper pulp produced using conventional and innovative processes (Jones and Walsh 2001).

Building materials

Miscanthus has been also investigated as a source of fibre to be used in building materials (Jones and Walsh 2001). *Miscanthus* fibre is particularly suitable for the production of medium density fibre-board (MDF) with features comparable with those made from wood chips. A light natural sandwich material (LNS) with wood-based layers and a core of *Miscanthus* stalks has been developed in the Wilhelm-Klauditz-Institute at Braunschweig in Germany (Visser and Pignatelli 2001). LNS could have a wide range of application, substituting plastic or light metal materials as well as wood-based ones.

For centuries *Miscanthus* has been used as thatching material in Japan (Visser and Pignatelli 2001). In Denmark local thatchers have shown interest in substituting reed (*Phragmites australis* L.) with *Miscanthus*, as the quality seems similar. Plots of selected clones of *M. sinensis* have been grown, since this species looks more suitable than *M. ×giganteus*, whose stems are too thick. The stems from these plots were used to thatch small huts. The stability is thought to be the same as for reed. *Miscanthus tinctorius* is also used in Japan for thatching. Both *M. sinensis* and *M. tinctorius* are now planted at the Research Centre Foulum, DIAS (Danish Institute of Agricultural Science, Denmark). Harvesting tests were also carried out showing that it is possible to use existing machines for harvesting *Miscanthus* for thatching with few modifications. In comparison with reed, *Miscanthus* grows also on dry land and the cost for harvesting has been estimated to be at least 50% lower than reed (Jones and Walsh 2001).

Bioremediation

Intensive agriculture, industrialization, and other factors has led in last centuries to the expansion of areas that are badly damaged, contaminated or destroyed by human activity. Since contaminated lands are not suitable for food production, they can be converted to production of non-food crops, such as *Miscanthus,* avoiding aerial dispersion, runoff and improving visual impact (Visser and Pignatelli 2001). *Miscanthus ×giganteus* is able to grow on heavy metal polluted soils as in Cornwall, UK, where its growth and heavy metal uptake was tested on lands subjected for centuries to intense tin mining activity to study the implication of the combustion of such plants for energy production (Visser and Pignatelli

2001). Results show that the uptake of heavy metals was not higher than plants grown on unpolluted soils, even if biomass production was lower, and that heavy metals content is not related to soil concentration.

Another study in Monte de Caparica in Portugal was performed to investigate the use of sewage sludge as fertiliser for *Miscanthus*, monitoring accumulation of heavy metals in the above-ground and below-ground biomass. At harvest, only roots and rhizomes contained a significant amount of metals from the sewage sludge, whereas there was no difference in metal concentration in the above ground fraction between plants grown on polluted and unpolluted soils, allowing the use of such plants for energy production (Jones and Walsh 2001).

Composting

Miscanthus has been also tested as a component for composts, mulches and plant growth substrates. Another possible use is as biological consolidation of wet organic waste. A decrease of organic matter was observed after six months by co-composting *Miscanthus* with sewage sludge and paper pulp effluents (Jones and Walsh 2001). This compost could be good as fertiliser, but there are still problems in cadmium and chromium content (Visser and Pignatelli 2001).

1.1.4 Molecular markers for breeding

Despite all the promising features of *Miscanthus*, breeding for biomass in Europe is still in its infancy. Huge phenotypic variation has been observed in *M. sinensis* and *M. sacchariflorus* for all the interesting traits for biomass production. Starting from its putative parents, breeding of new and improved genotypes of *M. ×giganteus* suitable for different growing conditions throughout Europe, is feasible and underway in several research institutes such as Teagasc, The Institute of Biological, Environmental and Rural Sciences (IBERS) in Aberystwyth Wales, and Plant Research International (Wageningen, the Netherlands). However, there is a need for molecular tools that allow for quick selection of hybrids with desirable traits.

Molecular markers such as simple sequence repeats (SSRs) or single nucleotide
8

polymorphism (SNPs) could be used for Marker Assisted Selection (MAS) (Ribaut and Hoisington 1998) to associate genotypic and phenotypic differences in order to screen new genotypes with no need to wait for the plants to reach a mature phenotype (usually three years in temperate zones) for selection, as well as a first step to map genes of interest along the *Miscanthus* genome (Clifton-Brown et al. 2008). For example a recent paper by Kim et al. (2012) used SSR markers to map 261 loci spanning 40 linkage groups and 1,998.8 cM, covering an estimated 72.7% of the genome.

1.2 General aim of the thesis

The primary aim of this work was to characterise the morphological and genetic diversity of a collection of *Miscanthus* established in Teagasc, Oak Park. Chloroplast and nuclear microsatellite markers, and morphological characters were used to determine genetic diversity, to assess the relationships between genotypes, to classify unidentified individuals, and to develop markers suitable for plant breeding initiatives such as quantitative trait loci (QTL) mapping and MAS.

In detail, the objectives of this thesis were to:

- assess morphological diversity in *Miscanthus*, using measurements of morphological characters from a collection of plants (Chapter 2);
- compare the morphological variation of the collection with the variation observed in herbarium specimens of selected *Miscanthus* species (Chapter 2);
- determine the ploidy level of the accessions in the collection (Chapter 2);
- investigate nuclear DNA variation in a collection of *Miscanthus* accessions using DNA sequencing of the highly polymorphic nuclear ribosomal region (ITS) (Chapter 2);
- design and optimize a new set of chloroplast simple sequence repeat (cpSSR) markers for *Miscanthus* (Chapter 3);
- describe cpDNA allelic and haplotypic diversity and assess the potential of the set of cpSSR markers for the definition of cytoplasmic pools (Chapter 3);
- design and optimize a new set of nuclear simple sequence repeat (nSSR) markers for *Miscanthus* (Chapter 4);
- assess genotypic variation in the collection and relationships between genotypes (Chapter 4)

A peer-reviewed publication has already been published in an international journal from Chapter 3 of this thesis (de Cesare et al. 2010) and others are in preparation for each of the other chapters.

Chapter 2

Morphological and cytological characterization of a collection of *Miscanthus*

2.1 Introduction

2.1.1 Origin and distribution of *Miscanthus*

The genus name *Miscanthus* (from the Greek *mischos* = pedicel and *anthos* = flower) was first used by Andersson in 1855 referring to 15 grass species (Andersson 1855). The genus belongs to the tribe Andropogoneae, subtribe Saccharinae, in the family Poaceae and it is closely related to other genera of the "*Saccharum* complex" including *Erianthus, Narenga, Saccharum,* and *Sclerostachya* (Hodkinson et al. 2002c).

The description of the genus by Chen and Renvoize (2006) is:

"Perennial, tufted or rhizomatous. Culms slender to robust, erect, solid. Leaves basal or cauline; leaf blades large, linear, flat, broad or narrow; ligule membranous. Inflorescence a panicle, often large and plumose, of racemes arranged on a long or short axis; raceme axis tough, internodes slender, spikelets paired, both spikelets pedicelled, pedicels slender, flattened, slightly clavate. Spikelets similar, lanceolate, dorsally compressed; callus bearded with hairs shorter than, as long as, or longer than the spikelet; glumes papery or membranous; lower floret usually represented by a hyaline sterile lemma; upper floret bisexual, lemma hyaline, awned or awnless. Stamens 2–3. Caryopsis oblong or ellipsoid. Fourteen species, mostly in SE Asia and the Pacific Islands, extending to tropical Africa; seven species (two endemic) in China. This genus is readily recognized by its paniculate inflorescence of racemes, which have a tough rachis, and also by its paired spikelets, both of which are pedicelled."

In 1930 Honda divided the genus into two sections, *Triarrhena* and *Eumiscanthus*, including 20 species and 10 varieties. Afterwards, several researchers (Ohwi 1942; Keng 1957; Swallen 1961) disagreed with this classification, reducing the number of recognised species in the genus.

On the basis of both cytological and morphological studies on the Japanese *Miscanthus* species, Adati (1962) divided the genus *Miscanthus* into 3 sections:

1.   *Triarrhena* Honda;
2.   *Eumiscanthus* Honda;
3.   *Kariyasua* Ohwi.

The section *Triarrhena* includes several varieties of *M. sacchariflorus* and it is characterized by creeping stout rhizomes, dense bristles on the leaf sheath when young and culm nodes from which aerial branches and roots develop.

The section *Eumiscanthus* includes *M. sinensis* and *M. floridulus* and lacks branching from culm nodes.

The section *Kariyasua* includes the species *M. tinctorius, M. oligostachyus* and *M. intermedius* all of which are endemic to Japan.

In 1964, Lee separated the Asian species of *Miscanthus* into four sections:

1.   *Triarrhena* Honda;
2.   *Miscanthus* – Eumiscanthus Honda;
•   Subsect. *Sinensis*
•   Subsect. *Miscanthus*
•   Subsect. *Condensati*
3.   *Kariyasua* Ohwi ex Hirayoshi;
4.   *Diandra* Keng.

The section *Diandra* included *M. nepalensis.*

The use of molecular phylogenetics (Hodkinson et al. 2002c) shows that some species included in *Miscanthus s.l.* (*s.l.*: in broad sense) are more closely related to other genera than *Miscanthus.* For example the 4 species from Africa, sometimes classified under the genus *Miscanthidium*, were clearly separate from the Asian *Miscanthus.* These also differ in their basic chromosome number (x=10 instead of x=19). Synonymy is a large problem in the genus. To illustrate the problem of synonymy, Clifton-Brown et al. (2008) listed the species names given in 'The International Plant Names Index' (IPNI at

http://www.ipni.org/index.html). Over 60 *Miscanthus* species are listed but only 11– 12 of these arerecognized as valid names under *Miscanthus s.s.*.

The genus *Miscanthus sensu stricto* can be restricted to a dozen species and one hybrid that are morphologically well characterized (Clifton-Brown et al. 2008):

- *M. condensatus* Hack. (=*M. sinensis* ssp. *condensatus* (Hack.) T. Koyama)
- *M. floridulus* (Labill.) Warb.
- *M. intermedius* (Honda) Honda
- *M. longiberbis* Nakai
- *M. lutarioparius* (B. S. Sun) Renvoize & S. L. Chen
- *M. oligostachyus* Stapf.
- *M. paniculatus* (B. S. Sun) Renvoize & S. L. Chen
- *M. sacchariflorus* (Maxim.) Hack.
- *M. sinensis* Anderss.
- *M. tinctorius* (Steud.) Hack.
- *M. transmorrisonensis* Hayata
- *M. ×giganteus* Greef & Deuter ex Hodkinson and Renvoize

They are all perennial rhizomatous, or sometimes tufted, grasses with erect cane-like stems up to 7m tall (*M. lutarioparius*). The inflorescence is terminal with a cluster of plumose racemes bearing awned or awnless spikelets in pairs, both pedicellate. The inflorescence axis may be short with the inflorescence subdigitate with long racemes, as found in *M. sinensis*, or it may be long bearing short racemes, as in *M. floridulus*.

Miscanthus was introduced in Europe in the 19th century as an ornamental plant. Later, in 1935, the hybrid *M. ×giganteus*, was collected in Yokohama, Japan, by the Danish botanist Aksel Olsen and was distributed throughout Europe (Clifton-Brown et al. 2008).

2.1.2 Morphological description of *Miscanthus s.s.* species

Modified from Osada et al. (1989); Chen and Renvoize (2006)

Miscanthus condensatus

Miscanthus condensatus is similar to *M. sinensis* in its gross morphology and many taxonomists considered this a variety of *M. sinensis*. However some evidence pointed out that this species is completely independent from *M. sinensis* and may derive from *M. floridulus* (Hodkinson et al. 2002a). *Miscanthus condensatus* has a densely tufted culm, erect, 1-2.5 m tall and with a diameter over 2 cm. The leaf blades are 20-80 cm long and 15-40 mm wide, flat, light green in colour, glossy on the above surface and glaucous beneath. The margins appear smooth but look dentate at magnification. The ligules are 2 mm tall and truncate and glabrous. The panicle is made up of densely clustered racemes, thicker and denser than in *M. sinensis*. The spikelets are 5-7 mm long and bearded at their base with 5-8 mm long hairs. Awns are exserted. For distributions of species see Figure 2.1.1.

Miscanthus floridulus

Miscanthus floridulus is densely tufted in large clumps 1.5-3.5 m tall. The leaf blades are flat, 30-80 cm long and 15-35 mm wide with very rough margins, white midrib, glabrous except on their base. The ligules are 2 mm tall and fimbriate on upper margin. The radical leaves in *Miscanthus floridulus* could remain green during the winter. The panicle is two times longer than wide, with numerous racemes 8-20 cm long on an axis 30-50 cm long. The spikelets are 3-3.5 mm long with white hairs 4-6 cm long. The glumes are alike, as long as the spikelets, glabrous or with short hairs on their back. The lower lemma is hyaline and nerveless, the upper one is deeply bifid and bears an awn 8-15 mm long. For distributions of species see Figure 2.1.1.

Miscanthus intermedius

Miscanthus intermedius has tufted or solitary culms, 1-1.8 m tall and with a diameter of 5 mm. The leaf blades are 20-60 cm long and 1-2.5 cm wide, rough on margins, glaucous and sparsely pilose beneath. The ligules are truncate, 1-2 mm tall. The panicle bears 6-10 racemes 10-15 cm long digitated on a short axis. The spikelets are alike, 7-8 mm long with white hairs 5-7 mm long at their base. The two glumes are alike, both 3- or 5- nerved with sparse hairs on their back. The awns are shorter than in *M. oligostachyus* and hardly exserted. The upper

14

lemma is hyaline with a hardly exserted awn 4-7 mm long. *Miscanthus intermedius* has a larger vegetative part than *M. oligostachyus*. For distributions of species see Figure 2.1.1.

Miscanthus lutarioriparius

Miscanthus lutarioriparius is characterised by stout cane-like culms, 3-7 m tall and with a diameter of 10-20 mm at the base, branching at nodes. The nodes are glabrous in the lower part of culms and hairy in the upper part, with lower nodes bearing adventitious roots. The leaf blades are flat and linear, 50-90 cm long and 1.5-3 cm wide, with a prominent midrib, pilose at base and acuminate at apex. The ligule is 0.5mm tall with pilose margin. The panicle is large and with a glabrous main axis and 20-40 racemes 10-30 cm long. The spikelets are 4-6.5 mm long, pilose and without awns, with hairs at base exceeding the spikelet. The glumes are unequal but both with attenuate apex. The lower is 3-5- veined, pilose on its back with 10 mm long hairs; the upper is 3-veined, glabrous on back but with pubescent margins. The lower lemma is lanceolate and hyaline, nerveless and pilose; the upper is similar but smaller. For distributions of species see Figure 2.1.1.

Miscanthus oligostachyus

Miscanthus oligostachyus is characterised by a few tufted or solitary culms, 60-90cm tall and with a diameter of 1-3 mm. The 4-6 nodes along it are tomentose. The blades of the lower leaves are linear, 13-40 cm long and 6-14 mm wide, whereas the upper ones are narrowly lanceolate with an obtuse base and less than 10 cm long. Leaves are thinner and softer than in *M. sinensis*, with smooth margins and are loosely pilose on the lower or on both surfaces. The ligule has a lacerate and ciliate upper margin. The panicle is composed by 2-5 racemes 7-15 cm long, subdigitatelly arranged on the main axis. The spikelets are alike, tawny, 7-8 cm long and with hairs at their base. The glumes are as long as spikelets, the lower 3-nerved and with two teeth and the upper 3- or 5-nerved and acuminate at apex. The lower lemma is hyaline, faintly 1-nerved; the upper lemma is 4-5 mm long, hyaline, awned at apex with an awn 8-15 mm long and exserted. For distributions of species see Figure 2.1.1.

Miscanthus paniculatus

The culms of *M. paniculatus* are 30-100 cm tall with a diameter of 3-4 mm and 3-4 nodes. The leaf blades are flat and linear, 10-40 cm long and 2-8 mm wide, glabrous or pilose, acuminate at apex and narrowed to midrib at base. The ligules are 0.5-1.5 mm tall and ciliate. The panicle is 5-15 cm long, with a glabrous main axis bearing short racemes of 2-6 cm. The spikelets are lanceolate, 5-6 mm long with short hairs at base. The two glumes are unequal. The lower glume, 5 mm long, is pilose on back, faintly 2- or 3- nerved or nerveless, with apex 2-toothed. The upper glume is 6 mm long, faintly 3- or 5- nerved, with pilose margins and apex acuminate. The lower lemma is lanceolate and 4.5-6 mm long, 3- or 5- nerved, the upper one is ~5 mm long, 1- or 3- nerved, acuminate with pilose margins and a straight awn, 2 mm long. For distributions of species see Figure 2.1.1.

Miscanthus sacchariflorus

The culms in *M. sacchariflorus* are solitary and erect, 1-2.5 m tall and with a diameter of 1-1.5 cm at their base. The leaf blades are 20-80 cm long and 1-3 cm wide, with scabrous margins and glaucous beneath. The ligules bear a fringe of short hairs. The panicle is composed by numerous racemes 20-35 cm long and pendulous, subdigitated on a short central axis. Spikelets are paired, 5-6 mm long and bearded on base with white hairs 10-15 mm long. The glumes are both 3- nerved, the lower long as the spikelet, the upper shorter, rounded and hairy on back. The lower lemma is nerveless or faintly nerved, whereas the upper lemma is awnless or short-awned. For distributions of species see Figure 2.1.1.

Miscanthus sinensis

The species *M. sinensis* is characterised by a densely tufted culms, erect, 0.6-2 m. tall and with a diameter of 3-7 mm at their base. The leaf blades are 20-60 cm long and 6-20 mm wide, with rough margins and a prominent white midrib. The ligules are ~1.5 mm tall and shortly ciliate on the upper margin. The panicle is nodding, bearing 10-25 racemes 10-30 cm long on a central axis shorter than racemes. The spikelets are paired, alike, a short and a long pedicellate one, 5-7 mm long with white or purplish hairs 7-12 mm long at their base. The two glumes are equal, as long as the spikelets. The lower is 5- or 7-nerved, whereas the upper

is 3-nerved. The lower lemma is membranous, hyaline and nerveless. The upper lemma is bifid with an awn 8-15 mm long, exserted and geniculate. For distributions of species see Figure 2.1.1.

Miscanthus tinctorius

Miscanthus tinctorius has a loosely tufted culm, 60-100 cm tall, with a diameter of 2-4 mm and tomentose nodes. The leaf blades are 8-20 cm long and 6-12 mm wide, glabrous except at their base. The sheath can be hairy or not. The ligules are rounded and 2-3 mm tall. The panicle is composed by 3-10 racemes 7-12 cm long on a short axis. The spikelets are 5-6 mm long, with short hairs at the base. The glumes are tawny, as long as the spikelets and pilose on their back. The lower is 3- nerved and bifid at apex, the upper one acute. The lower lemma is lanceolate, faintly 1- nerved; the upper lemma is 1-nerved, unawned or short- awned. *M. tinctorius* is smaller than *M. sinensis*, with shorter but erect racemes and lanceolate uppermost leaf. For distributions of species see Figure 2.1.1.

Miscanthus transmorrisonensis

The species *Miscanthus transmorrisonensis* is similar to *M. sinensis* but characterised by leaves less than 5 mm wide and panicle-branches usually not tufted. The panicle has a main axis two thirds as long as the inflorescence and is usually purplish. The spikelets are less than 4 mm long. For distributions of species see Figure 2.1.1.

Miscanthus ×giganteus

Miscanthus ×giganteus has erect culms, unbranched, ~2 m tall with a diameter of 5-10 mm with cauline leaves. The leaf blades are flat, 50-66 cm long and 2.2-2.5 cm wide, with scabrid margins, glabrous and acuminate. The ligules are membranous, 2 mm long and ciliate dorsally. The panicle holds 24 racemes 10-20 cm long on a glabrous axis 12 cm long. The spikelets have glabrous pedicels 1-3.5 mm long. Each spikelet is 2-flowered, lanceolate and 4.5-5.5 mm long. The glumes are both coriaceous, as long as spikelet, acuminate. The lower glume has hairs on the back; the upper one is ciliate on the upper margin. The lemma is hyaline, with ciliate margins. For distributions of species see Figure 2.1.1.
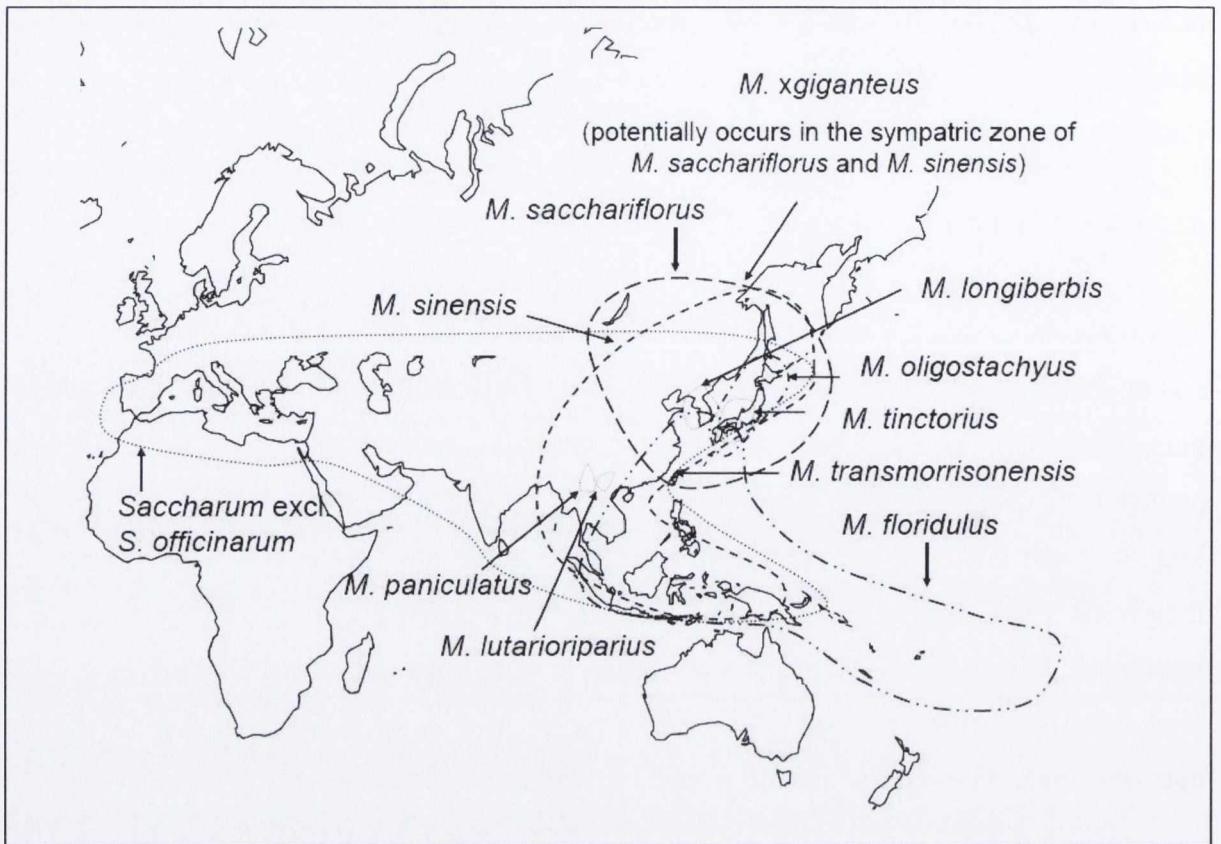
Figure 2.1.1 Geographical distribution of *Miscanthus s.s.* species (source Hodkinson, unpublished; with permission).

2.1.3 Cytogenetics of the genus *Miscanthus*

The basic chromosome number in the genus *Miscanthus* is $x = 19$ (Clayton and Renvoize 1986). *Miscanthus sinensis* usually has $2n = 38$. This species exhibits a regular meiosis. All 46 *M. sinensis* pollen mother cells (PMC) analysed (Linde-Laursen 1993) showed 19 ring bivalents at diakinesis, indicating that this species is a diploid. The karyotype obtained through a chromosome spread performed on root tips includes 2 metacentric satellite chromosomes (SAT- chromosome) with a long proximal nucleolar constriction in their short arms. A diploid number of $2n = 38$ has been reported also for *M. sinensis* var. *gracillimus*, *M. sinensis* var. *variegatus* and *M. sinensis* var. *zebrinus*. Aneuploids in this species have also been observed, ranging from 35-41 chromosomes (Takizawa et al., 1952).

Miscanthus sacchariflorus has a chromosome number of $2n = 76$. 38 bivalents were observed in most meiotic preparations examined by Adati (1958). As expected, *M. sacchariflorus* has only one pair of SAT- chromosomes morphologically similar to the one in *M. sinensis*

18

(Linde-Laursen 1993). A 3*x*- hybrid was produced crossing a diploid *M. sinensis* var. *condensatus* and a tetraploid *M. sacchariflorus* (Hirayoshi et al. 1955). Meiosis in this hybrid showed 17-21 bivalents, univalent and occasionally trivalents at first division. This can be explained assuming that *M. sacchariflorus* has 2 genomes, one from *M. sinensis* and one from an unidentified species with partial homology with *M. sinensis* (it would now be considered to be *M. ×giganteus*).

Miscanthus ×giganteus is sterile: it produces few seeds which give rise to a highly variable offspring (Nielsen 1987) and can be propagated only vegetatively. A first cytogenetic study showed that metaphase preparations from root tips have 2n = 58 or in few cells, 2n = 57 chromosomes (Linde-Laursen 1993). The chromosomes are all metacentric except 8 acrocentric ones, and only 1 SAT- chromosome is present. Some metaphases exhibit small bodies positive to the Feulgen staining used that are thought to be accessory (B) chromosomes. The analysis of PMCs at diakinesis shows few trivalents but an equal number of bivalents and univalents, suggesting the presence of two highly homologous genomes and a third genome with low homology with the two. *M. ×giganteus* is probably an hybrid between a diploid and a tetraploid having a genome in common. The diploid parent is supposed to be *M. sinensis*, whereas the tetraploid one is probably a *M. sacchariflorus*, which has stout rhizomes like *M. ×giganteus* and unlike all the other Japanese species.

A subsequent study disagreed with this finding, suggesting a karyotype for *M. ×giganteus* of 2n = 57 (Lafferty and Lelley 1994), with only metacentric and submetacentric chromosomes found. The presence of 2 SAT- chromosomes very similar in morphology suggests an allotriploid origin for this hybrid. Furthermore, no B-chromosome was observed. Meiosis was irregular and characterized by stickiness of chromosomes.

To confirm the hybrid origins of *M. ×giganteus* a molecular study was necessary (Hodkinson et al. 2002b). AFLP data were used to build a neighbour joining (NJ) tree for *M. ×giganteus* and its putative parental species, *M. sinensis* and *M. sacchariflorus* (Hodkinson et al. 2002b) . DNA fragments obtained from the AFLP analysis with four primer pairs were scored. The NJ tree shows that *M. ×giganteus* is equally distant from both *M. sinensis* and *M. sacchariflorus*, in contrast with a higher distance between the two species. DNA sequencing and cytogenetic analysis using *in situ* hybridisation also confirmed the hybrid nature of *M. ×giganteus*

19

(Hodkinson et al. 2002c). Two different 2n sets has been found in *M. condensatus* (=*M. sinensis ssp. condensatus*): 2n = 38 and 57. The latter exhibits an irregular meiosis with a high number of trivalents (Adati and Mitsuishi 1956; Adati 1958) suggesting that this might be an autotriploid.

Miscanthus oligostachyus and *M. tinctorius* have a diploid set of 2n = 38 with normal bivalents in meiosis and one pair of SAT- chromosome each in somatic cells, whereas *M. intermedius* has been found to be hexaploid with a 2n = 114, of which 6 are SAT-chromosomes, with formation of multimers at diakinesis (Adati 1958).

2.1.4 Genome size studies in *Miscanthus* using flow cytometry

To estimate the ploidy level in plants flow cytometry has become the most popular method. The process requires only a small quantity of fresh leaf material and the results are ready in a few minutes (Doležel et al. 2007). The preparation of the tissue can be divided in two phases: extraction and staining. During the first phase, a small piece of leaf is chopped with a razor blade in a suitable buffer to extract whole nuclei from the tissue. The liquid obtained is then filtrated and stained with a fluorochrome that binds specifically DNA. For ploidy estimation, the fluorochrome of choice is usually DAPI (4',6-diamidino-2-phenylindole), that binds preferentially AT-rich regions, while for measuring the genome size of a species, intercalating stains such as propidium iodide (PI) with no base preference are more suitable. The fluorescence emitted by the stained nuclei is proportional to the DNA amount and it is measured through a flow cytometer (Doležel et al. 2007).

The genome size of *M. ×giganteus*, *M. sacchariflorus* and *M. sinensis* has been evaluated using flow cytometry (Rayburn et al. 2008). The nuclear DNA content was found to be 7.0 pg in triploid *M. ×giganteus,* 5.5 pg and 4.5 pg for diploid samples of *M. sinensis* and *M. sacchariflorus* respectively, even though they share the same chromosome number 2n=2x=38. The DNA content of *M. ×giganteus* is in accordance with the postulated hybrid origin of *M. ×giganteus* resulting from the union of a haploid genome of *M. sinensis* with a diploid genome of *M. sacchariflorus* (Linde-Laursen 1993).

2.1.5 Sequencing of the internal transcribed spacer (ITS) of the nuclear ribosomal DNA

The nuclear ribosomal DNA (nrDNA) is composed by a highly tandem repeated cluster of genes that code for the ribosomal RNA (rRNA) (Brown and Shaw 1998). Multiple copies of the region are homogenized through concerted evolution leading to uniformity in the sequence of rDNA loci. Each cluster contains the genes for the rRNA 18S, 5.8S and 26S, separated by two spacers, ITS-1 and ITS-2, which are transcribed together with the genes but are not part of the final product, allowing them to diverge more quickly compared to the rDNA (Baldwin et al., 1995). Nevertheless, these regions seem to play a role in the cleavage of the rRNA, thus being subject to a certain level of conservation. The ITS sequences shows low level of length variation in closely related taxa. The conservation of length make it easier to compare sequences, that are variable enough to be interesting for phylogenetic analyses (Baldwin et al. 1995; Hodkinson et al. 2010).

Sequencing of the ITS regions have been previously used to investigate the origin of *M. ×giganteus* (Hodkinson et al. 2002c). As pointed out by Hodkinson et al. (2002b), homogenization could occur only through gene conversion, but not unequal crossing-over in sterile hybrids like *M. ×giganteus* and two different parental ITS sequences were still detectable.

2.2 Aims

The aim of this chapter was to assess morphological variation in *Miscanthus* using measurements of morphological characters from a collection of plants using summary statistics and multivariate ordination (PCA) in comparison with herbarium specimens. It also aimed to compare the morphological results to ploidy variation in the collection and DNA sequencing of the internal transcribed spacer (ITS) of the nuclear rDNA.

In detail the objectives were:

- To describe morphological diversity in a collection of *Miscanthus* established in Oak Park, Carlow;
- To compare the morphological data to herbarium specimens to help classification of unidentified individuals;
- To assess ploidy variation in the collection;
- To determine if morphological information describes a similar pattern of diversity as DNA content and sequence.

2.3 Material and methods

2.3.1 Plant material

Rhizomes of 33 *Miscanthus sinensis* were provided by Svalöf Weibull, Sweden; 80 individuals of *M. ×giganteus*, *M. sacchariflorus* and *M. sinensis*, including different ornamental varieties, were collected from TCD Botanic Gardens, Dublin, Ireland; 15 additional genotypes of the three species were made available by the University of Hohenheim, Germany (Clifton-Brown and Lewandowski, 2002).

All the rhizomes collected were potted in the autumn of 2007 and plants were kept in a glasshouse for the winter. Plants were transferred in the field in Oak Park, Carlow, in spring 2008. Plants were spaced 2m apart from each other and arranged in different order in two replicates of three rows of 42 plots each. Stems were cut every year in late winter before the new growing season began. A full list of accessions is given in Table A in the Appendix.

2.3.2 Morphological characterization

Scoring of characters

Each plant was scored in the field in late summer 2009 for the following morphological characters:

- growth habit (spread or clumped);
- space between culms (cm);
- culm wax (yes/no (Y/N));
- maximum culm width (cm);
- culm buds or branching (Y/N);
- plant height (cm);
- approximate leaf number;
- leaf variegation (Y/N);
- internode length (cm);
- maximum leaf length (cm);

23

- maximum leaf width (cm);
- leaf hair (Y/N);
- inflorescence production (Y/N).

Where present, inflorescences were collected, pressed and dried. The following characters were subsequently scored:

- inflorescence length (cm);
- inflorescence axis length (cm);
- axis hairs (Y/N);
- raceme length (cm);
- raceme number;
- raceme hairs (Y/N);
- raceme internode length (cm);
- upper pedicel length (cm);
- lower pedicel length (cm);
- pedicel hairs (Y/N);
- spikelet length (cm);
- spikelet callus hair length (cm);
- awn length (where present).

Herbarium specimens

Specimens belonging to *Miscanthus* were collected from the following herbarium:

- Royal Botanic Gardens, Kew, UK
- Trinity College, Dublin, Ireland
- University of Copenhagen, Denmark
- Royal Botanic Gardens, Edinburgh, UK
- British Museum, UK

The specimens were grouped according to species and the following characters were scored for each sample:

- culm height (cm);

24

- culm width (cm);
- leaf length (cm);
- leaf width (cm);
- leaf hairs (Y/N);
- inflorescence length (cm);
- inflorescence axis length (cm);
- raceme number;
- raceme length (cm);
- raceme internode length (cm);
- peduncle hairs (Y/N);
- axis hairs (Y/N);
- raceme axis hairs (Y/N);
- upper pedicel length (cm);
- lower pedicel length (cm);
- raceme internode length (cm);
- pedicel hairs (Y/N);
- spikelet length (cm);
- spikelet callus hair length (cm);
- awn length (where present).

The species scored for this study, based on availability, are: *M. condensatus*, *M. floridulus*, *M. nepalensis, M. nudipes. M. oligostachyus, M. sacchariflorus, M. sinensis, M. sorghum* and *M. Tinctorius* (Table A in Appendix).

Data analysis

Descriptive statistics for morphological characters were calculated using Minitab® 16.2.0 (2007) to assess the basic properties of data distributions. For quantitative data, means and standard deviation were computed and a histogram for each character was used to display the distribution of data. To determine if characters follow a normal distribution, a normality test was performed using the Kolmogorov-Smirnov (KS) test: characters with a p-value greater than the value from the normality test were considered normally distributed. Where

25

characters were shown not to be normally distributed, transformations were attempted using natural log transformation to achieve a normal distribution for the transformed dataset.

A principal component analysis (PCA) was performed on both quantitative data from herbarium specimens and the field collection in Oak Park, Carlow. A two-dimension scatterplot was constructed to visualise the distribution of the eigenvalues and eigenvectors for each sample using Minitab® 16.2.0 statistical software.

2.3.3 Evaluation of ploidy

A few cm^2 of fresh leaf material from each accession was chopped with a razor blade in a Petri dish together with leaf material of a non-Miscanthus plant (*Ilex*) as internal standard and an ice-cold buffer to extract intact nuclei from the plant cells. The DNA buffer contained 5 mM Hepes, 10 mM magnesium sulphate heptahydrate, 50 mM Potasium chloride, 0.2 % Triton X-100, 0.1 % DTT (Dithiothreitol), 2 mg/l DAPI at pH 8, modified after (Arumuganathan and Earle 1991). DAPI is a fluorescent dye which complexes with double-stranded DNA to give a product that fluoresces at 465 nm. After chopping, 2 ml of the buffer solution is passed through a nylon filter of 50 μm mesh size. The solution with stained nuclei is sent through a CyFlow ML (Partec GmbH, Otto Hahnstrasse 32, D-4400 Münster, Germany) flowcytometer with a high-pressure mercury lamp. When the samples are run with the appropriate filter-settings for excitation, the fluorescence of the stained nuclei is measured by a photomultiplier and converted into voltage pulses. These voltage pulses are electronically processed to yield peak signals and to produce DNA histograms that are then analysed using Flomax version 2.4d (Partec).

A ratio between the fluorescence of the sample and the internal standard was calculated for each accession and compared to the ratio from three samples of known ploidy used as references: *M. sinensis* 'Strictus' as diploid standard, *M. ×giganteus* as triploid standard, and *M. sacchariflorus* as tetraploid standard.

2.3.4 DNA sequencing

The internal transcribed spacer (ITS) of the nuclear ribosomal DNA (nrDNA) was amplified by PCR for 81 accessions using the primer pairs 17SE-ITS2 and ITS3-26SE (White et al. 1990; Sun et al. 1994) for the ITS-1 and ITS-2 regions respectively. A template DNA volume of 5 µl ($40ng\mu l^{-1}$) was amplified with an initial denaturation of 1 min at 97°C followed by 30 cycles each with a denaturation of 1 min at 97°C, 1 min at a 51°C and an extension of 3 min at 72°C, followed by a final extension at 72°C for 7 min. The reaction mixture (final volume of 50 µl) contained 10× reaction buffer (New England Biolabs) containing 2mM $MgSO_4$, 0.125µM dNTPs, 0.25µM of each primer, 0.5U of *Taq* DNA polymerase (New England Biolabs). The PCR products were sequenced by a commercial sequencing company (AGOWA GmbH, Germany) and the electropherograms were viewed by using Chromas Lite version 2.01 (Technelysium Pty Ltd, Australia). Sequences were subsequently aligned using ClustalW (Thompson et al. 1994) as implemented in MEGA version 3.1 (Kumar et al. 2001).

2.4 Results

2.4.1 Morphological characterization

Summary statistics for herbarium specimens

Mean values and standard deviation were computed for a set of herbarium specimens of the species *M. condensatus*, *M. floridulus*, *M. nepalensis*, *M. nudipes. M. oligostachyus*, *M. sacchariflorus*, *M. sinensis*, *M. sorghum* and *M. tinctorius* (Table 2.4.1, raw data in Table B in Appendix) and fitted curves, as implemented in Minitab® 16.2.0 statistical software, were built for each character to display the results (Figure 2.4.1).

Table 2.4.1 Summary statistics for qualitative traits in herbarium specimens of *M. condensatus, M. floridulus, M. nepalensis, M. nudipes. M. oligostachyus, M. sacchariflorus, M. sinensis, M. sorghum* and *M. tinctorius*. N = number of samples; N*= number of missing value; SE Mean= standard error of mean; StDEv= standard deviation; Min= lowest value; Median= middle of the range data; Max= higher value; Q1-Q3= first and third quartile; * = absent.

Species	N	N*	Mean	SE Mean	StDev	Min	Q1	Median	Q3	Max
Culm height										
M. sinensis	4	6	62.3	25	49.9	0	12	66.5	108.3	116
M. sacchariflorus	8	2	97.4	14.9	42.1	40	62.3	90	137.8	163
M. floridulus	8	3	65.63	3.38	9.55	50	57.5	66.5	73.75	79
M. condensatus	9	2	87.1	15.7	47.2	35	46	80	118.5	183
M. oligostachyus	9	0	71.44	8.45	25.34	37	51	72	83	122
M. tinctorius	4	0	117.6	18	36	74	81.5	120.8	150.6	155
M. sorghum	3	0	133.3	28.5	49.4	92	92	120	188	188
M. nudipes	6	1	68.2	15.6	38.3	25	25	75	103.8	106
M. nepalensis	13	3	69.62	7.18	25.9	30	52	64	90	126
Culm width										
M. sinensis	10	0	0.5	0.0537	0.17	0.2	0.4	0.5	0.625	0.8
M. sacchariflorus	10	0	0.32	0.0389	0.1229	0.2	0.2	0.3	0.425	0.5
M. floridulus	11	0	0.4909	0.061	0.2023	0.2	0.4	0.4	0.7	0.9
M. condensatus	10	1	0.76	0.113	0.357	0.5	0.5	0.7	0.825	1.7
M. oligostachyus	9	0	0.1889	0.0111	0.0333	0.1	0.2	0.2	0.2	0.2
M. tinctorius	4	0	0.3	0.0408	0.0816	0.2	0.225	0.3	0.375	0.4
M. sorghum	3	0	0.4667	0.0333	0.0577	0.4	0.4	0.5	0.5	0.5
M. nudipes	7	0	0.2357	0.0322	0.0852	0.15	0.2	0.2	0.3	0.4
M. nepalensis	16	0	0.2625	0.0315	0.1258	0.1	0.2	0.25	0.3	0.5

Table 2.4.1 (*continued*)

Species	N	N*	Mean	SE Mean	StDev	Min	Q1	Median	Q3	Max
Leaf length										
M. sinensis	8	2	53.25	5.59	15.82	27	38	58	63	75
M. sacchariflorus	9	1	41.56	4.25	12.76	24	29	39	53.5	57
M. floridulus	11	0	39.82	5.39	17.87	20	20	40	47	71
M. condensatus	9	2	52.22	5.52	16.57	32	41.5	47	63	87
M. oligostachyus	9	0	24.94	2.66	7.99	12	19.75	23	30	40
M. tinctorius	4	0	39.75	4.21	8.42	32	32.25	39.5	47.5	48
M. sorghum	3	0	62	6.43	11.14	52	52	60	74	74
M. nudipes	7	0	21.14	4.43	11.73	7	10	22	30	40
M. nepalensis	14	2	27.71	3.95	14.79	3	14.75	29.5	39.25	50
Leaf width										
M. sinensis	10	0	0.93	0.175	0.552	0.4	0.575	0.8	1.075	2.3
M. sacchariflorus	10	0	1.03	0.18	0.568	0.3	0.55	0.9	1.525	2
M. floridulus	11	0	1.145	0.233	0.772	0.3	0.8	0.8	1.1	2.7
M. condensatus	10	1	1.59	0.209	0.662	0.6	1.2	1.45	2.05	3
M. oligostachyus	9	0	1.0333	0.0707	0.2121	0.6	0.9	1.1	1.2	1.2
M. tinctorius	4	0	1.225	0.202	0.403	0.8	0.85	1.2	1.625	1.7
M. sorghum	2	1	0.45	0.15	0.212	0.3	*	0.45	*	0.6
M. nudipes	7	0	0.4714	0.036	0.0951	0.4	0.4	0.4	0.6	0.6
M. nepalensis	14	2	0.843	0.227	0.851	0.1	0.35	0.75	1	3.5
Inflorescence length										
M. sinensis	9	1	29.11	4.2	12.61	13	19.5	28	39	50
M. sacchariflorus	10	0	21.9	1.46	4.63	16	17.5	22	25.25	30
M. floridulus	11	0	30.09	2.57	8.53	17	23	33	36	41
M. condensatus	9	2	26.78	2.49	7.48	18	20.5	24	34	38
M. oligostachyus	9	0	11.89	1.27	3.82	5	10	11	15	18
M. tinctorius	4	0	17.88	1.94	3.88	14	14.5	17.25	21.88	23
M. sorghum	3	0	46.67	5.55	9.61	38	38	45	57	57
M. nudipes	7	0	11.93	1.61	4.27	8	8	10.5	16	19
M. nepalensis	16	0	14.84	1.5	6	5	12.25	14.5	16.75	31
Inflorescence axis										
M. sinensis	9	1	12.94	4.05	12.16	0	2	12.5	23.5	33
M. sacchariflorus	9	1	6.59	1.01	3.02	2	5	5.8	9	11.5
M. floridulus	10	1	15.75	2.75	8.7	6	9.38	11.5	22.25	34
M. condensatus	7	4	12.07	2.43	6.44	7	7	9.5	20	22.5
M. oligostachyus	5	4	2.42	0.825	1.846	1	1	2.1	4	5.5
M. tinctorius	3	1	1.333	0.167	0.289	1	1	1.5	1.5	1.5
M. sorghum	3	0	33	6.43	11.14	21	21	35	43	43
M. nudipes	6	1	4.22	1.69	4.13	0.8	1.33	3	6.75	12
M. nepalensis	16	0	5.63	1.12	4.48	0.5	3	4.5	7.42	19
Raceme number										
M. sinensis	10	0	27.6	4.55	14.38	8	17	27.5	35	50
M. sacchariflorus	10	0	22.8	4.58	14.5	6	9.5	22.5	32	50
M. floridulus	11	0	47.27	9.05	30.03	20	20	50	60	100
M. condensatus	9	2	38.89	6.76	20.28	20	25	30	55	80
M. oligostachyus	9	0	2.778	0.324	0.972	1	2	3	3.5	4
M. tinctorius	4	0	7	1.87	3.74	2	3.25	7.5	10.25	11
M. sorghum	3	0	66.67	3.33	5.77	60	60	70	70	70
M. nudipes	7	0	11.29	3.28	8.67	4	4	6	20	25
M. nepalensis	16	0	28.94	7.33	29.33	5	11.25	20	30	100

29

Table 2.4.1 (*continued*)

Species	N	N*	Mean	SE Mean	StDev	Min	Q1	Median	Q3	Max
Raceme length										
M. sinensis	10	0	14.9	0.951	3.007	8	13.125	16	17	18
M. sacchariflorus	10	0	13.95	0.973	3.077	9	11.75	13.5	16.875	18
M. floridulus	11	0	14.91	1.54	5.11	7	12	14	19	23
M. condensatus	9	2	14.78	1.16	3.49	8	12	16	18	18
M. oligostachyus	9	0	10	0.799	2.398	5	8.5	11	11.5	13
M. tinctorius	4	0	14.88	1.05	2.1	13.5	13.63	14	17	18
M. sorghum	3	0	8	1.53	2.65	6	6	7	11	11
M. nudipes	7	0	7.071	0.727	1.924	4	6	7	9	9.5
M. nepalensis	16	0	9.219	0.855	3.42	4	6.25	10	11.75	16
Raceme internode length										
M. sinensis	10	0	0.69	0.0605	0.1912	0.5	0.5	0.7	0.825	1
M. sacchariflorus	10	0	0.58	0.0467	0.1476	0.4	0.5	0.5	0.7	0.9
M. floridulus	11	0	0.4909	0.0368	0.1221	0.3	0.4	0.5	0.5	0.7
M. condensatus	9	2	0.3222	0.0364	0.1093	0.2	0.2	0.3	0.4	0.5
M. oligostachyus	9	0	0.9444	0.0784	0.2351	0.6	0.75	1	1.05	1.4
M. tinctorius	4	0	0.6	0.0408	0.0816	0.5	0.525	0.6	0.675	0.7
M. sorghum	3	0	0.6	0.115	0.2	0.4	0.4	0.6	0.8	0.8
M. nudipes	7	0	0.5571	0.0481	0.1272	0.3	0.5	0.6	0.6	0.7
M. nepalensis	15	1	0.3467	0.0274	0.106	0.2	0.3	0.3	0.4	0.6
Upper pedicel length										
M. sinensis	10	0	0.145	0.0263	0.0832	0	0.1	0.125	0.2	0.3
M. sacchariflorus	10	0	0.195	0.0157	0.0497	0.1	0.1875	0.2	0.2	0.3
M. floridulus	11	0	0.1545	0.0157	0.0522	0.1	0.1	0.2	0.2	0.2
M. condensatus	9	2	0.15	0.0167	0.05	0.1	0.1	0.15	0.2	0.2
M. oligostachyus	9	0	0.2111	0.0261	0.0782	0.1	0.15	0.2	0.3	0.3
M. tinctorius	3	1	0	0.0577	0.1	-0.1	-0.1	0	0.1	0.1
M. sorghum	3	0	0.2333	0.0441	0.0764	0.15	0.15	0.25	0.3	0.3
M. nudipes	7	0	0.1929	0.0202	0.0535	0.15	0.15	0.2	0.2	0.3
M. nepalensis	15	1	0.1333	0.0174	0.0673	0	0.1	0.15	0.2	0.2
Lower pedicel length										
M. sinensis	10	0	0.35	0.0619	0.1958	0	0.2	0.4	0.525	0.6
M. sacchariflorus	10	0	0.46	0.034	0.1075	0.3	0.4	0.4	0.6	0.6
M. floridulus	11	0	0.3455	0.0282	0.0934	0.2	0.3	0.3	0.4	0.5
M. condensatus	9	2	0.3444	0.0377	0.113	0.2	0.25	0.3	0.45	0.5
M. oligostachyus	9	0	0.6667	0.0471	0.1414	0.5	0.5	0.7	0.8	0.8
M. tinctorius	4	0	0.35	0.0645	0.1291	0.2	0.225	0.35	0.475	0.5
M. sorghum	3	0	0.483	0.117	0.202	0.25	0.25	0.6	0.6	0.6
M. nudipes	7	0	0.45	0.0393	0.1041	0.3	0.35	0.5	0.5	0.6
M. nepalensis	15	1	0.28	0.0145	0.0561	0.2	0.2	0.3	0.3	0.4
Spikelet length										
M. sinensis	10	0	0.49	0.0314	0.0994	0.3	0.4	0.5	0.6	0.6
M. sacchariflorus	10	0	0.43	0.0213	0.0675	0.3	0.4	0.4	0.5	0.5
M. floridulus	11	0	0.4	0.0357	0.1183	0.3	0.3	0.4	0.5	0.6
M. condensatus	9	2	0.4889	0.0261	0.0782	0.4	0.4	0.5	0.55	0.6
M. oligostachyus	9	0	0.8556	0.0475	0.1424	0.6	0.75	0.9	1	1
M. tinctorius	4	0	0.575	0.025	0.05	0.5	0.525	0.6	0.6	0.6
M. sorghum	3	0	0.4	0	0	0.4	0.4	0.4	0.4	0.4
M. nudipes	7	0	0.4429	0.0429	0.1134	0.2	0.4	0.5	0.5	0.5
M. nepalensis	16	0	0.3	0.0242	0.0966	0.2	0.2	0.3	0.3	0.5

Table 2.4.1 (*continued*)

Species	N	N*	Mean	SE Mean	StDev	Min	Q1	Median	Q3	Max
Spikelet hairs length										
M. sinensis	10	0	0.52	0.0249	0.0789	0.4	0.5	0.5	0.525	0.7
M. sacchariflorus	10	0	0.98	0.102	0.322	0.6	0.7	0.9	1.3	1.6
M. floridulus	11	0	0.4636	0.0388	0.1286	0.3	0.4	0.4	0.6	0.7
M. condensatus	9	2	0.5	0.0236	0.0707	0.4	0.45	0.5	0.55	0.6
M. oligostachyus	9	0	0.4889	0.0309	0.0928	0.3	0.45	0.5	0.55	0.6
M. tinctorius	4	0	0.3	0.0408	0.0816	0.2	0.225	0.3	0.375	0.4
M. sorghum	3	0	0.2667	0.0333	0.0577	0.2	0.2	0.3	0.3	0.3
M. nudipes	7	0	0.4857	0.0553	0.1464	0.4	0.4	0.4	0.5	0.8
M. nepalensis	16	0	0.675	0.0393	0.1571	0.5	0.525	0.65	0.775	1
Awn length										
M. sinensis	10	0	0.51	0.0407	0.1287	0.3	0.4	0.5	0.625	0.7
M. sacchariflorus	0	10	*	*	*	*	*	*	*	*
M. floridulus	11	0	0.6273	0.0506	0.1679	0.4	0.5	0.6	0.8	0.9
M. condensatus	9	2	0.8556	0.0766	0.2297	0.5	0.7	0.8	1.05	1.2
M. oligostachyus	9	0	0.9222	0.0813	0.2438	0.5	0.75	1	1.05	1.3
M. tinctorius	0	4	*	*	*	*	*	*	*	*
M. sorghum	3	0	0.3333	0.0333	0.0577	0.3	0.3	0.3	0.4	0.4
M. nudipes	7	0	0.871	0.119	0.315	0.5	0.6	0.8	1.3	1.3
M. nepalensis	16	0	1.0313	0.053	0.212	0.5	0.9	1	1.2	1.3
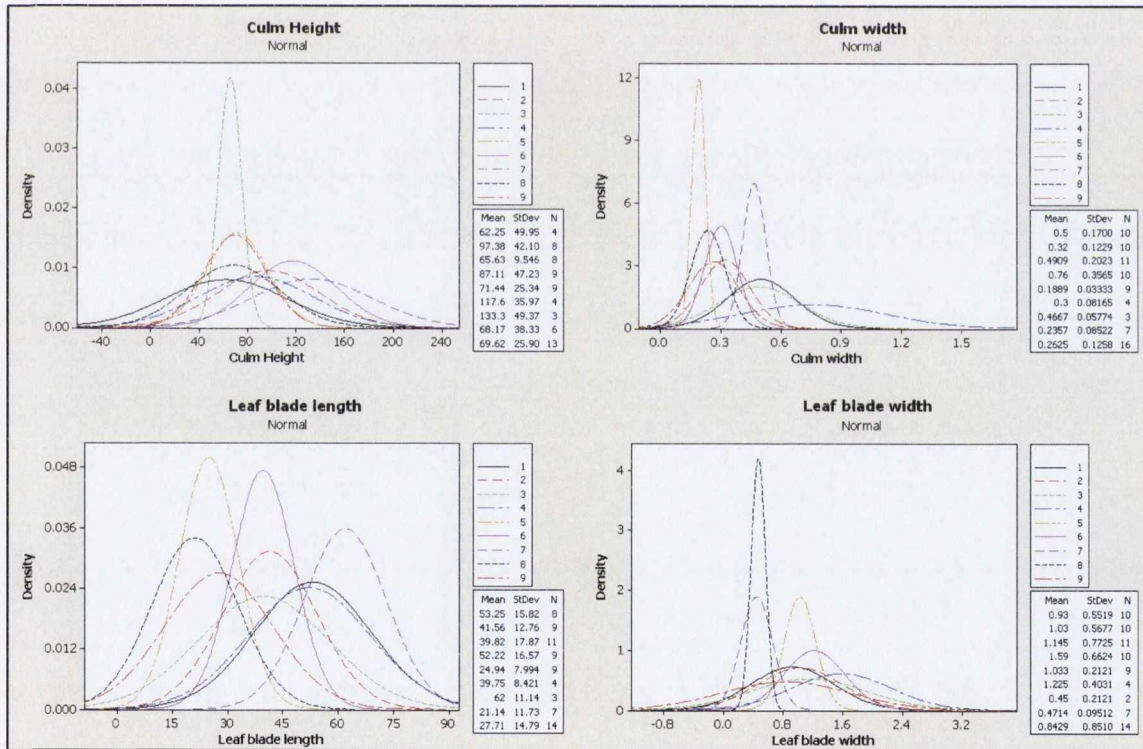
Figure 2.4.1 (*continued*)
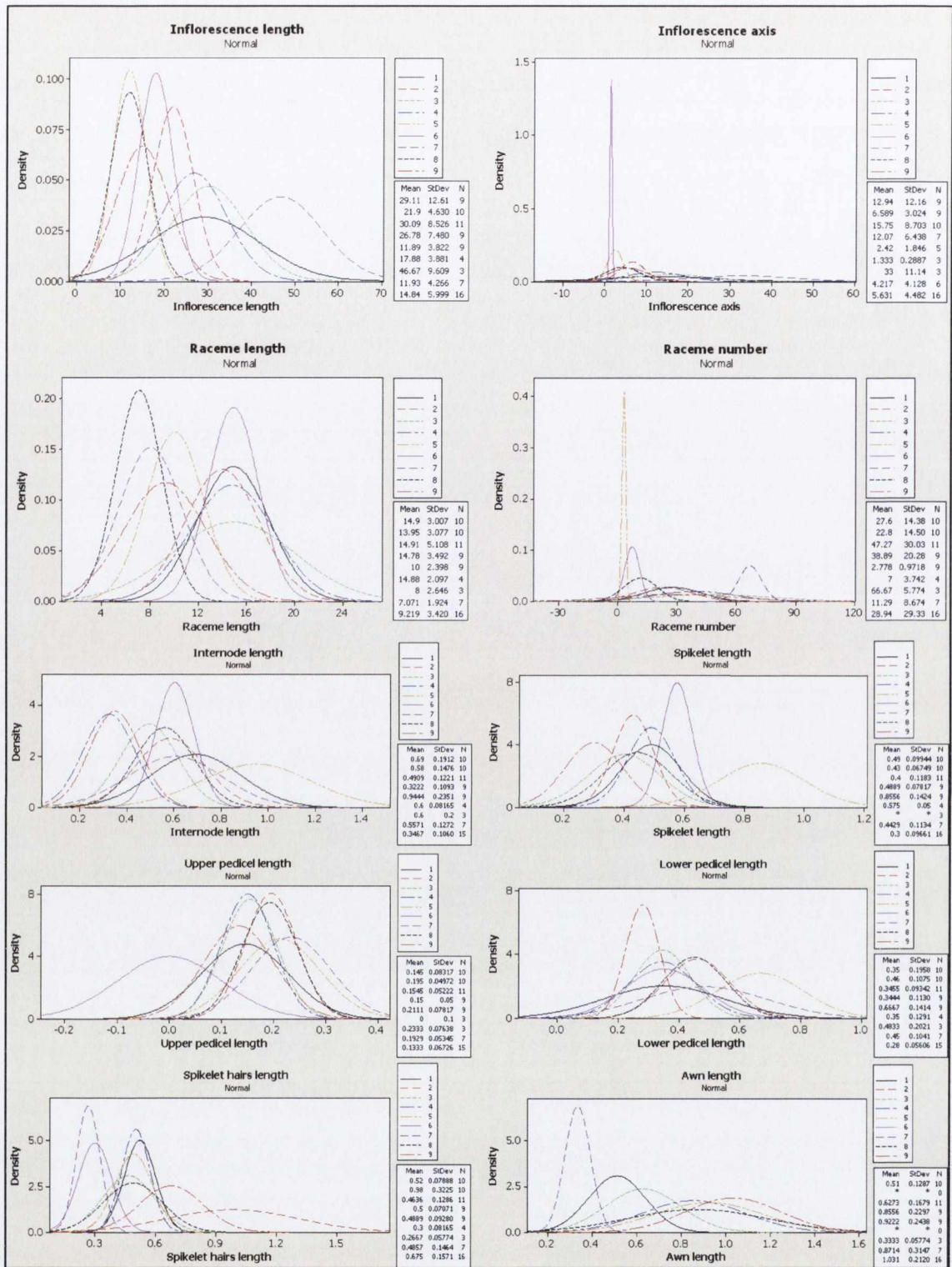
Figure 2.4.1 Fitted curves displaying the distribution of data for herbarium specimens of nine *Miscanthus* species: *1- M. sinensis, 2- M. sacchariflorus, 3- M. floridulus, 4- M. condensatus, 5- M. oligostachyus, 6- M. tinctorius, 7- M. sorghum, 8- M. nudipes, 9-M. nepalensis.* x-axis: length of characters (cm), y-axis: frequency.

32

All the characters analysed showed different mean values among species as well as a high standard deviation. It should be taken into account in the interpretation of these results, that only a small number of plants (and sometimes not an entire plant) was represented in the herbarium samples, thus leading to a possible overestimation of the variation within species.

For the culm height, the highest mean was found in *M. sorghum* (133.3 cm) and the lowest in *M. sinensis* (62.3 cm), the shortest plant being was a *M. nudipes* with a culm length of 25 cm and the tallest an *M. sorghum* of 188 cm. The culm width ranged from 0.1 cm in *M. oligostachyus* (mean = 0.19 cm) to 1.7 cm in *M. condensatus* (mean = 0.76 cm). Where entire leaves were present, length and width at the widest point were measured. The mean length values span from 21.14 cm for *M. nudipes* to 62 cm for *M. sorghum*, and width values range from 0.45 cm to 1.59 cm.

The mean length of the inflorescence varies between 11.89 cm for *M. oligostachyus* and 46.67 cm for *M. sorghum*. *Miscanthus sorghum* also had the longest mean inflorescence axis (33 cm) with the shortest mean value belonging to *M. tinctorius* (1.33 cm). The mean number of racemes for inflorescence ranged from between 2.78 for *M. oligostachyus* to 66.67 for *M. sorghum*, while their length varied between 7.07 cm for *M. nudipes* to 14.9 cm for *M. floridulus* and *M. sinensis*. Along the racemes, the raceme internode length between spikelet pairs was found to be between 0.32 cm for *M. condensatus* and 0.94 cm for *M. oligostachyus*. In the spikelet pairs, the length of the pedicels ranged between 0.28 cm in *M. nepalensis* and 0.67 cm in *M. oligostachyus* for the lower pedicel, and from the highest mean value of 0.23 cm in *M. sorghum* to virtually no pedicel in *M. tinctorius* (for the upper one). The length of the spikelets was found in the range of 0.4-0.6 cm for most of the species, with the exception of *M. nepalensis* (mean = 0.3 cm) and *M. oligostachyus* (mean = 0.86 cm). *Miscanthus sorghum* and *M. tinctorius* had the shortest spikelet callus hairs among species (approximately 0.3 cm on average), whereas *M. sacchariflorus* is characterised by long hairs up to 1.6 cm (mean = 0.98 cm). Both *M. sacchariflorus* and *M. tinctorius* have no awn in their spikelets. Where present, the awn length ranges between 0.33 cm for *M. sorghum* and 0.92 cm for *M. oligostachyus*.

A principal component analysis (PCA) on the data was performed. As shown in Table 2.4.2, the first component accounts for 26% of the total variation, the second for 21% (cumulative 47%) and the third for an additional 12% (cumulative 59%).

Table 2.4.2 Eigenvalues for the first three components of the PCA on herbarium specimens dataset with relative percentage of variation.

Components	1st	2nd	3rd
Eigenvalue	3.6418	2.9088	1.7098
%of variation	0.26	0.208	0.122
Cumulative	0.26	0.468	0.59

The eigenvectors were plotted in a two-dimensional scatterplot (Figure 2.4.2).
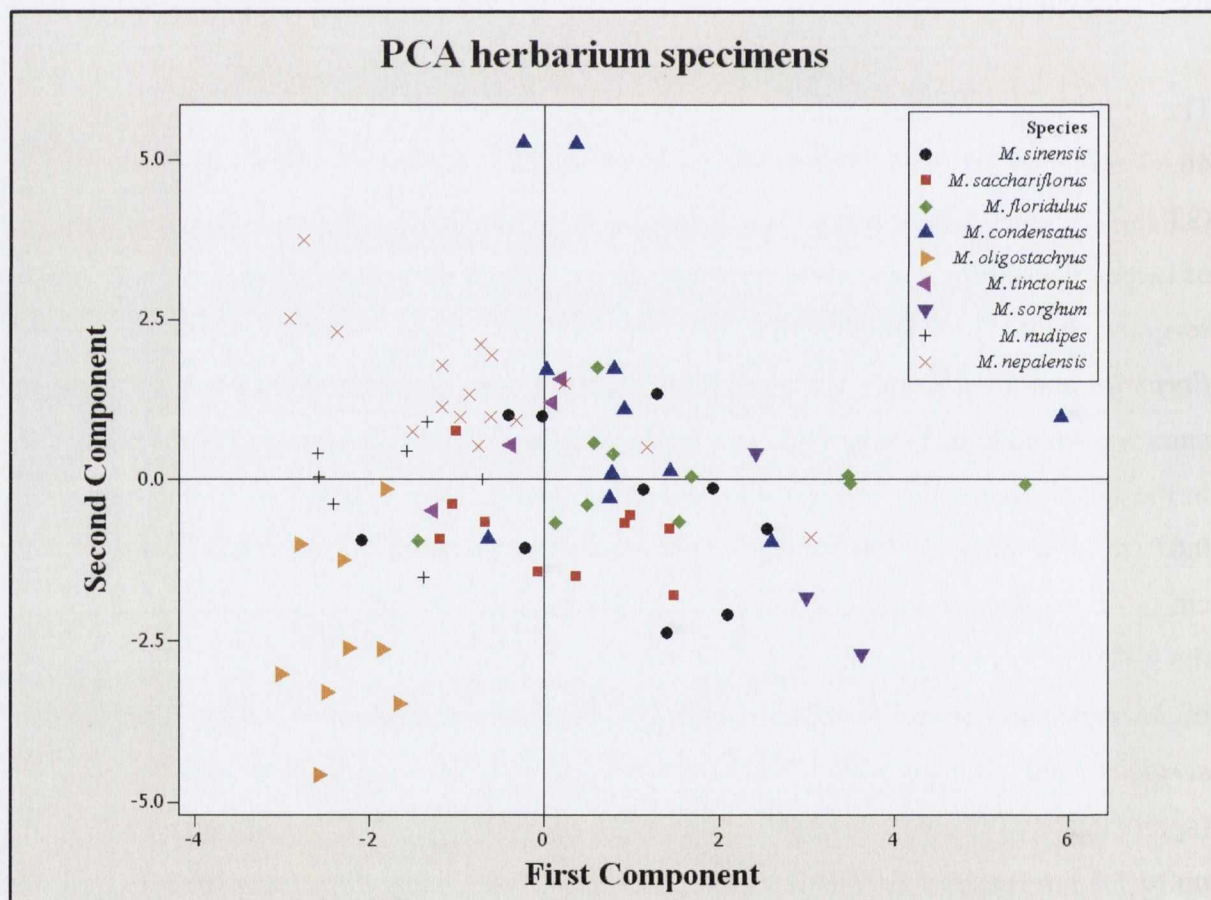
Figure 2.4.2 Principal component analysis displaying the morphological variation in herbarium specimens of nine *Miscanthus* species: *1- M. sinensis, 2- M. sacchariflorus, 3- M. floridulus, 4- M. condensatus, 5- M. oligostachyus, 6- M. tinctorius, 7- M. sorghum, 8- M. nudipes, 9-M. nepalensis.*

34

Some species appeared to be resolved along the two axis (especially in pairwise comparisons): *M. oligostachyus* and *M. nudipes* are well separated between them and from *M. sacchariflorus, M. floridulus, M. condensatus* and *M. sorghum* along the first axis, and from of *M. nepalensis* along the second axis. The two dimensions were not able to separate *M. sacchariflorus, M. floridulus, M. condensatus* and *M. sinensis.*

Summary statistics for the Oak Park collection

Mean values and standard deviations were calculated for quantitative traits in both replicates of the *Miscanthus* collection in Oak Park, Carlow (Table 2.4.3). Histograms were built to display the results (Figure 2.4.3).

Table 2.4.3 Summary statistics for 17 morphological traits in the Oak Park collection for each of the two replicates established in the field. N = number of samples; N*= number of missing values in the dataset; SE Mean= standard error of mean; StDEv= standard deviation; Min= lowest value; Median= middle of the range data; Max= higher value; Q1-Q3= first and third quartile; /1 and /2= from replicate 1 and 2 respectively..

Replicate 1	N	N*	Mean	SE Mean	StDev	Min	Q1	Median	Q3	Max
Leaf length/1	118	2	52.57	1.2	13.02	29	42.19	50.88	64.19	80.75
Leaf width/1	118	2	1.7323	0.0443	0.4813	0.2875	1.4188	1.7375	2.0875	3.075
Space between culms/1	116	4	8.323	0.638	6.875	0	4	5.75	10	36
Plant height/1	116	4	131.05	3.85	41.5	2.2	100	130	160	230
Max culm width/1	116	4	0.7871	0.0259	0.2793	0.3	0.5	0.7	1	1.5
Leaf number/1	116	4	183.9	10.9	117.8	8	90	160	240	600
Internode length/1	115	5	9.859	0.355	3.808	2.8	7	9.5	12	22
Inflorescence length/1	59	0	25.975	0.648	4.974	14	23	26	29	38
Inflorescence axis/1	59	0	10.61	0.601	4.616	0.5	8	11	13	19
Raceme length/1	59	0	15.5	0.362	2.781	11	13	15	18	22
Raceme number/1	59	0	28.22	1.22	9.4	8	22	26	35	53
Internode length/1	59	0	0.578	0.0149	0.1146	0.4	0.5	0.6	0.6	1.1
Upper pedicel length/1	59	0	0.4932	0.0132	0.1015	0.3	0.4	0.5	0.6	0.7
Lower pedicel length/1	59	0	0.2051	0.0082	0.0628	0.1	0.2	0.2	0.2	0.3
Spikelet hairs length/1	59	0	0.7051	0.018	0.1382	0.5	0.6	0.7	0.8	1
Spikelet length/1	59	0	0.4559	0.0074	0.0565	0.4	0.4	0.5	0.5	0.6
Awn length/1	52	7	0.4808	0.0213	0.1534	0.2	0.4	0.5	0.6	0.8

Replicate 2	N	N*	Mean	SE Mean	StDev	Min	Q1	Median	Q3	Max
Leaf length/2	117	7	49.63	1.06	11.51	30	39	48.75	60.25	74
Leaf width/2	117	7	1.7439	0.0348	0.3769	0.875	1.5	1.725	1.9875	2.7
Space between culms/2	107	17	7.86	0.527	5.454	1	4	6	10	32
Plant height/2	109	15	134.68	4.06	42.37	30	110	140	160	220
Max culm width/2	109	15	0.8009	0.024	0.2504	0.3	0.6	0.8	1	1.4
Leaf number/2	109	15	166.3	10.6	110.4	7	80	150	240	520
Internode length/2	109	15	9.138	0.333	3.478	2	6	9	11	20
Inflorescence length/2	56	0	27.161	0.748	5.601	15	23	27.5	30	39
Inflorescence axis/2	56	0	13.179	0.565	4.23	4	11	13.5	16	23
Raceme length/2	56	0	15.161	0.398	2.977	11	13	15	17	21
Raceme number/2	56	0	32.84	1.67	12.49	10	22.25	34	40	65
Internode length/2	56	0	0.5625	0.0162	0.1214	0.3	0.5	0.5	0.6	1
Upper pedicel length/2	56	0	0.5	0.0132	0.0991	0.3	0.4	0.5	0.6	0.7
Lower pedicel length/2	56	0	0.2036	0.0095	0.0713	0.1	0.2	0.2	0.3	0.3
Spikelet hairs length/2	56	0	0.7214	0.0163	0.1217	0.5	0.6	0.7	0.8	1
Spikelet length/2	56	0	0.4732	0.009	0.0674	0.4	0.4	0.5	0.5	0.6
Awn length/2	49	7	0.4755	0.0181	0.1267	0.2	0.4	0.5	0.6	0.8
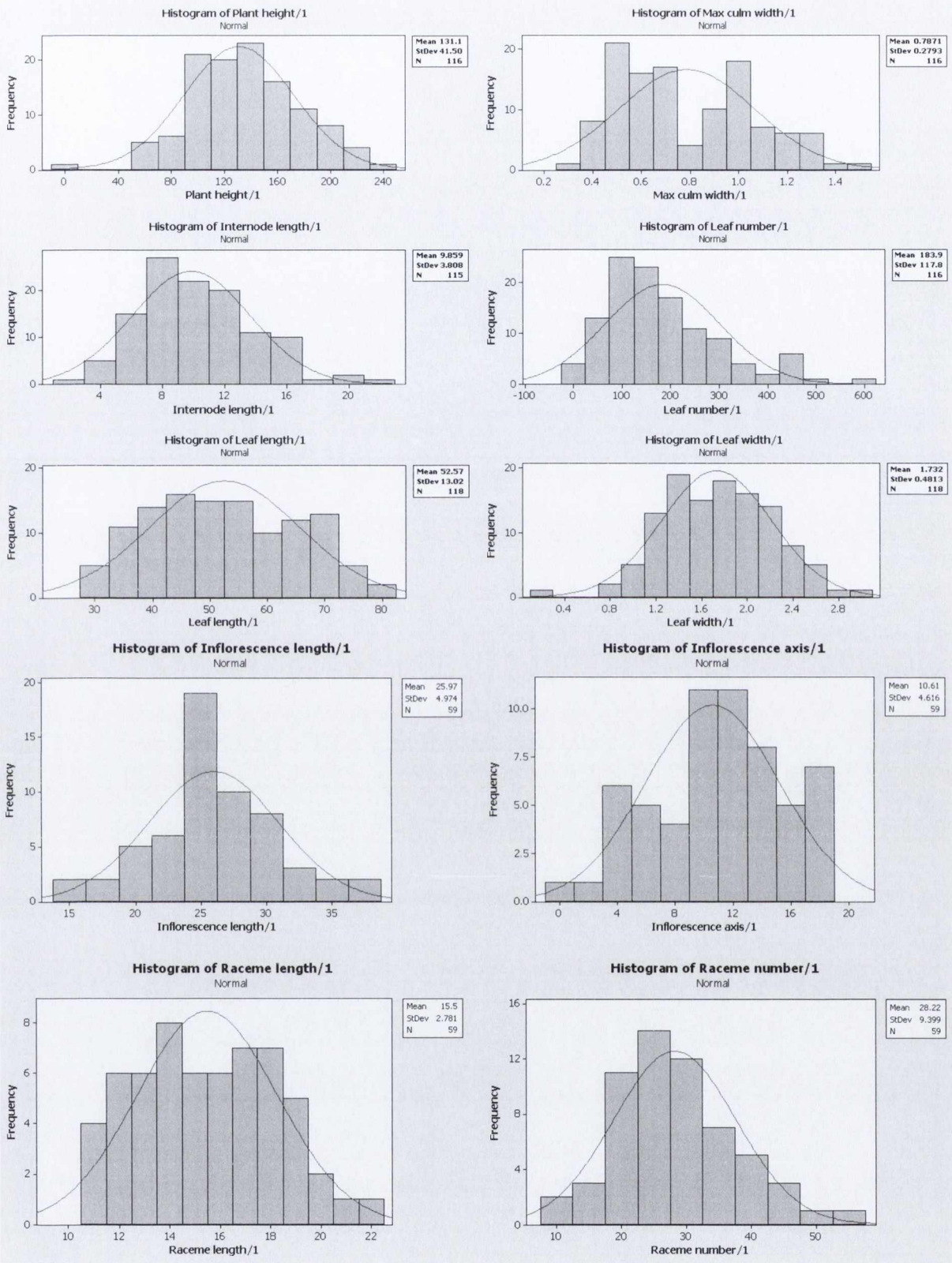
Figure 2.4.3 (*continued*)
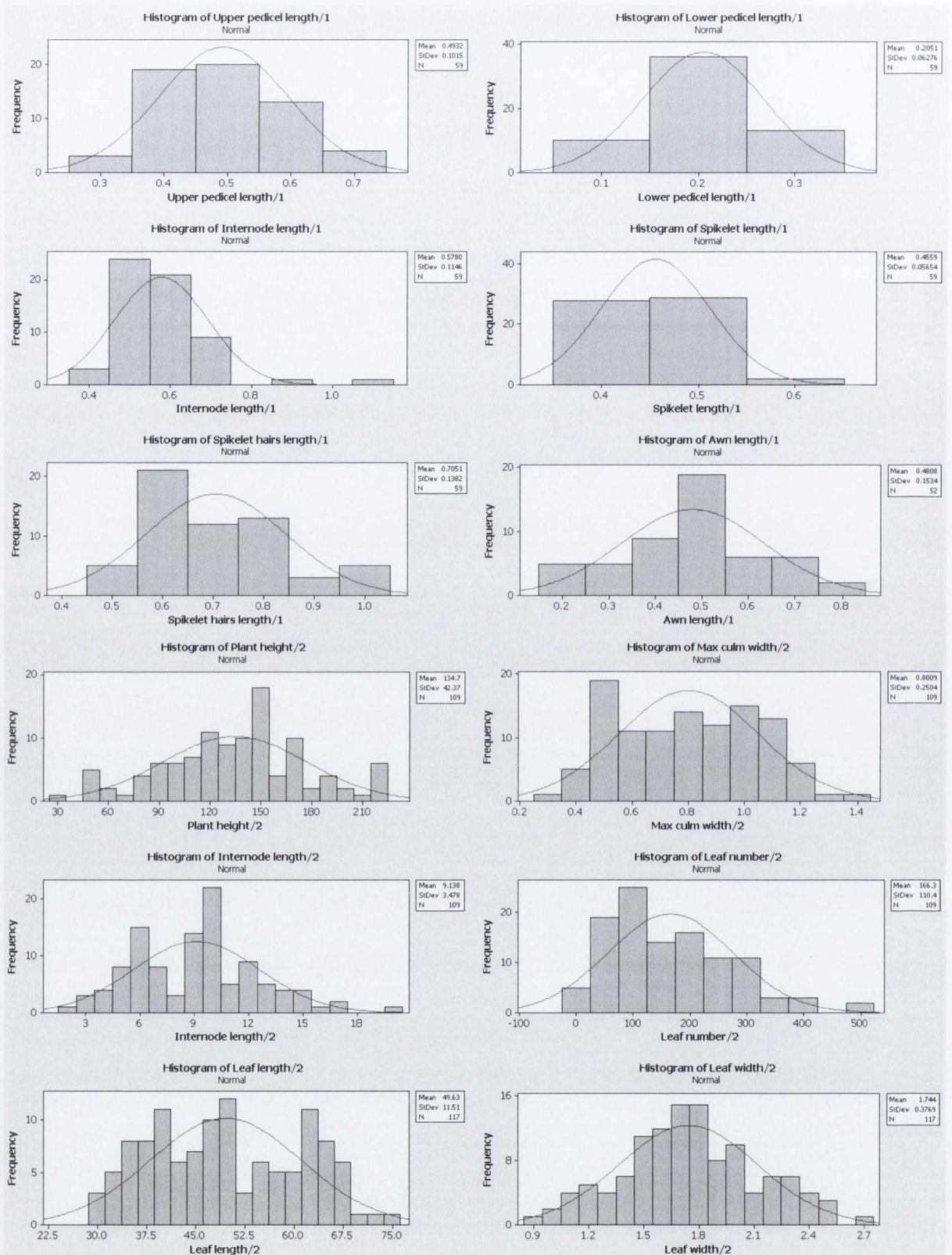
Figure 2.4.3 (*continued*)
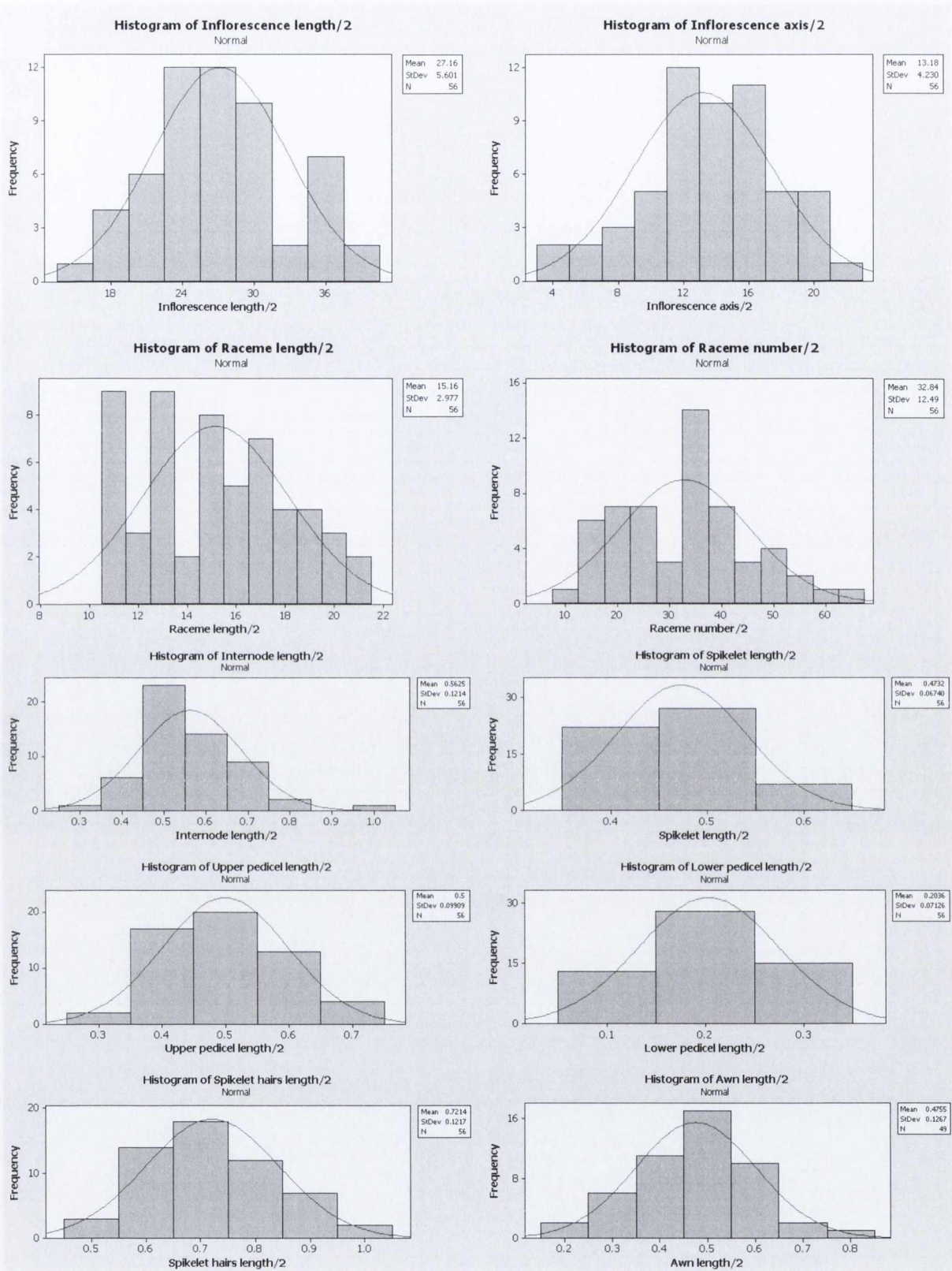
Figure 2.4.3 Histograms and relative fitting curves displaying the distribution of data for the Oak Park collection for each of the two replicates established in the field. x-axis: length of characters for the first (/1) and second (/2) field replicates (cm); y-axis: frequency.

None of the histograms showed a clear normal distribution appearance in the shape of a bell curve. To test for normality, a Kolmogorov-Smirnov test was performed. The results of the test are summarized in Table 2.4.4 and probability plots are displayed in Figure 2.4.4.

Table 2.4.4 Kolmogorov-Smirnov (KS) statistics and p-values for each log transformed character in the two replicates. /1 and /2= from replicate 1 and 2 respectively.

Replicate 1	KS	p-value
Leaf length/1	0.085	0.04
Leaf width/1	0.065	>0.150
Plant height/1	0.07	>0.150
Max culm width/1	0.165	<0.010
Leaf number/1	0.132	<0.010
Internode length/1	0.095	<0.010
Inflorescence length/1	0.091	>0.150
Inflorescence axis/1	0.093	>0.150
Raceme length/1	0.112	0.066
Raceme number/1	0.126	0.029
Internode length/1	0.237	<0.010
Upper pedicel length/1	0.194	<0.010
Lower pedicel length/1	0.312	<0.010
Spikelet hairs length/1	0.217	<0.010
Spikelet length/1	0.313	<0.010
Awn length/1	0.184	<0.010

Replicate 2	KS	p-value
Leaf length/2	0.093	0.02
Leaf width/2	0.058	>0.150
Plant height/2	0.093	0.03
Max culm width/2	0.119	<0.010
Leaf number/2	0.112	<0.010
Internode length/2	0.118	<0.010
Inflorescence length/2	0.092	>0.150
Inflorescence axis/2	0.09	>0.150
Raceme length/2	0.141	<0.010
Raceme number/2	0.077	>0.150
Internode length/2	0.232	<0.010
Upper pedicel length/2	0.196	<0.010
Lower pedicel length/2	0.252	<0.010
Spikelet hairs length/2	0.195	<0.010
Spikelet length/2	0.262	<0.010
Awn length/2	0.189	<0.010

According to the Kolmogorov-Smirnov test, the length of the inflorescence and of its axis had a normal distribution, together with the leaf width. The result of the test for the plant height and the raceme number is unclear since they appear to have a normal distribution in only one out of two replicates.
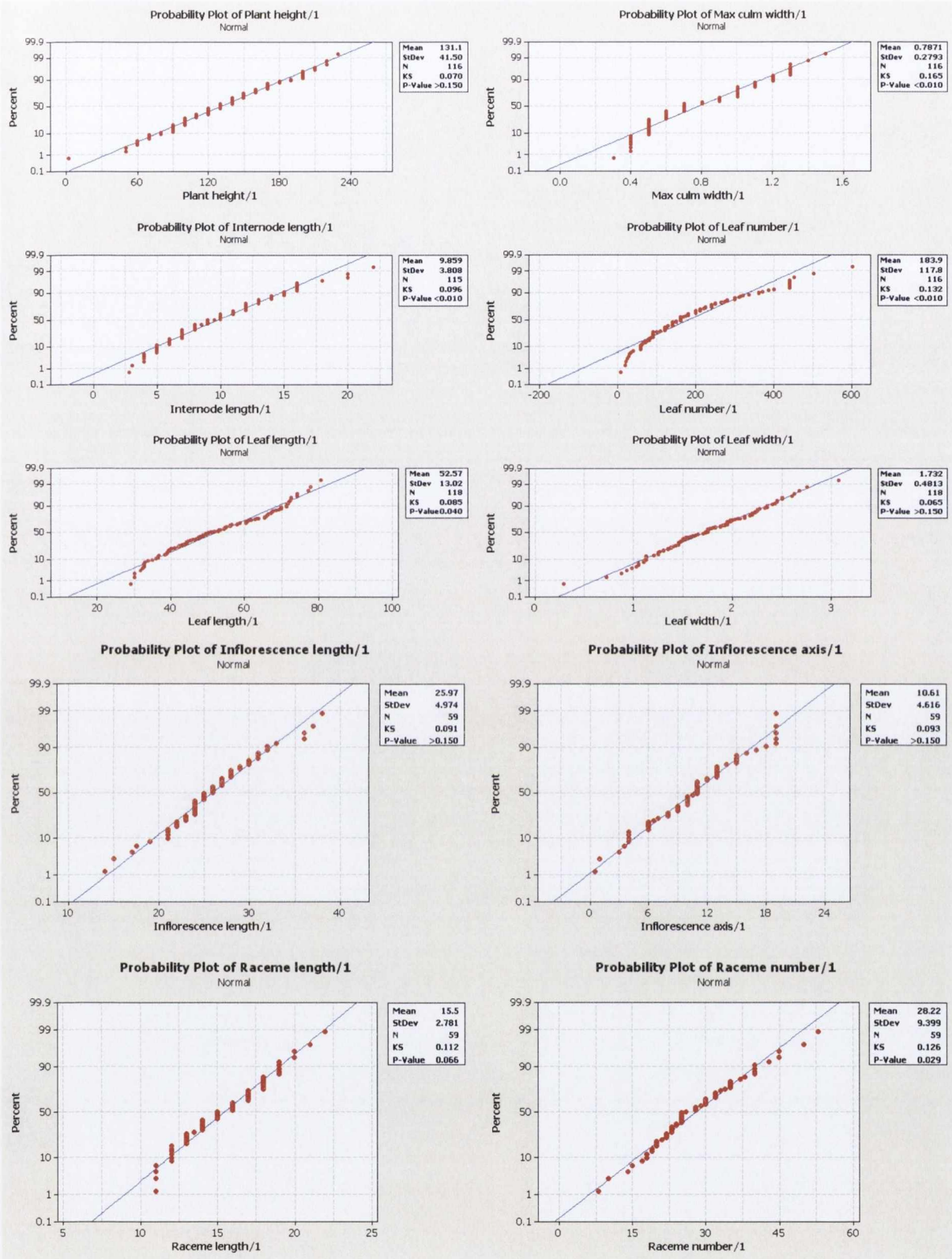
Figure 2.4.4 (*continued*)
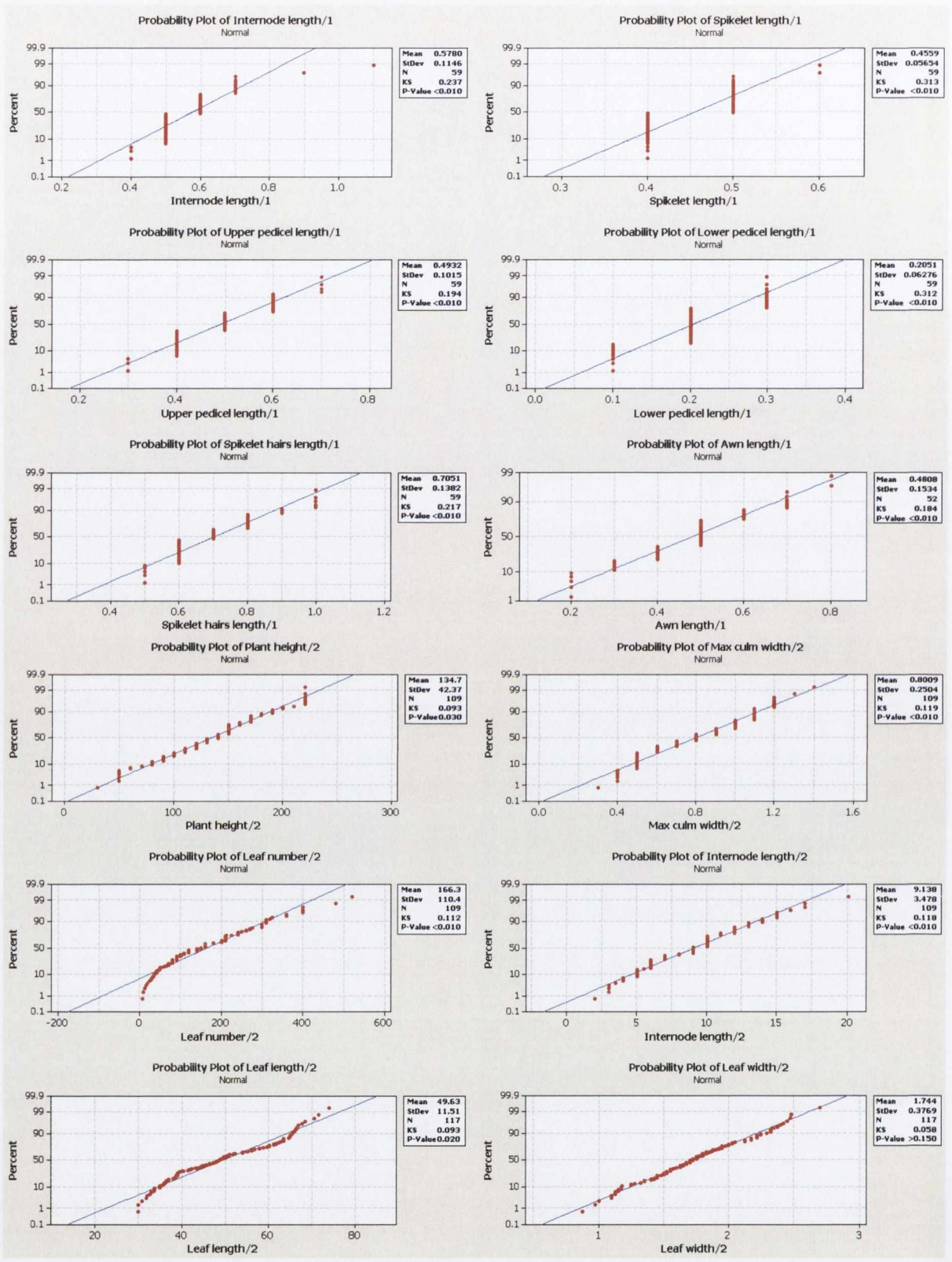
Figure 2.4.4 (*continued*)
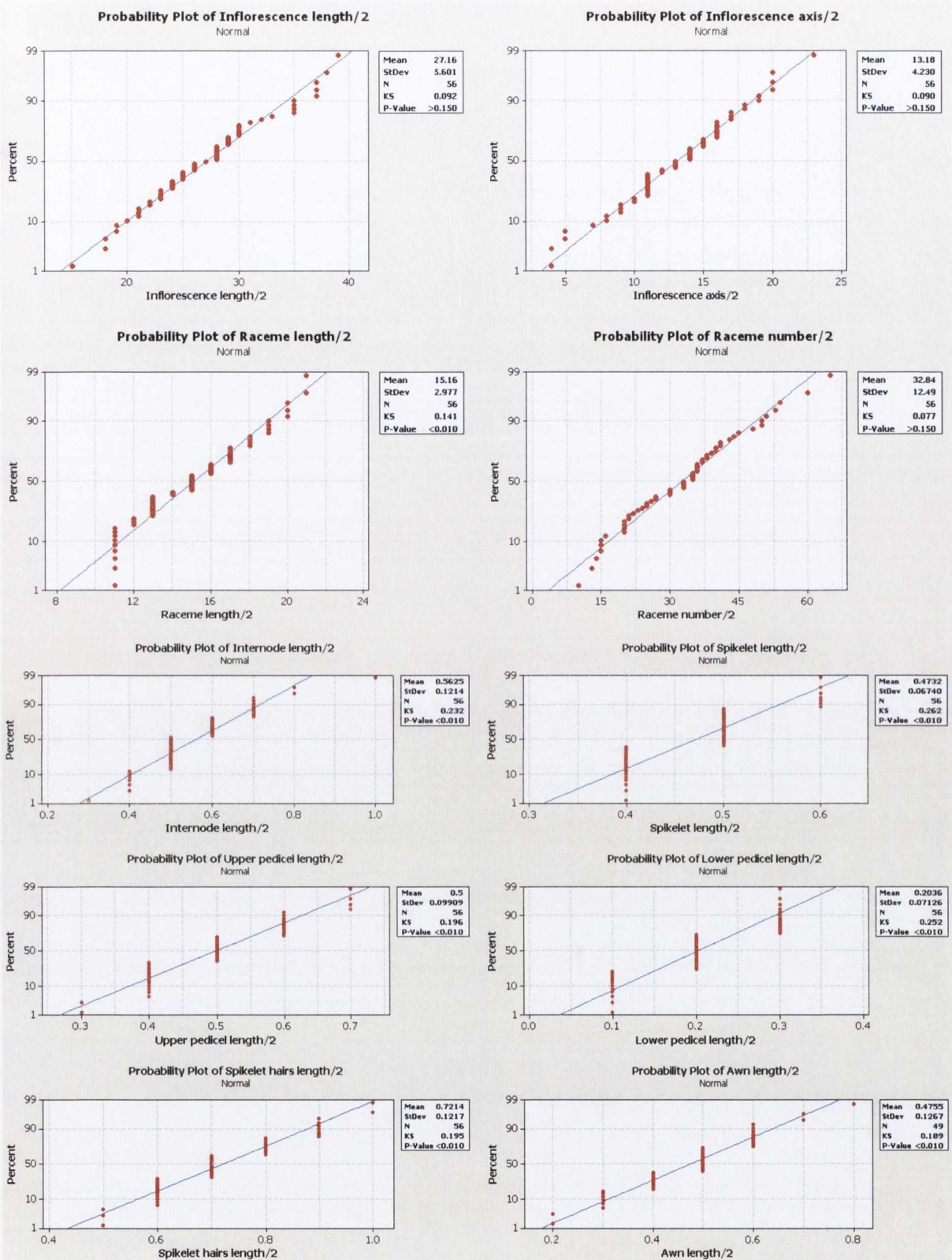
Figure 2.4.4 Plots using the Kolmogorov-Smirnov test for each character in the two replicates. x-axis: length of characters for the first (/1) and second (/2) field replicates (cm); y-axis: percentile.

Data for non-normally distributed characters were transformed using a log transformation in the attempt to obtain a normal distribution for data that showed a skewed distribution: histograms were constructed and normality was tested using the Kolmogorov-Smirnov test statistics. Only the log transformed data for the raceme number in replicate 1 showed a normal distribution with a p-value greater than the KS statistic (Figure 2.4.5).
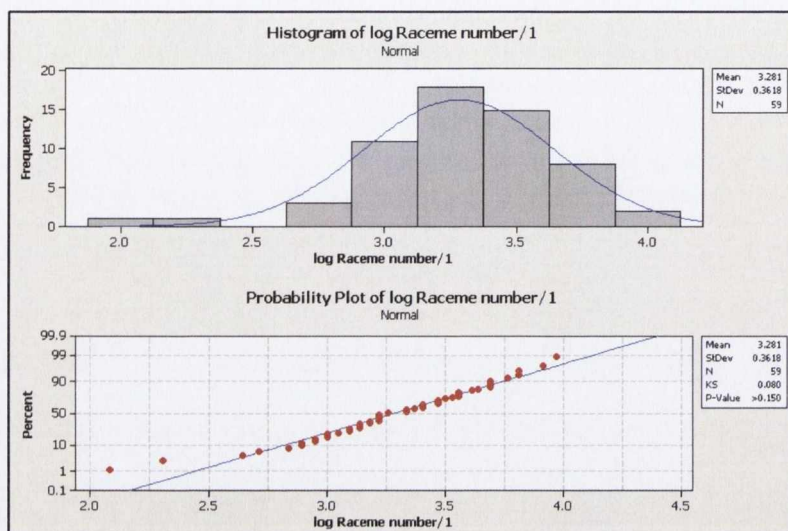
Figure 2.4.5 Histogram with fitted normal distribution curves and plots using the Kolmogorov-Smirnov test for the log transformed character raceme number in the first replicate. x-axis: natural logarithm of length of character (cm); y-axis histogram: frequency; y-axis plot: percentile.

A principal component analysis was undertaken on data from the first replicate including the herbarium samples of *M. sinensis* and *M. sacchariflorus* in the calculation for comparison. The resulting eigenvalues and percentage of variation for the first three components are shown inTable 2.4.5.

Table 2.4.5 Eigenvalues of the PCA of field measurements from the first replicate.

Components	1st	2nd	3rd
Eigenvalue	9.1698	1.7565	0.6618
Proportion	0.655	0.125	0.047
Cumulative	0.655	0.78	0.828

The first component accounts for 66% of the total variation in the dataset, the second component explains 13% of variation (cumulative 78%) and the third an additional 5% (cumulative 83%). A two-dimensional scatterplot of the eigenvectors is displayed in Figure 2.4.6.
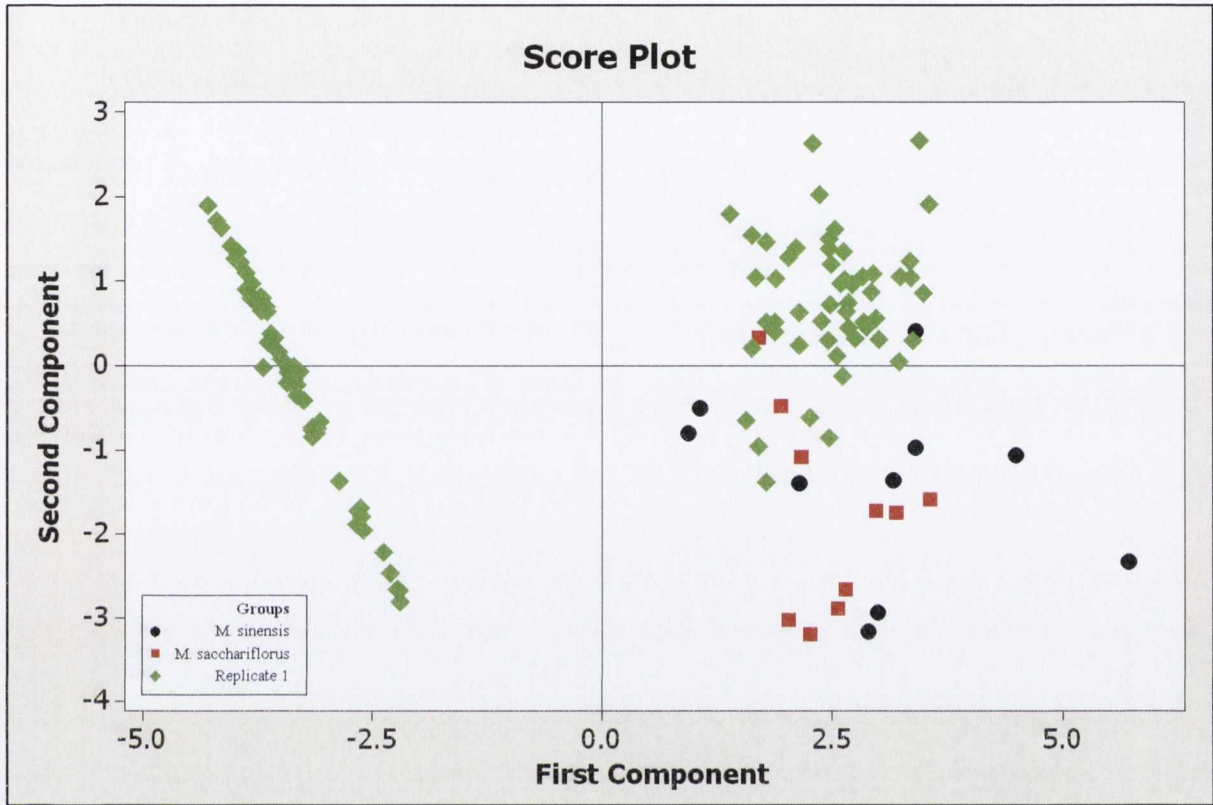
Figure 2.4.6 Principal component analysis displaying the morphological variation in the first replicate of the Oak Park collection (green). Data for herbarium specimens of *M. sinensis* (black) and *M. sacchariflorus* (red) were included as reference. % of variation displayed in the scatterplot = 78%.

As expected from the high value of the first eigenvalue, two groups are clearly separated along the x-axis, the one on the left side of the graph including all the plants that did not flower during the season 2009. For the plants that produced inflorescences (right side of the scatterplot) no clear pattern is visible as well as for the specimens used as reference.

2.4.2 Evaluation of ploidy
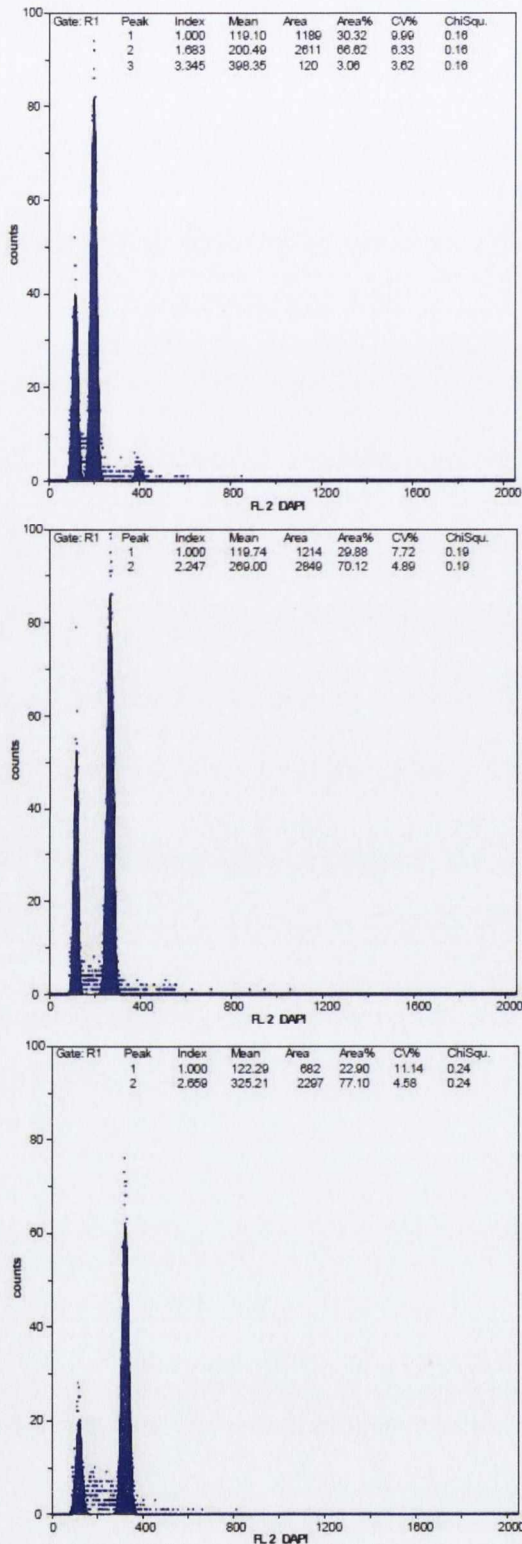
The ploidy of the *Miscanthus* collection established in Oak Park were evaluated using flow cytometry. Three samples for different levels of ploidy were used as standards: a diploid *M. sinensis* 'Strictus', a triploid *M. ×giganteus* and a tetraploid *M. sacchariflorus* (Figure 2.4.7).

Figure 2.4.7 Flow cytometry results for, from top to bottom, a diploid *M. sinensis* 'Strictus', a triploid *M. ×giganteus* and a tetraploid *M. sacchariflorus* used as standard to estimate the ploidy.

The level of ploidy was estimated as ratio between the fluorescence of each samples and the fluorescence of an internal standard not related to *Miscanthus* (*Ilex*). The obtained ratio was then compared to the ones from the three *Miscanthus* standards of known ploidy. The results are summarised in Table 2.4.6.

Table 2.4.6 Ploidy variation in the Oak Park collection. The ratio of the fluorescence of each samples and the fluorescence of an internal standard not related to *Miscanthus* is reported together with the ploidy estimated by comparing the ratio of each sample with the diploid *M. sinensis* Strictus', triploid *M. ×giganteus* and tetraploid *M. sacchariflorus* used as reference.

ID	Ratio	Ploidy
M. sinensis 'strictus' 2X	1.68	Diploid standard
M. xgiganteus 3X	2.25	Triploid standard
M. sacchariflorus 4X	2.66	Tetraploid standard
M. condensatus Tea-44	1.77	2x
M. sinensis 'gross fontane' Tea-35	1.72	2x
M. sinensis 'gross fontane' Tea-36	1.73	2x
M. sinensis 'malaparteus' Tea-61	1.74	2x
M. sinensis 'sirene' Tea-58	1.69	2x
M. sinensis Tea-100	1.74	2x
M. sinensis Tea-101	1.80	2x
M. sinensis Tea-102	1.78	2x
M. sinensis Tea-103	1.74	2x
M. sinensis Tea-104	1.67	2x
M. sinensis Tea-105	1.77	2x
M. sinensis Tea-106	1.70	2x
M. sinensis Tea-107	1.72	2x
M. sinensis Tea-108	1.72	2x
M. sinensis Tea-109	1.72	2x
M. sinensis Tea-110	1.76	2x
M. sinensis Tea-111	1.71	2x
M. sinensis Tea-112	1.74	2x
M. sinensis Tea-113	1.78	2x
M. sinensis Tea-114	1.73	2x
M. sinensis Tea-115	1.76	2x
M. sinensis Tea-13	1.74	2x
M. sinensis Tea-14	1.73	2x
M. sinensis Tea-30	1.70	2x
M. sinensis Tea-40	1.77	2x
M. sinensis Tea-76	1.74	2x
M. sinensis Tea-77	1.74	2x
M. sinensis Tea-78	1.78	2x
M. sinensis Tea-79	1.76	2x
M. sinensis Tea-80	1.77	2x

Table 2.4.6 (*continued*)

ID	Ratio	Ploidy
M. sinensis Tea-86	1.74	2x
M. sinensis Tea-88	1.73	2x
M. sinensis Tea-95	1.71	2x
M. sinensis Tea-96	1.76	2x
M. sinensis Tea-97	1.76	2x
M. sinensis Tea-98	1.76	2x
M. sinensis Tea-99	1.76	2x
M. sinensis 'zebrinus' Tea-3	1.70	2x
Miscanthus sp. Tea-130	1.73	2x
Miscanthus sp. Tea-16	1.82	2x
Miscanthus sp. Tea-18	1.71	2x
Miscanthus sp. Tea-22	1.76	2x
Miscanthus sp. Tea-24	1.76	2x
Miscanthus sp. Tea-25	1.75	2x
Miscanthus sp. Tea-26	1.72	2x
Miscanthus sp. Tea-27	1.72	2x
Miscanthus sp. Tea-29	1.73	2x
Miscanthus sp. Tea-38	1.73	2x
Miscanthus sp. Tea-41	1.75	2x
Miscanthus sp. Tea-42	1.70	2x
Miscanthus sp. Tea-43	1.71	2x
Miscanthus sp. Tea-45	1.72	2x
Miscanthus sp. Tea-46	1.79	2x
Miscanthus sp. Tea-47	1.74	2x
Miscanthus sp. Tea-73	1.76	2x
Miscanthus sp.Tea-37	1.73	2x
Miscanthus sp.Tea-39	1.75	2x
Miscanthus sp.Tea-54	1.74	2x
M. sacchariflorus x M. sinensis Tea-116	2.29	3x
M. sacchariflorus x M. sinensis Tea-117	2.31	3x
M. sacchariflorus x M. sinensis Tea-118	2.22	3x
M. sacchariflorus x M. sinensis Tea-119	2.27	3x
M. sacchariflorus x M. sinensis Tea-120	2.26	3x
M. sacchariflorus x M. sinensis Tea-121	2.22	3x
M. sacchariflorus x M. sinensis Tea-122	2.25	3x
M. sacchariflorus x M. sinensis Tea-123	2.23	3x
M. sacchariflorus x M. sinensis Tea-124	2.23	3x
M. sacchariflorus x M. sinensis Tea-125	2.31	3x
M. sacchariflorus x M. sinensis Tea-126	2.30	3x
M. sacchariflorus x M. sinensis Tea-127	2.22	3x
M. sinensis Tea-62	2.31	3x
M. xgiganteus Tea-17	2.22	3x
M. xgiganteus Tea-20	2.27	3x
M. xgiganteus Tea-31	2.24	3x
M. xgiganteus Tea-4	2.27	3x
M. xgiganteus Tea-5	2.28	3x
M. xgiganteus Tea-64	2.31	3x
M. xgiganteus Tea-65	2.27	3x
M. xgiganteus Tea-66	2.30	3x

Table 2.4.6 (*continued*)

ID	Ratio	Ploidy
M. xgiganteus Tea-74	2.31	3x
M. xgiganteus Tea-81	2.29	3x
M. xgiganteus Tea-82	2.27	3x
M. xgiganteus Tea-83	2.34	3x
M. xgiganteus Tea-93	2.32	3x
M. xgiganteus Tea-94	2.29	3x
Miscanthus sp. Tea-1	2.27	3x
Miscanthus sp. Tea-10	2.25	3x
Miscanthus sp. Tea-131	2.31	3x
Miscanthus sp. Tea-15	2.23	3x
Miscanthus sp. Tea-21	2.29	3x
Miscanthus sp. Tea-28	2.25	3x
Miscanthus sp. Tea-32	2.24	3x
Miscanthus sp. Tea-34	2.29	3x
Miscanthus sp. Tea-48	2.29	3x
Miscanthus sp. Tea-49	2.32	3x
Miscanthus sp. Tea-50	2.34	3x
Miscanthus sp. Tea-51	2.30	3x
Miscanthus sp. Tea-52	2.32	3x
Miscanthus sp. Tea-53	2.30	3x
Miscanthus sp. Tea-6	2.24	3x
Miscanthus sp. Tea-68	2.32	3x
Miscanthus sp. Tea-69	2.27	3x
Miscanthus sp. Tea-7	2.25	3x
Miscanthus sp. Tea-70	2.30	3x
Miscanthus sp. Tea-71	2.31	3x
Miscanthus sp. Tea-72	2.29	3x
Miscanthus sp.Tea-55	2.31	3x
Miscanthus sp. Tea-11	2.34	3x ?
M. sacchariflorus Tea-128	2.79	4x
M. sacchariflorus Tea-84	2.78	4x
M. sinensis 'goliath' Tea-57	2.54	4x
M. sinensis goliath-like Tea-19	2.63	4x
M. sinensis goliath-like Tea-85	2.55	4x
M. sinensis 'goliath'Tea-56	2.55	4x
M. sinensis 'zebrinus' Tea-33	2.51	4x
Miscanthus sp. Tea-23	2.55	4x
Miscanthus sp. Tea-8	2.78	4x
Miscanthus sp. Tea-9	2.68	4x
Miscanthus sp. Tea-90	3.19	4x
Miscanthus sp. Tea-91	3.19	4x
Miscanthus sp. Tea-92	3.04	4x
M. sacchariflorus x M. sinensis Tea-75	1.59	Aneuploid
M. sacchariflorus x M. sinensis Tea-87	1.57	Aneuploid
M. sinensis 'strictus' Tea-59	1.84	not reliable
M. sinensis 'strictus' Tea-60	1.90	not reliable

All the plants labelled as *M. ×giganteus* were found to be triploid, together with the hybrids between *M. saccharifloris* and *M. sinensis,* with the exception of Tea-75 and Tea-87 that showed a lower ratio than the diploid standard and were estimated as aneuploid genotypes. The accessions of *M. sacchariflorus* were estimated as tetraploid plants, as well as a group of *M. sinensis* 'Goliath' and the *M. sinensis* 'Zebrinus' Tea-33. It should be noticed that the ratio for the tetraploid *M. sinensis* is lower than the one of the tetraploid standard. With the only exception of accession Tea-62, all the remaining *M. sinensis* were found to be diploid. As for the *Miscanthus* sp. genotypes, they are almost equally divided between di- and triploids, with a few tetraploid, too.

2.4.3 DNA sequencing

The sequencing of the internal transcribed spacer nrDNA reveals the presence of base substitutions in the sequence of some accessions. Among these single nucleotide polymorphisms (SNPs), four revealed an interesting peculiarity: plants classified as *M. sacchariflorus* and *M. sinensis* clearly differed for the nucleotides in these positions, whereas the accessions belonging to *M. ×giganteus* showed double peaks of comparable intensity at these sites for both nucleotides present in *M. saccharifloris* and *M. sinensis* sequences (Figure 2.4.8).
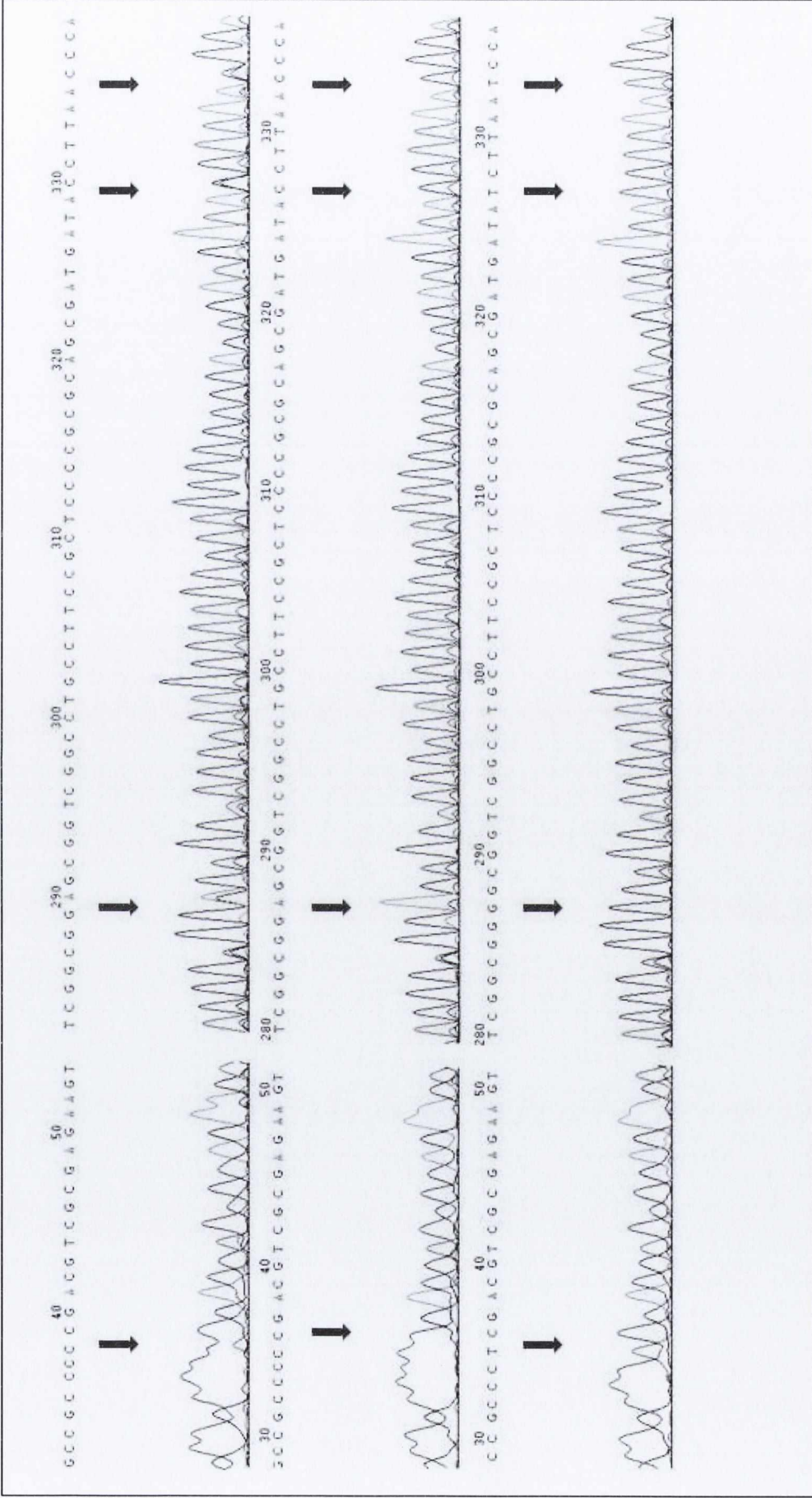
Figure 2.4.8 Internal transcribed spacer sequences of four polymorphic sites (shown by arrows) in *M.* ×*giganteus*, *M. saccharifloris* and *M. sinensis*. Where *M. sacchariflorus* and *M. sinensis* differs in their sequence, *M.* ×*giganteus* shows sequence heterogeneity.

51

A summary of the four nucleotides found in all the genotypes sequenced is shown in Table 2.4.7.

Table 2.4.7 Nucleotides in four polymorphic positions of the ITS-1 for 76 genotypes of the Oak Park collection.

ID	Nucleotide			
	nt 38	nt 290	nt 330	nt 336
M. sinensis goliath-like Tea-19	T	G	T	T
M. sinensis goliath-like Tea-85	T	G	T	T
M. sinensis 'goliath'Tea-56	T	G	T	T
M. sinensis 'sirene' Tea-58	T	G	T	T
M. sinensis 'zebrinus' Tea-2	T	G	T	T
M. sinensis Tea-14	T	G	T	T
M. sinensis Tea-30	T	G	T	T
M. sinensis Tea-40	T	G	T	T
M. sinensis Tea-77	T	G	T	T
M. sinensis Tea-79	T	G	T	T
M. sinensis Tea-80	T	G	T	T
M. sinensis Tea-86	T	G	T	T
M. sinensis Tea-88	T	G	T	T
M. sinensis Tea-95	T	G	T	T
M. sinensis Tea-96	T	G	T	T
M. sinensis Tea-97	T	G	T	T
M. sinensis Tea-98	T	G	T	T
M. sinensis Tea-99	T	G	T	T
M. sinensis Tea-100	T	G	T	T
M. sinensis Tea-101	T	G	T	T
M. sinensis Tea-102	T	G	T	T
M. sinensis Tea-103	T	G	T	T
M. sinensis Tea-104	T	G	T	T
M. sinensis Tea-105	T	G	T	T
M. sinensis Tea-106	T	G	T	T
M. sinensis Tea-107	T	G	T	T
M. sinensis Tea-108	T	G	T	T
M. sinensis Tea-109	T	G	T	T
M. sinensis Tea-110	T	G	T	T
M. sinensis Tea-111	T	G	T	T
M. sinensis Tea-112	T	G	T	T
M. sinensis Tea-113	T	G	T	T
M. sinensis Tea-114	T	G	T	T
M. sinensis Tea-115	T	G	T	T

Table 2.4.7 (*continued*)

ID	Nucleotide			
	nt 38	nt 290	nt 330	nt 336
M. sacchariflorus x M. sinensis Tea-75	T/C	G/A	T	T
M. sacchariflorus x M. sinensis Tea-87	C	G/A	T	T
M. sacchariflorus x M. sinensis Tea-117	T/C	G/A	T/C	T/C
M. sacchariflorus x M. sinensis Tea-118	T/C	G/A	T/C	T/C
M. sacchariflorus x M. sinensis Tea-119	T/C	G/A	T/C	T/C
M. sacchariflorus x M. sinensis Tea-120	T/C	G/A	T/C	T/C
M. sacchariflorus x M. sinensis Tea-121	T/C	G/A	T/C	T/C
M. sacchariflorus x M. sinensis Tea-122	T/C	G/A	T/C	T/C
M. sacchariflorus x M. sinensis Tea-123	T/C	G/A	T/C	T/C
M. sacchariflorus x M. sinensis Tea-124	T/C	G/A	T/C	T/C
M. sacchariflorus x M. sinensis Tea-126	T/C	G/A	T/C	T/C
M. xgiganteus Tea-4	T/C	G/A	T/C	T/C
M. xgiganteus Tea-5	T/C	G/A	T/C	T/C
M. xgiganteus Tea-64	T/C	G/A	T/C	T/C
M. xgiganteus Tea-66	T/C	G/A	T/C	T/C
M. xgiganteus Tea-81	T/C	G/A	T/C	T/C
M. xgiganteus Tea-82	T/C	G/A	T/C	T/C
M. xgiganteus Tea-83	T/C	G/A	T/C	T/C
M. xgiganteus Tea-93	T/C	G/A	T/C	T/C
M. sacchariflorus Tea-84	C	A	C	C
Miscanthus sp. Tea-1	T/C	G/A	T/C	T/C
Miscanthus sp. Tea-9	T/C	G/A	T/C	T/C
Miscanthus sp. Tea-10	T/C	G/A	T/C	T/C
Miscanthus sp. Tea-11	T/C	G/A	T/C	T/C
Miscanthus sp. Tea-18	T	G	T	T
Miscanthus sp. Tea-21	T/C	G/A	T/C	T/C
Miscanthus sp. Tea-22	T	G	T	T
Miscanthus sp. Tea-25	T	G	T	T
Miscanthus sp. Tea-24	T	G	T	T
Miscanthus sp. Tea-26	T	G	T	T
Miscanthus sp. Tea-28	T/C	G/A	T/C	T/C
Miscanthus sp. Tea-32	T/C	G/A	T/C	T/C
Miscanthus sp. Tea-34	T/C	G/A	T/C	T/C
Miscanthus sp. Tea-38	T	G	T	T
Miscanthus sp. Tea-42	T	G	T	T
Miscanthus sp. Tea-43	T	G	T	T
Miscanthus sp. Tea-46	T	G	T	T
Miscanthus sp. Tea-53	T/C	G/A	T/C	T/C
Miscanthus sp.Tea-54	T	G	T	T
Miscanthus sp. Tea-89	T/C	G/A	T/C	T/C
Miscanthus sp. Tea-130	T	G	T	T
Miscanthus sp. Tea-131	T/C	G/A	T/C	T/C

2.5 Discussion

2.5.1 Morphological diversity

Different traits were scored to evaluate the morphological variation in a collection of *Miscanthus*. As a reference, the same characters were measured in 79 herbarium specimens belonging to the species *M. condensatus, M. floridulus, M. nepalensis, M. nudipes. M. oligostachyus, M. sacchariflorus, M. sinensis, M. sorghum,* and *M. tinctorius*. The values scored were consistent with the description of the species given by Ohwi et al. (1965) and Chen and Renvoize (2006).

In particular, *M. sinensis* is characterized by a culm height between 48 and 116 cm and thickness of 2-8 mm, with leaves 27-75 cm long and 4-23 mm wide. The panicle is composed of a long inflorescence axis bearing 8-50 racemes 8-18 cm long. The spikelets are 5 mm long on average, with hairs of similar length and with an awn up to 7 mm long. In contrast, *M. sacchariflorus* has a culm up to 163 cm high and 2-5 mm wide, bearing leaves 24-57 cm long and 3-16 mm wide. The panicle is composed by a short inflorescence axis and 8-50 racemes long between 9-18 cm. The spikelets are 4 mm long with hairs 9 mm long and no awn.

Miscanthus condensatus and *M. floridulus* could not be clearly separated from *M. sinensis* based on quantitative traits only, whereas *M. oligostachyus* with its characteristic low raceme number in the inflorescence, and awned spikelets, was resolved from the other species of the *Miscanthus s.s.* group in the set of specimens analysed. Nevertheless when single traits were compared, they showed a different distribution among species.

The same traits were measured in the Oak Park collection established in Carlow. The statistical analyses revealed that only a few of them are normally distributed in the collection. This could be explained by the presence of different species in the field each with different distributions for the traits in question, as highlighted by the herbarium specimens, which give rise to a concealed multimodal distribution of the data.

Since most of the individuals in Oak Park presumably belongs to *M. sacchariflorus, M. sinensis* and *M. ×giganteus,* a principal component analysis (PCA) was performed including

54

the data for the *M. sacchariflorus* and *M. sinensis* specimens. Most of the variation in the data was due to the presence/absence of inflorescences, as half of the plants did not flower in the second growing season from the settling of the *Miscanthus* field, leading to two separate groups of individuals according to the flowering. The PCA could not resolve the specimens either, and no grouping was possible even for the plants that flowered.

Among the plants that did flower, the majority showed a *sinensis*-like inflorescence, with few exceptions, where the spikelets were awnless, as in *M. sacchariflorus*. The *sacchariflorus*-like inflorescence were found in some of the *M. sacchariflorus*×*M. sinensis* hybrids, and in one of the unidentified genotypes (*Miscanthus* sp. Tea-41) in one of the two replicates.

The morphological data alone were not conclusive in the aim to classify the unidentified individuals, mostly because of the unavailability of inflorescence for the whole collection.

2.5.2 Ploidy and molecular variation

The estimated ploidy levels for the Oak Park collection are consistent with the studies on the cytogenetics of *Miscanthus* by Adati and Shiotani (1962) and Linde-Laursen (1993). The plants classified as *M. sacchariflorus* were all found to tetraploid. Furthermore, all the individuals belonging to *M. ×giganteus* showed a triploid status, together with the new *M. sacchariflorus*×*M. sinensis* hybrids.

As for the known *M. sinensis,* the genotypes were mostly diploid, with a few exceptions. In *M. sinensis* Tea-62 the ratio of the fluorescence emitted by the nuclei is comparable to the value of the *M. ×giganteus* used as triploid standard. A different situation was found for other non-diploid *M. sinensis,* in particular four *M. sinensis* 'Goliath' and the *M. sinensis* 'Zebrinus' Tea-33: in this case the ratio measured by the flow cytometer was in between the values of the triploid and tetraploid standards.

According to Clifton-Brown and Lewandowski (2002), the genotype GOFAL7 (i.e. *M. sinensis* 'Goliath-like' Tea-85) is a triploid *sinensis* hybrid, as shown by a chromosome count of 2n = 57. The higher value in DNA content in this individual compared to the one of the triploid *M. ×giganteus* used as reference could be explained by the different composition in

55

haploid sets in the two genotypes: the 'Goliath-like' hybrid is an autotriploid with three *M. sinensis* haploid sets, whereas *M.* ×*giganteus* is an allotriploid that is supposed to have two genomes from *M. sinensis* and one from *M. sacchariflorus*, which has a lower amount of DNA per haploid genome (Rayburn et al. 2008).

It could be postulated that all the other *M. sinensis* with similar fluorescence ratio to Tea-85 are triploid. The triploid nature of *M. sinensis* 'Zebrinus' Tea-33 and the absence of the white stripes on the leaves typical of the 'Zebrinus' variety, suggest that this plant could have been misclassified.

As for the *Miscanthus* sp. genotypes, the information about the ploidy is not sufficient to support a tentative classification, but it could be a useful tool in addition to supplementary data, such as the sequencing of the ITS region.

Genotypes whose sequence for the ITS-1, the internal transcribed spacer between the genes for the nrRNA 18S and 5.8S, was obtained, showed a preference in base composition according to species: *M. sacchariflorus* and *M. sinensis* differ for the nucleotides present in four positions, whereas *M.* ×*giganteus* sequence is ambiguous for the presence of both nucleotides found in the two putative parents in each of the four positions. This seems to confirm the hypothesis that *M.* ×*giganteus* is an interspecific hybrid between *M. sacchariflorus* and *M. sinensis* (Linde-Laursen 1993) in which concerted evolution has not homogenized the sequences in the rDNA clusters yet, as expected for a sterile hybrid where unequal crossing-over is not possible (Hodkinson et al. 2002c).

When compared to the ploidy of the genotypes, the ITS in diploid genotypes showed a preference for a *sinensis*-like sequence, both in individuals known to be *M. sinensis* and in diploid *Miscanthus* sp., suggesting these individual could be *M. sinensis* too. Two out of three triploid *M. sinensis* 'Goliath' exhibit a *sinensis*-like sequence, while the genotype Tea-56 had a mixture of the two different sequences, as observed in the other three groups of triploids: *M.* ×*giganteus, M. sacchariflorus*×*M. sinensis* hybrids and 3x-*Miscanthus* sp.

The tetraploid *M. sacchariflorus* Tea-84 is the only genotype with a sacchariflorus-like sequence.

These data, together with the information from the morphological characterization, suggest that the Oak Park collection is composed by diploid plants mostly belonging to *M. sinensis*, three tetraploid *M. sacchariflorus*, 18 triploids belonging to the variety *M. sinensis* 'Goliath' and to the group of *M. sacchariflorus×M. sinensis* hybrids, while the remaining 3x plants are mostly *M. ×giganteus*.

2.6 Conclusions

The morphology of a collection of *Miscanthus* has been evaluated through the measurement of key qualitative and quantitative traits and then compared with the morphology of herbarium specimens of relevant *Miscanthus* species. In addition, the ploidy of the plants has been estimated. The results showed great morphological diversity among individuals and different levels of ploidy, with the presence of tetraploid *M. sacchariflorus,* triploid *M. ×giganteus,* and both diploid and triploid *M. sinensis.* A classification of unidentified genotypes has been attempted with the support of DNA sequencing, which proved to be a useful tool to discriminate between species, thanks to single nucleotide polymorphisms (SNPs) species-specific identified in the ITS-1 region of the nrDNA.

Chapter 3

Characterisation of genetic diversity and structure in a collection of Miscanthus and related species using newly developed chloroplast DNA microsatellite markers

3.1 Introduction

3.1.1 The chloroplast genome

For a long time the chloroplasts were regarded as organs that differentiated in an ancient cell from colourless protoplasm (Wilson, 1902). In 1885, Shimper demonstrated that although chlorophyll arises *de novo,* the plastids (chloroplasts) are already in the plant as leucoplasts and never appear *de novo.* This contrasted with the then used definition of an organ as a discrete part of an organism arising from primordia in the germplasm.

Following this evidence, Mereschkowsky (1905) speculated that chloroplasts (chromatophore) were foreign organisms (i.e. a cyanobacteria prokaryotes) that invaded the cell and entered into a symbiotic existence with it. To support what is today known as the 'endosymbiotic theory', Mereschkowsky reported five observations: 1) chloroplasts never arise *de novo*, but always through division of pre-existing plastids; 2) chloroplasts are highly independent of the nucleus; 3) chloroplasts possess a complete analogy with zoochlorellae; 4) Cyanophyceae are organisms that can be regarded as free-living chloroplasts; 5) Cyanophytes actually live as symbionts in cell protoplast.

The presence of DNA inside plastids was first demonstrated by Ris and Plaut (1962). They found, in the chloroplasts of Feulgen-stained Chlamydomonas cells, one or more small bodies with an intensity of reaction similar to the one observed in the nucleus. These bodies disappeared after treatment of the cells with ribonuclease. The endosymbiotic origin of plastids is now well documented and primary and secondary endosymbiotic events are believed to have resulted in the green and red algal symbionts and glaucophytes (Primoplantae) and several other Eukaryotic plant lineages (in the Chromalveolates, Excavates, and Rhizaria supergroups) (Palmer et al. 2004).

The first chloroplast genome (cpDNA) sequences became publicly available in 1986 (Ohyama et al. 1986; Tanaka et al. 1986) and the number has increased exponentially in the last decades. So far, over 250 plastid genomes have been published (http://www.ncbi.nlm.nih.gov/), providing general information about the structure of the chloroplast genome. Mainly the molecules are circular and they contain much less DNA than their contemporary prokaryotic relatives. This discrepancy in size is due to DNA transfer between chloroplast, mitochondria and nucleus, leading to the acquisition of genes important for the plastid biochemistry by the nuclear genome (Kleine et al. 2009). In *Arabidopsis* at least 2000 nuclear genes are of cynobacterial origin (Martin et al. 2002). In angiosperms, the size of the cpDNA ranges between 115 and 165 kb and consists of two large inverted repeats (IRs) dividing the single-copy part of genome into a small and a large region (SSC and LSC respectively) (Diekmann et al. 2009). The gene content and order is highly conserved, with genes coding for ribosomal and transfer RNA, ribosomal proteins, RNA polymerase subunits, and most of the polypeptides of Photosystem I, Photosystem II, the cytochrome b_6f and the ATP synthase (Green 2011). A typical plastid DNA genome is shown in Figure 3.1.1. This is from (Diekmann et al. 2009) for *Lolium perenne* (perennial ryegrass).
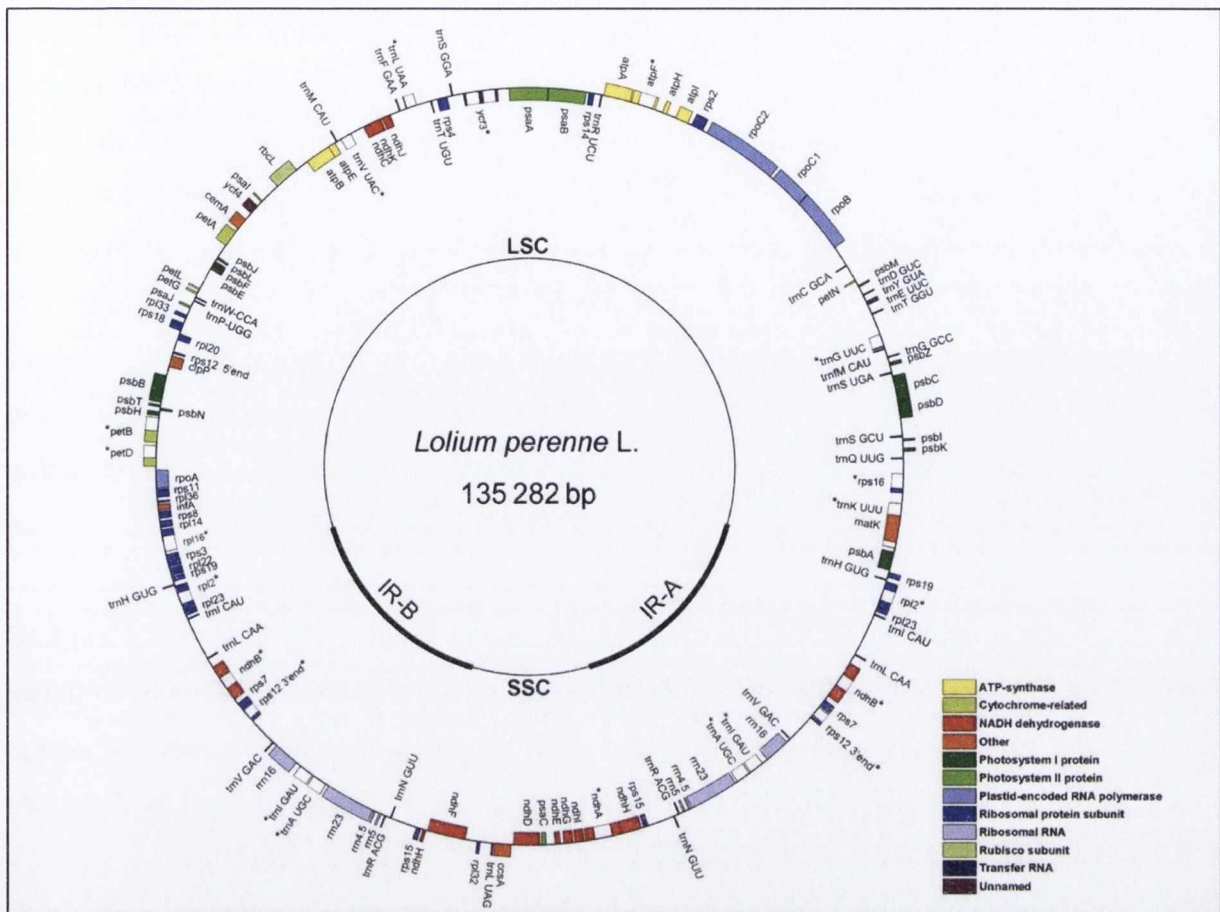
Figure 3.1.1 Circular structure of the chloroplast genome of *Lolium perenne*. Genes written on the outside are transcribed clockwise, genes on the inside counter-clockwise, annotated genes are colour coded according to their function, genes containing introns are highlighted with an asterisk; LSC, large single copy region; SSC, small single copy region; IR, inverted repeat From Diekmann et al. (2009) with permission.

The cpDNA is present in high copy number in the cell and it is uniparentally inherited. These features, together with the general homoplasmy and lack of recombination, have made the chloroplast genome an eligible tool for phylogenetic analyses in plants (Provan 2001) and the genetic resource characterization of cytoplasmic DNA diversity in crop plants (Flannery et al. 2006).

3.1.2 Chloroplast molecular markers

Different approaches have been used for phylogenetic and genetic resource characterization of germplasm with cpDNA (Olmstead and Palmer 1994; Flannery et al. 2006; McGrath et al.

2007). One of the earliest methods used was the analysis of restriction fragment variation. Despite the advantage of being simple to set up and the presence of many independent restriction sites along the genome, this methodology shows a lower limit due to the high level of conservation in the cpDNA among closely related species, and an upper limit where restriction site homology becomes difficult to be assessed (Olmstead and Palmer 1994).

More recently chloroplast genome sequences have proved useful in phylogenetic analysis including grasses (GPWG 2001; Hodkinson et al. 2002a). Several genes and intergenic spacer have been used, including genes the *rbcL, ndhF, matK, atpB, rpl16* and the non-coding regions *trnL* intron and *trnL-trnF* intergenic region (Ravi et al. 2007).

The discovery of mononucleotide repeats in the chloroplast genome of all the partial and total cpDNAs sequenced so far widened the possibility to large scale screening of polymorphism associated with the chloroplast genome. The need for DNA sequence in order to design specific primers to amplify microsatellite containing regions is counterbalanced by the cross-amplification in related species (Provan 2001).

Despite its conserved gene order and a lack of recombination, chloroplast DNA (cpDNA) shows length polymorphism associated with mononucleotide and less commonly polynucleotide repeats (Provan 2001; Diekmann et al. 2009). Non-coding intron and intergenic spacers are particularly variable and contain microsatellite and non-microsatellite polymorphisms even between closely related individuals and taxa in a range of plant groups (Provan 2001; Flannery et al. 2006; McGrath et al. 2007).

Chloroplast simple sequence repeat (cpSSR; also known as chloroplast microsatellite) markers have been useful to assess genetic variation in plants (Provan 2001). SNPs and indels are also common (Kelchner 2000) and provide useful markers. cpSSRs have been proven useful in gene flow studies to estimate seed and pollen contribution (McCauley 1995) and in phylogeographic analyses (McGrath et al. 2007).

Two chloroplast loci, the *trnL* intron and the *trnL-F* intergenic spacer, have previously been used to reconstruct the phylogenetic relationships in *Miscanthus* (Hodkinson et al. 2002a) and demonstrated considerable variation in the *Miscanthus* plastid genome. There is a shortage of reliable plastid genome markers available for plant genetic resource activity. The availability of plastid markers for this genus would facilitate the selection of parental lines with distinct plastid genomes. It will also help understand the complex polyploid and hybrid origins of some of its taxa.

3.2 Aims

The aims of the chapter were to develop new chloroplast SSR markers for the genus *Miscanthus* and to determine genetic diversity in a collection of *M.* ×*giganteus*, *M. sacchariflorus* and *M. sinensis* established in Teagasc, Oak Park, Carlow.

In detail, the objectives were:

- To design and optimise new primer pairs to amplify regions containing microsatellites;
- To determine the informativeness of the newly developed SSRs by testing them on several species of the genus *Miscanthus* and on representative species of related genera;
- To assess the genetic variation in the *Miscanthus* collection in Teagasc;
- To clarify the taxonomic status of unknown accessions in the collection.

3.3 Materials and methods

3.3.1 Plant material and DNA isolation

Fresh leaf material from a collection of 128 individuals of the genus *Miscanthus* was used for this study. Rhizomes of 33 *M. sinensis* were provided by Svalöf Weibull, Sweden; 80 individuals of *M. ×giganteus*, *M. sacchariflorus* and *M. sinensis*, including different ornamental varieties, were collected from TCD Botanic Gardens, Dublin, Ireland; 15 additional genotypes of the three species were made available by the University of Hohenheim, Germany (Clifton-Brown and Lewandowski 2002). Specimens for other *Miscanthus*, *Saccharum* and related grasses (subfamily Panicoideae) were collected from the living collections at the Royal Botanic Gardens, Kew, Surrey, UK and ADAS, Arthur Rickwood Research Station, Cambridge, UK. Details on the number of individuals per species analysed are shown in Table 2. Fresh leaves were frozen in liquid nitrogen and ground manually to a fine powder. Total genomic DNA was extracted following a modified cetyltrimethyl ammonium bromide (CTAB) method (Doyle and Doyle 1987) or by following the CTAB protocol in Hodkinson et al. (2002a).

3.3.2 Primer design, amplification and SSR selection

Regions of cpSSRs were identified in the complete chloroplast genome sequence of *Saccharum officinarum* (GenBank Accession No. AP006714.1) using 'find_microsat_Win32' (Salamin, personal communication) and primers were developed to amplify 30 loci, each, with one exception, including a A/T repetition ranging from 8bp to 15bp (Table 3.3.1).

65

Table 3.3.1 List of primer pairs developed for cpSSR amplification in *Miscanthus*. Locus position, SSR motifs and expected length of the PCR products refer to the *Saccharum officinarum* chloroplast genome used as template for primers design. In grey PCR products that were sequenced, darker for primer pairs used for genotyping.

Name	Locus position	cpSSR	Forward Primer	Revers Primer	Length	Ampl.
Sac-1	IGSpsbE-petL	$(A)_9t(A)_8$	tgaccaaacagcaagcatgg	agaaaagcattccacgaccc	230	Y
Sac-2	Int-trnK	$(A)_{15}$	acctttccctgcattaggca	agaccacgactgatcctcaa	260	Y
Sac-3	IGSrps16-trnQ	$(T)_{14}$	accaaagaaaagaaggacgc	tctacggatagggactctat	269	Y
Sac-4	Int-rps12;infA	$(T)_{10}+(T)_{10}$	ggatgatcgtgtcattctctagg	ggtaatggccgcacctatag	214	Y
Sac-5	IGSytrnT-trnE	$(T)_{14}$	gggcttttcacttagtggtag	gcaattttaactctgtgcttcg	207	Y
Sac-6	IGSndhK-ndhC	$(T)_{14}$	cttgtttggtcgagtaacgg	cgacgtattgggtgtatccg	197	Y
Sac-7	Rpl16-Int-rps12	$(T)_{13}+(T)_8$	cagtttattaacccggctgctc	ggattcccatgtgtatgg	173	Y
Sac-8	IGSpsbM-petN	$(A)_{12}$	aaggtgcgagatgcttcaatcga	gaaagggcgattatagtaact	226	N
Sac-9	IGSrpl33-rps18	$(T)_{12}$	cttcttctggttctggatca	acgcctacgaaaaggttgtttgg	218	Y
Sac-10	IGSrps11-rpl36	$(T)_{12}+(T)_9$	taaaacgggcattcctacgc	ccgaagcataaacaaagacag	285	Y
Sac-11	IGSpsbK-psbI	$(A)_{11}$	ccaatcgtggatgttatgc	taggccaggggtagaaag	259	Y
Sac-12	IGSrps16-trnQ	$(T)_{11}$	cttcttcgactcgaataaca	gcgaaaacgatctcgatctgtg	238	Y
Sac-13	IGStrnG-trnfM	$(T)_{11}$	gtgcttcgaggacgcaaatt	cagagcggagtagagcagtt	293	Y
Sac-14	IGSatpI-atpH	$(T)_{11}$	aggtgaatccatggaaggtcatc	cctacaactctaggtgtat	158	Y
Sac-15	trnR;IGStrnR-	$(T)_{11}+(T)_{11}$	cgtgtgtagtagaagaggaatcg	attatttctcttgttccgag	238	Y
Sac-16	IGSpsaI-ycf4	$(T)_{11}$	gattgtctagaaacgacggg	gccttcagtagctagtactcct	282	Y
Sac-17	IGSmatK-rps16	$(A)_{10}$	tacaccggacgctcctgtcaaa	ttgccctcttgcatgtact	232	Y
Sac-18	IGSorf147-trnT	$(A)_{10}$	gttcttattgcccctttggc	aacacgatctcgtacgactcac	132	Y
Sac-19	rpoc2	$(A)_{10}$	aaaagctgccctacgaggtc	gtaaacggggtctctgatgt	261	Y
Sac-20	intpetB;intrps12	$(A)_{10}+(A)_{10}$	gaaagggcctgttatctcta	gagttctcttttgggcaaac	222	Y
Sac-21	IGSndh5-rpl32	$(A)_{10}$	cagaatgggtttagttactg	caattacgaaacaacagagctc	175	
Sac-22	psbC	$(G)_{10}$	gttaggtctaggtgcttttc	acaatccatccttctcccc	174	Y
Sac-23	orf43	$(T)_{10}$	tagatcgcgcaaggcaagaa	gctctctattgcatgggtgg	203	Y
Sac-24	IGStrnC-rpoB	$(T)_{10}$	tggatttccagtcgcaggct	ccgatttaagagtcgttcac	133	Y
Sac-25	IGSaptI-aptH	$(T)_{10}$	cccgatagagcttagaagttgg	agcagtaccttgaccaactc	182	Y
Sac-26	intatpF	$(T)_{10}$	gagtgtgtgcgagttgtcta	accaatgaatcgcgaaatgc	175	Y
Sac-27	IGSatpB-rbcL	$(T)_{10}$	gacgcgaagtagtaggattg	gcaacgaaatcaagtgcgag	191	Y
Sac-28	IGSatpB-rbcL	$(T)_{10}$	gaacgtacacagggtgtaca	cagggtctactcgatatgga	161	Y
Sac-29	IGSpetG-trnW	$(T)_{10}$	agcgggattattcgtgactg	cgatgtcgtaggttcaaatcc	214	Y
Sac-30	IGSpsaJ-rpl33	$(T)_{10}$	gaattcttcgtgacatgacg	ctttgcccttggccatgaac	357	Y

DNA from twelve samples was amplified with an initial denaturation of 5 min at 95°C followed by 35 cycles each with a denaturation of 1 min at 95°C, 1 min at a primer-specific annealing temperature (see results section 3.4), and an extension of 1 min at 72°C, followed by a final extension at 72°C for 10 min. The reaction mixture (10 μl) contained 10× reaction buffer (New England Biolabs) containing 2 mM $MgSO_4$, 0.125μM dNTPs, 0.25 μM of each primer, 0.5 U of *Taq* DNA polymerase (New England Biolabs) and 40 ng template DNA.

The PCR products were loaded on 3% MetaPhor® Agarose (Lonza, Rockland, ME, USA) gels. For primers which produced weak amplification, PCR conditions were optimized using a gradient PCR, with temperatures ranging from 48°C to 60°C, and the amplification test was repeated using the optimal annealing temperature. Twelve primer pairs producing the best amplification (highlighted in grey in Table 3.3.1) were selected and used to amplify a set of 24 genotypes. The PCR products were subsequently sequenced by a commercial sequencing company (AGOWA GmbH, Germany) to confirm length polymorphism in the microsatellite motifs. The sequences were aligned using ClustalW (Thompson et al. 1994) as implemented in MEGA version 3.1 (Kumar et al. 2004). Six markers did not show polymorphism in the microsatellite sequence and where therefore discarded. Further analyses were carried out on the remaining six cpSSRs (in dark grey in Table 3.3.1).

3.3.3 Genotyping

Forward primers were fluorescently labelled so that they could be used for automated genotyping. A polyA treatment at 65°C was applied for 30 min to the PCR products. The PCR products were then sized using the LIZ500 internal sizing standard on an ABI 3130xl automated DNA sequencer with GENEMAPPER ™ V4.0 software (Applied Biosystems).

3.3.4 Data analyses

Genetic distance

Allele number and size range were calculated for each locus. Due to the haploid nature of the chloroplast markers, it was necessary to transform the data matrix into a binary matrix scoring 1 for presence of alleles and 0 for absence. Genetic similarity (GS) indices were calculated using the Jaccard's coefficient for all possible pairwise comparisons. The Jaccard's coefficients disregards the conjoint absence of alleles in the pairwise comparison, reducing the risk of over-estimating similarity. Jaccard's coefficients were calculated using the software FreeTree (Pavlícek et al. 1999) and used to cluster genotypes according to similarity. The UPGMA (unweighted pair group method using arithmetic means) clustering

tree building approach was used, with internal support assessed using 1000 bootstrap replicates. UPGMA tree was visualized using FIGTREE 1.2.1 (Rambaut 2007).

Structure

The software STRUCTURE 2.3.3 (Pritchard et al. 2000) was used to infer the genetic structure of the collection. A series of simulations were run with the number of clusters K ranging from 1 to 8, with three independent runs for each K value. Each run consisted of a burn-in period of 10,000 steps and 100,000 MCMC (Monte Carlo Markov Chain) replicates, assuming an admixture model and uncorrelated allele frequencies. No prior information about the structure of the population was defined. The most likely value of K was chosen following (Evanno et al. 2005) and used to run a simulation with a burn-in period of 10,000 steps followed by 100,000 MCMC replicates.

AMOVA

An analysis of molecular variance (AMOVA; Excoffier et al. (1992)) was carried out with GenAlEx 6 (Peakall and Smouse 2006) starting from the raw data to estimate the components of genetic variation between and within groups as observed in the output of the cluster analysis performed with STRUCTURE 2.3.3. A value of 999 permutations was used to test for statistical significance.

PCA

A principal coordinates analysis (PCA) was carried out with GenAlEx 6 (Peakall and Smouse 2006). Starting from raw data, a Nei genetic distance matrix was calculated as

$$GD = -\ln \frac{J_{xy}}{\sqrt{J_x J_y}} \, ,$$

where:

$$J_{xy} = \sum_{i=1}^{k} p_{ix} p_{iy}$$

$$J_x = \sum_{i=1}^{k} p_{ix}^2$$

68

$$J_y = \sum_{i=1}^{k} p_{iy}^2$$

with p_{ix} and p_{iy} as the frequencies of the i_{th} allele in populations x and y. For multiple loci, J_{xy}, J_x and J_y are calculated by summing over all loci and alleles and dividing by the number of loci. These average values are then used to calculate the genetic distance matrix used to perform the PCA with the algorithm described in Orloci, 1978 (cited in GenAlEx 6) as implemented in GenAlEx 6.

3.4 Results

3.4.1 Analysis of the cpSSR sequences

Sequencing of the *Miscanthus* accessions revealed the nature of the detected length variation. They contained mononucleotide repeats ranging from 10bp to 21bp in all six loci (Table 3.3.1), with species-specific length polymorphisms due to A/T indels of 1-4bp. The region amplified with the Sac-2 primer pair included two mononucleotide repeats which were both polymorphic and were separated by 68bp (Figure 3.4.1a). The combination of the length variation of the two microsatellites results in an allele number of four at this locus in the sequenced genotypes. All other loci sequenced included a single microsatellite with two to three alleles (Figure 3.4.1b).

In addition to microsatellite length polymorphism, two species-specific SNPs were detected in the sequence amplified with the marker Sac-10 (Figure 3.4.1b). The hybrid *M. ×giganteus* and the species *M. sacchariflorus* share the same sequence at this locus, whereas *M. sinensis* may be differentiated.

Figure 3.4.1 Variation in chloroplast simple sequence repeat (cpSSR) motifs and flanking regions of cpSSR markers Sac-2 (a) and Sac-10 (b) in a range of *Miscanthus* accessions. Grey boxes: allele groups, lined boxes: simple sequence polymorphisms (SNPs). In (a) the alignment shows two polymorphic regions. Within the flanking markers, the sequence is interrupted by a dotted line to indicate the presence of a non-displayed and non-polymorphic sequence between the two microsatellite motifs.

70

The number of alleles detected by genotyping of 165 *Miscanthus* accessions ranged from four within locus Sac-26 to ten within locus Sac-10. Additional alleles were detected in related grass species (Table 3.4.1).

Table 3.4.1 List of cpSSRs with locus name, GenBank accession number, chloroplast genome region in *Saccharum*, primer sequences, dye for the 5'-labelling of the F primer, annealing temperature, SSR motif, size range of the PCR product and allele number in all species analysed and in the genus *Miscanthus* only (in brackets).

cpSSR marker	GenBank Accession numbers	Chloroplast genome region	Dye	T_a (°C)	Repeat motif	Size range (bp) (Miscanthus)	Allele no. (Miscanthus)
Sac-2	FN64379 to 82	trnK intron	Pet	58	$(T)_{11}/(A)_{21}$	236-257 (247-256)	11 (9)
Sac-3	FN646383 to 86	rps16-trnQ IGS	Vic	56	$(T)_{16}$	262-272 (262-270)	10 (9)
Sac-10	FN646387 to 90	rps11-rpl36 IGS	Fam	52	$(T)_{12}$	269-294 (271-287)	16 (10)
Sac-13	FN646391 to 94	trnG-trnfM IGS	Ned	62	$(T)_{15}$	283-291 (283-291)	9 (8)
Sac-17	FN646395 to 98	matK-rps16 IGS	Fam	60	$(A)_{12}$	217-238 (227-235)	9 (7)
Sac-26	FN646399 to 02	atpF intron	Fam	58	$(T)_{10}$	169-178 (172-177)	7 (4)

The six markers were tested on 73 individuals of *M. sinensis,* nine *M. sacchariflorus* individuals and 15 *M. ×giganteus* individuals. 14 out of 15 *M. ×giganteus* analysed shared the same haplotype. The other two species showed a high level of polymorphism for all markers, but with a preference for the frequency of certain alleles (Table 3.4.2).

Table 3.4.2 Allele sizes of the six cpSSR loci grouped by species. Numbers of individuals per species are shown (N), the most frequent allele within a species is in bold.

Species	N	Marker Sac-2	Sac-3	Sac-10	Sac-13	Sac-17	Sac-26
Miscanthus species							
Miscanthus capense	2	251, 253	266, 267	273, 275	287	227, 229	176
M. condensatus	1	255	269	275	288	230	176
M. ecklonii	2	253, 255	266	275	287	229, 230	176
M. erectum	1	253	266	273	287	229	176
M. fusca	2	249, 251	265	287	291	228	175
M. ×giganteus	15	252	266	276, **277**	290	227	175
M. junceum	3	249, 251, 252	264, 268	273, 284	289	229	176, 177
M. nepalensis	1	250	262	284	288	229	175
M. nudipes	4	247, 248, 250	266	283, 286	285, 287	228, 229	176
M. oligostachyus	1	252	269	284	288	229	177
M. sacchariflorus	9	**252**, 255	264, 265, **266**, 269	271, 276, **277**	288, 289, 290	**227**, 230	**175**, 176
M. sinensis	73	251, 252, 253, **255**, 256	266, 267, 268, **269**, 270	274, **275**, 276, 277	283, 287, **288**, 290	227, 229, **230**, 231, 234, 235	172, 175, **176**, 177
M. sorghum	1	251	265	273	286	229	176
M. sp.	42	251, 252, 255, 256	265, 266, 267, 268, 269, 270	275, 276, 277	287, 288, 289, 290	227, 228, 229, 230, 234	175, 176
M. teretifolium	1	251	263	273	286	227	175
M. transmorrisonensis	1	255	269	275	288	230	176
M. tinctorius	2	252, 256	265, 269	275, 276	288	227, 230	176
M. violaceum	4	251	264, 267	273, 275	288, 290	227, 228	175
Related species (Panicoideae)							
Cymbopogon citratus	1	257	n.a.	279	n.a.	230	175
Eulalia quadrinervis	1	247	n.a.	283	283	234	174
Eulalia tripsicata	1	251	263	269	284	238	174
Eulalia villosa	1	249	262	278	n.a.	234	174
Pennisetum sp.	1	236	265	294	290	229	178
Saccharum contortus	1	250	264	283	n.a.	229	176

| Species | N | Marker | | | | | |
		Sac-2	Sac-3	Sac-10	Sac-13	Sac-17	Sac-26
Saccharum officinarum	2	252, 253	267	275, 284	291	228	176
Saccharum porphyrocoma	1	n.a.	266	277	288	n.a.	175
Saccharum spontaneum	1	252	263	289	291	229	176
Sorghum halpense	1	251	262	273	283	217	175
Spodiopogon rhizophorus	1	248	n.a.	278	n.a.	n.a.	169
Spodiopogon sibricus	1	248	n.a.	279	285	230	169
Zea diploperennis	1	250	272	285	290	227	175

3.4.3 Cluster analysis with STRUCTURE

The genetic structure of the population was detected using a model-based clustering method (Pritchard et al. 2000) as implemented in STRUCTURE 2.3.3. A series of three independent runs for each value of K (i.e. the number of populations in the collection) was run. For each run, the estimated log probability of data Pr(X|K) for each value of K is given (Table 3.4.3). The mean likelihood, indicated as L(K) afterwards, over the three runs for each K was first plotted against K (Figure 3.4.2). L(K) could be seen increasing dramatically until K=3, after which it slowly decrease. In order to harvest the true value for K, three additional steps were introduced, following (Evanno et al. 2005). In the second step, the mean difference between successive values of likelihood of K L'(K) = L(K)-L(K-1) was calculated and in the third step, the absolute value of the difference between successive values of L'(K), |L"(K) | = |L'(K+1)-L'(K) | (Table 3.4.3). Finally the value ΔK is estimated as the mean of |L"(K) | averaged over the three runs divided by the standard deviation of L(K), ΔK = m|L"(K) |/ s[L(K)].

Table 3.4.3 Evanno parameters calculated over three repetitions for each K value ranging from 1 to 8.

K	Reps	Mean LnP(K)	Stdev LnP(K)	Ln'(K)	|Ln''(K)|	Delta K
1	3	-1501.4	0.1	NA	NA	NA
2	3	-1103.5333	0.4163	397.86667	209.43333	503.0426
3	3	-915.1	1.3748	188.43333	181	131.65813
4	3	-907.6667	1.5044	7.433333	19.933333	13.249689
5	3	-920.1667	0.3215	-12.5	2.833333	8.814089
6	3	-935.5	0.781	-15.333333	4.333333	5.548265
7	3	-955.1667	2.9143	-19.666667	5.7	1.955851
8	3	-969.1333	2.2723	-13.966667	NA	NA
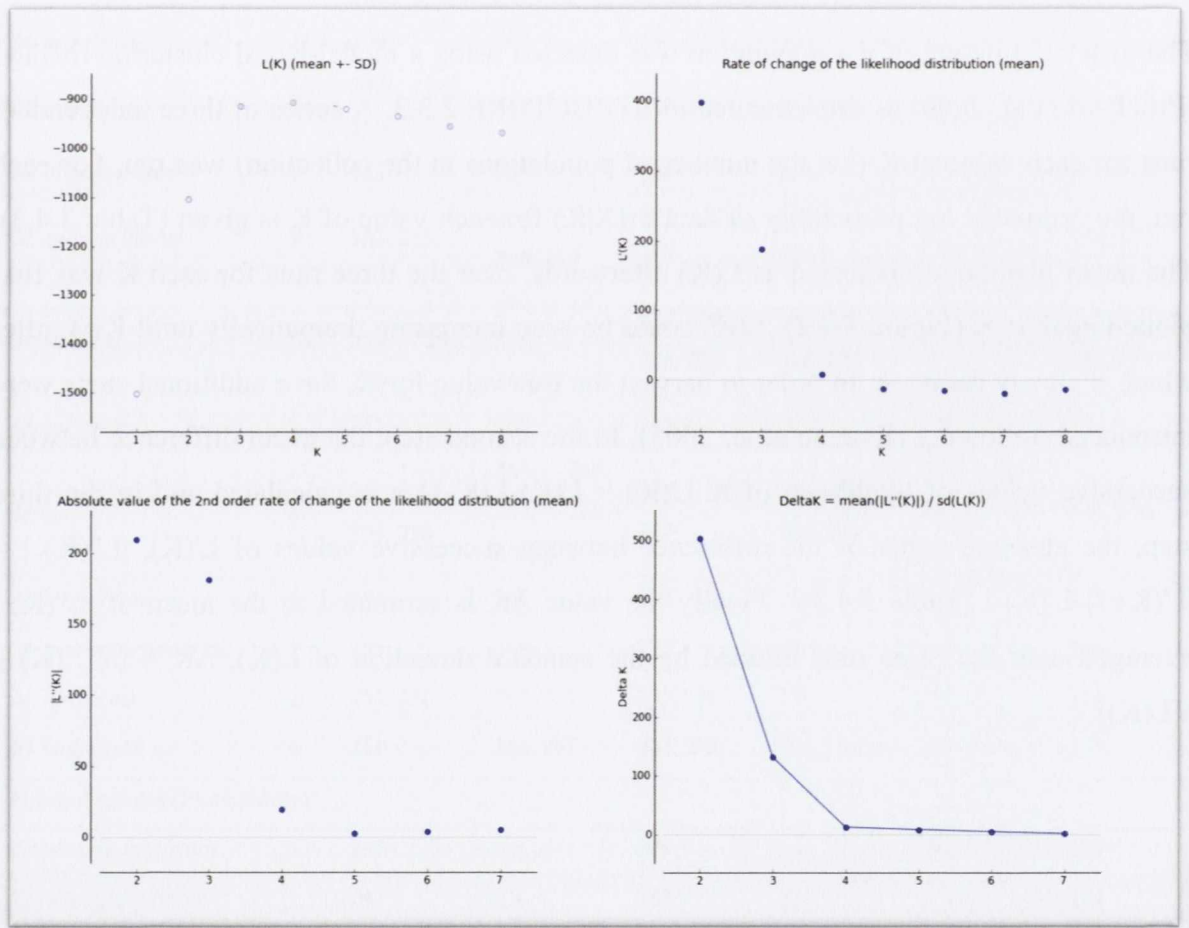
Figure 3.4.2 Graphical representation of the Evanno parameters for the estimation of the K value.

Plotting ΔK against the values of K, the highest value of ΔK represents the true value of K for the data, in our case at K=2. A simulation for this value of K was then run in order to

74

assign the individuals to each of the two clusters. The result is summarized in Table 3.4.4 and in Figure 3.4.3.
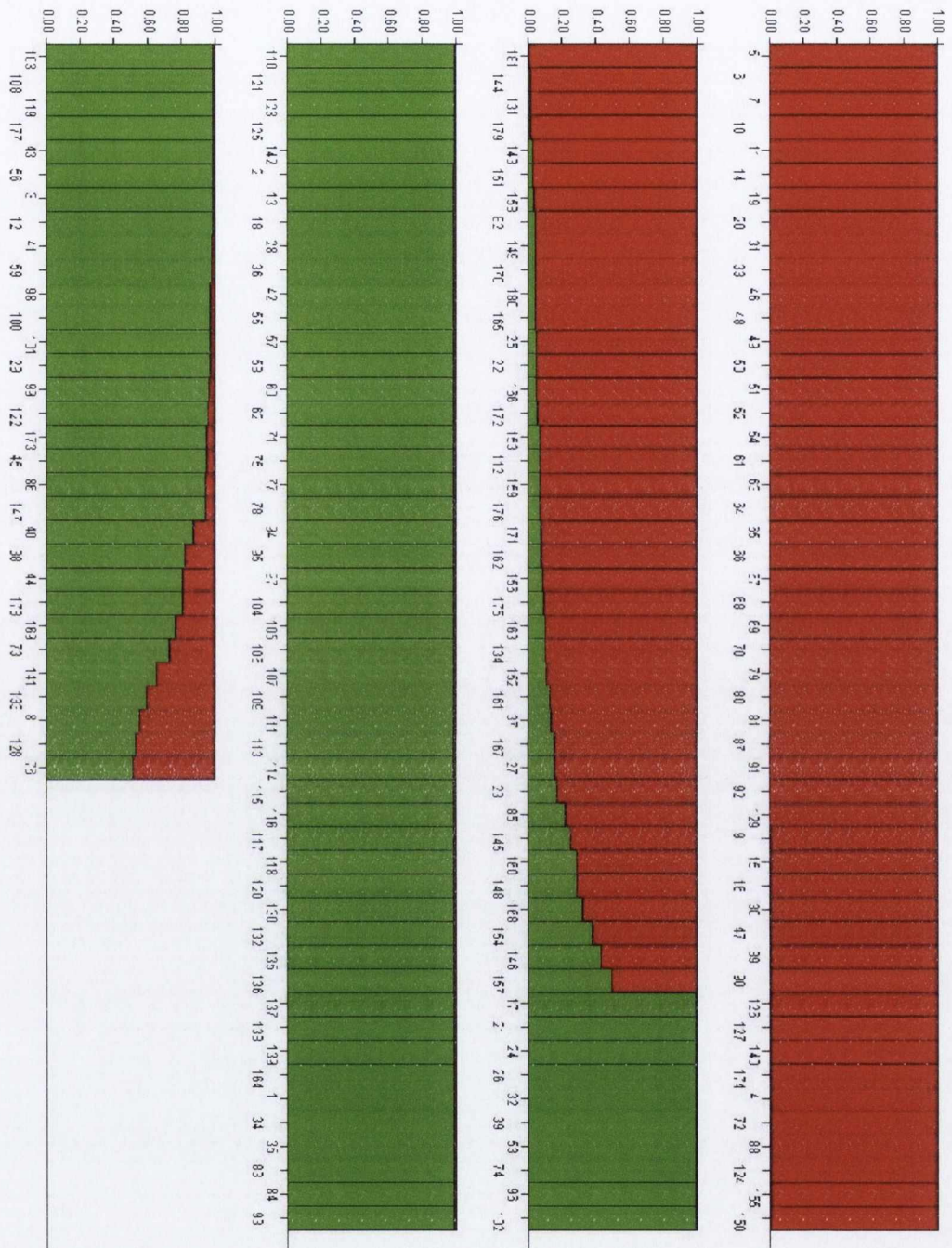
Figure 3.4.3 Structure barplot assigning each accession to cluster I (red) and cluster II (green).

Table 3.4.4 Accessions assigned to each cluster according to STRUCTURE analysis. Percentage of missing data is indicated. Shading indicates the assigned cluster.

ID	% Missing	Inferred cluster	
		I	II
Miscanthus sp. Tea-1	0	0.009	0.991
M. sinensis 'zebrinus' Tea-2	0	0.008	0.992
M. sinensis 'zebrinus' Tea-3	-33	0.012	0.988
M. xgiganteus Tea-4	-16	0.991	0.009
M. xgiganteus Tea-5	0	0.993	0.007
Miscanthus sp. Tea-6	0	0.993	0.007
Miscanthus sp. Tea-7	0	0.993	0.007
Miscanthus sp. Tea-8	-33	0.449	0.551
Miscanthus sp. Tea-9	0	0.992	0.008
Miscanthus sp. Tea-10	0	0.993	0.007
Miscanthus sp. Tea-11	0	0.993	0.007
M. sinensis Tea-13	0	0.013	0.987
M. sinensis Tea-14	0	0.008	0.992
Miscanthus sp. Tea-15	0	0.993	0.007
Miscanthus sp. Tea-16	0	0.992	0.008
M. xgiganteus Tea-17	0	0.992	0.008
Miscanthus sp. Tea-18	0	0.007	0.993
M. sinensis goliath-like Tea-19	0	0.008	0.992
M. xgiganteus Tea-20	0	0.993	0.007
Miscanthus sp. Tea-21	0	0.993	0.007
Miscanthus sp. Tea-22	0	0.007	0.993
Miscanthus sp. Tea-23	-16	0.948	0.052
Miscanthus sp. Tea-24	-16	0.827	0.173
Miscanthus sp. Tea-25	0	0.007	0.993
Miscanthus sp. Tea-26	0	0.949	0.051
Miscanthus sp. Tea-27	0	0.007	0.993
Miscanthus sp. Tea-28	-16	0.838	0.162
Miscanthus sp. Tea-29	0	0.008	0.992
M. sinensis Tea-30	-16	0.027	0.973
M. xgiganteus Tea-31	0	0.992	0.008
Miscanthus sp. Tea-32	0	0.993	0.007
M. sinensis 'zebrinus' Tea-33	0	0.007	0.993
Miscanthus sp. Tea-34	0	0.993	0.007
M. sinensis 'gross fontane' Tea-35	0	0.009	0.991
M. sinensis 'gross fontane' Tea-36	0	0.009	0.991
Miscanthus sp.Tea-37	0	0.008	0.992
Miscanthus sp. Tea-38	0	0.858	0.142
Miscanthus sp.Tea-39	-16	0.176	0.824
M. sinensis Tea-40	0	0.007	0.993
Miscanthus sp. Tea-41	0	0.124	0.876
Miscanthus sp. Tea-42	0	0.015	0.985
Miscanthus sp. Tea-43	0	0.008	0.992
M. condensatus Tea-44	-33	0.01	0.99

Table 3.4.4 (*continued*)

ID	% Missing	Inferred cluster	
		I	II
Miscanthus sp. Tea-45	0	0.187	0.813
Miscanthus sp. Tea-46	0	0.047	0.953
Miscanthus sp. Tea-47	0	0.993	0.007
Miscanthus sp. Tea-48	0	0.992	0.008
Miscanthus sp. Tea-49	0	0.993	0.007
Miscanthus sp. Tea-50	0	0.993	0.007
Miscanthus sp. Tea-51	0	0.993	0.007
Miscanthus sp. Tea-52	0	0.993	0.007
Miscanthus sp. Tea-53	0	0.993	0.007
Miscanthus sp.Tea-54	0	0.007	0.993
Miscanthus sp.Tea-55	0	0.993	0.007
M. sinensis 'goliath'Tea-56	0	0.008	0.992
M. sinensis 'goliath' Tea-57	0	0.011	0.989
M. sinensis 'sirene' Tea-58	0	0.008	0.992
M. sinensis 'strictus' Tea-59	0	0.008	0.992
M. sinensis 'strictus' Tea-60	0	0.019	0.981
M. sinensis 'malaparteus' Tea-61	0	0.008	0.992
M. sinensis Tea-62	0	0.993	0.007
M. sinensis 'sirene' Tea-63	0	0.008	0.992
M. xgiganteus Tea-64	0	0.993	0.007
M. xgiganteus Tea-65	0	0.993	0.007
M. xgiganteus Tea-66	0	0.993	0.007
Miscanthus sp. Tea-68	0	0.993	0.007
Miscanthus sp. Tea-69	0	0.993	0.007
Miscanthus sp. Tea-70	0	0.993	0.007
Miscanthus sp. Tea-71	0	0.993	0.007
Miscanthus sp. Tea-72	0	0.993	0.007
Miscanthus sp. Tea-73	0	0.008	0.992
M. xgiganteus Tea-74	-16	0.991	0.009
M. sacchariflorus x M. sinensis Tea-75	0	0.27	0.73
M. sinensis Tea-76	0	0.007	0.993
M. sinensis Tea-77	0	0.008	0.992
M. sinensis Tea-78	-16	0.479	0.521
M. sinensis Tea-79	0	0.008	0.992
M. sinensis Tea-80	0	0.008	0.992
M. xgiganteus Tea-81	0	0.993	0.007
M. xgiganteus Tea-82	0	0.993	0.007
M. xgiganteus Tea-83	0	0.993	0.007
M. sacchariflorus Tea-84	-16	0.957	0.043
M. sinensis goliath-like Tea-85	-16	0.009	0.991
M. sinensis Tea-86	0	0.009	0.991
M. sacchariflorus x M. sinensis Tea-87	0	0.772	0.228
M. sinensis Tea-88	0	0.056	0.944

Table 3.4.4 (*continued*)

ID	% Missing	Inferred cluster	
		I	II
Miscanthus sp. Tea-89	0	0.993	0.007
Miscanthus sp. Tea-90	-16	0.991	0.009
Miscanthus sp. Tea-91	0	0.992	0.008
Miscanthus sp. Tea-92	0	0.992	0.008
M. xgiganteus Tea-93	0	0.993	0.007
M. xgiganteus Tea-94	0	0.993	0.007
M. sinensis Tea-95	-16	0.009	0.991
M. sinensis Tea-96	0	0.008	0.992
M. sinensis Tea-97	0	0.008	0.992
M. sinensis Tea-98	0	0.007	0.993
M. sinensis Tea-99	0	0.008	0.992
M. sinensis Tea-100	0	0.024	0.976
M. sinensis Tea-101	0	0.032	0.968
M. sinensis Tea-102	0	0.025	0.975
M. sinensis Tea-103	0	0.025	0.975
M. sinensis Tea-104	0	0.007	0.993
M. sinensis Tea-105	0	0.009	0.991
M. sinensis Tea-106	0	0.008	0.992
M. sinensis Tea-107	0	0.008	0.992
M. sinensis Tea-108	0	0.008	0.992
M. sinensis Tea-109	0	0.008	0.992
M. sinensis Tea-110	-16	0.009	0.991
M. sinensis Tea-111	0	0.008	0.992
M. sinensis Tea-112	0	0.007	0.993
M. sinensis Tea-113	0	0.008	0.992
M. sinensis Tea-114	0	0.924	0.076
M. sinensis Tea-115	0	0.008	0.992
M. sacchariflorus x M. sinensis Tea-116	0	0.008	0.992
M. sacchariflorus x M. sinensis Tea-117	0	0.008	0.992
M. sacchariflorus x M. sinensis Tea-118	0	0.008	0.992
M. sacchariflorus x M. sinensis Tea-119	0	0.008	0.992
M. sacchariflorus x M. sinensis Tea-120	0	0.008	0.992
M. sacchariflorus x M. sinensis Tea-121	-16	0.009	0.991
M. sacchariflorus x M. sinensis Tea-122	0	0.008	0.992
M. sacchariflorus x M. sinensis Tea-123	0	0.007	0.993
M. sacchariflorus x M. sinensis Tea-124	0	0.041	0.959
M. sacchariflorus x M. sinensis Tea-125	0	0.007	0.993
M. sacchariflorus x M. sinensis Tea-126	-16	0.991	0.009
M. sacchariflorus x M. sinensis Tea-127	0	0.007	0.993
M. sacchariflorus Tea-128	0	0.992	0.008
M. sacchariflorus Tea-129	0	0.992	0.008
Miscanthus sp. Tea-130	0	0.469	0.531
Miscanthus sp. Tea-131	0	0.993	0.007

Table 3.4.4 (*continued*)

| ID | % Missing | Inferred cluster | |
		I	II
M. sinensis 'variegatus' Kew 1	0	0.008	0.992
Sorghum halpense Kew 6	0	0.981	0.019
M. condensatus Kew 7	0	0.008	0.992
M. oligostachyus Kew 16	0	0.404	0.596
M. nepalensis Kew 25	0	0.896	0.104
M. sinensis 'goliath' Kew 27	0	0.008	0.992
M. sinensis 'gracillimus' Kew 28	0	0.008	0.992
M. sinensis 'roland' Kew 29	0	0.008	0.992
M. sinensis Kew 30	0	0.008	0.992
M. sinensis 'gross fontane' Kew 31	0	0.008	0.992
M. sacchariflorus Kew 61	0	0.992	0.008
M. sinensis 'yakushimanum' Kew 63	0	0.345	0.655
M. transmorrisonensis Kew 65	0	0.007	0.993
M. fusca Kew 82	-50	0.97	0.03
M. violaceum Kew 84	-33	0.983	0.017
M. violaceum Kew 85	0	0.75	0.25
M. ecklonii Kew 86	0	0.568	0.432
M. ecklonii Kew 87	-16	0.057	0.943
M. junceum Kew 88	-83	0.709	0.291
M. junceum Kew 89	0	0.957	0.043
M. fusca Kew 91	0	0.987	0.013
M. violaceum Kew 92	-50	0.97	0.03
M. violaceum Kew 93	-16	0.887	0.113
M. capense Kew 94	0	0.925	0.075
M. capense Kew 95	-16	0.619	0.381
M. teretifolium Kew 96	0	0.99	0.01
M. junceum Kew 97	-16	0.909	0.091
Saccharum officinarum Kew 104	0	0.503	0.497
M. sorghum Kew 105	0	0.959	0.041
M. erectum Kew 106	0	0.922	0.078
M. yunnanensis Kew 107	-33	0.712	0.288
M. nudipes Kew 109	-16	0.868	0.132
M. nudipes Kew 110	0	0.92	0.08
M. nudipes Kew 111	0	0.898	0.102
M. tinctorius Kew 112	0	0.008	0.992
Saccharum spontaneum Kew 117	0	0.954	0.046
Narenga porphyrocoma Kew 120	-33	0.944	0.056
Saccharum contortus Kew 121	-16	0.842	0.158
Spodipogon rhizophorus Kew 125	-50	0.672	0.328
Spodipogon sibiricus Kew 128	-16	0.232	0.768
Eulalia villosa Kew 132	-16	0.957	0.043
Eulalia quadrinervis Kew 134	-16	0.921	0.079
Eulalia tripsicata Kew 138	0	0.942	0.058

Table 3.4.4 (*continued*)

ID	% Missing	Inferred cluster	
		I	II
M. sinensis 'morning light' Kew 155	0	0.044	0.956
M. Sacchariflorus Kew 159	0	0.992	0.008
M. sacchariflorus Kew 160	0	0.905	0.095
M. tinctorius 'nana variegata' Kew 161	0	0.922	0.078
M. sinensis 'goliath' Kew 194	-16	0.009	0.991
Cymbopogon citratus	-33	0.192	0.808
Pennisetum sp.	0	0.98	0.02
Saccharum officinarum	0	0.956	0.044
Zea diploperennis	0	0.987	0.013

The cluster analysis shows a clear separation of the *Miscanthus* collection in two clusters. All the *M. ×giganteus* and *M. sacchariflorus* clones belong to Cluster I, together with most of the *Miscanthus sensu lato* species (except *M. tinctorius* Kew 112 and *M. ecklonii* Kew 87) and the non-*Miscanthus* species (except only *Cymbopogon citratus* and *Spodipogon sibiricus*). In contrast, the *Miscanthus sensu stricto* can be found in Cluster II, together with most of the *M. sinensis* (apart from Tea-62 and Tea-114). The *M. sacchariflorus* × *M. sinensis* hybrids fall in Cluster II, with the exception of Tea-87 and Tea-126.

3.4.4 UPGMA tree

The matrix of Jaccard coefficients was calculated for all the accessions based on the cpSSR markers. The UPGMA tree (Figure 3.4.4) shows two clearly separated clusters: Cluster I, highlighted in yellow, and Cluster II in blue. Cluster I groups together all the *M. ×giganteus* and *M. sacchariflorus* clones. With the exception of only two genotypes (Tea-62 and Tea-78), all the *Miscanthus* classified as *sinensis* fall in the second cluster, as well as the *M. sacchariflorus* × *M. sinensis* hybrids (except Tea-126) and the clones belonging to other *Miscanthus s.s.* species. The *Miscanthus* species of the *sensu lato* group are equally divided between the two clusters. These results are consistent with the clustering obtained with STRUCTURE, except for two accessions of *M. sinensis*, Tea-78 and Tea-114, and the hybrid Tea- 87, that were assigned to a different cluster in the UPGMA analysis.

The *Saccarhum* species included in the study appear to be more closely related to *M. sinensis* than to *M. ×giganteus* and *M. sacchariflorus,* in contrast with the results from STRUCTURE. In both clusters subgroups can be recognized of individuals sharing the same haplotype (i.e. where no variation was detected among accessions).

80
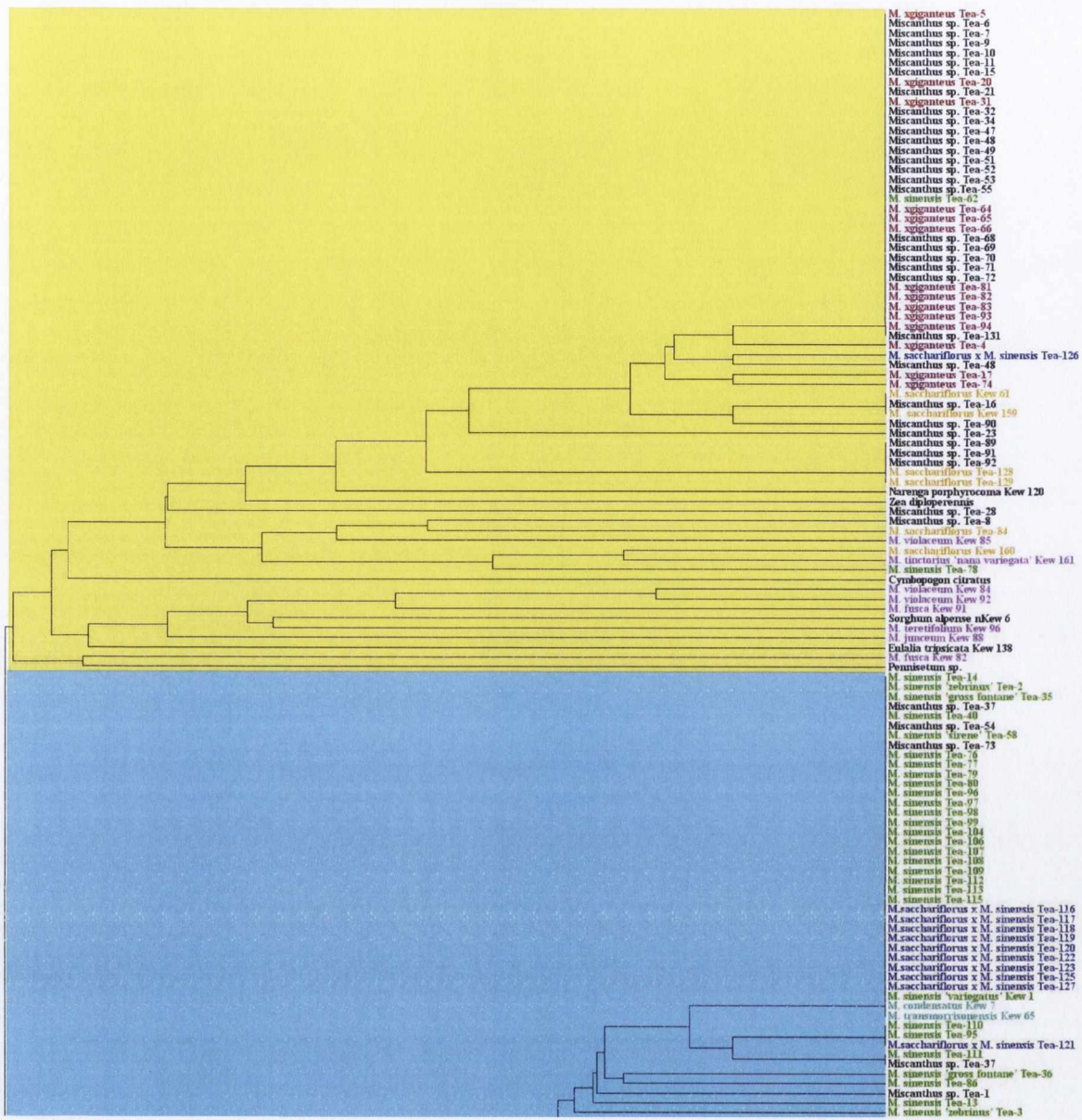
Figure 3.4.4 (*continued*)
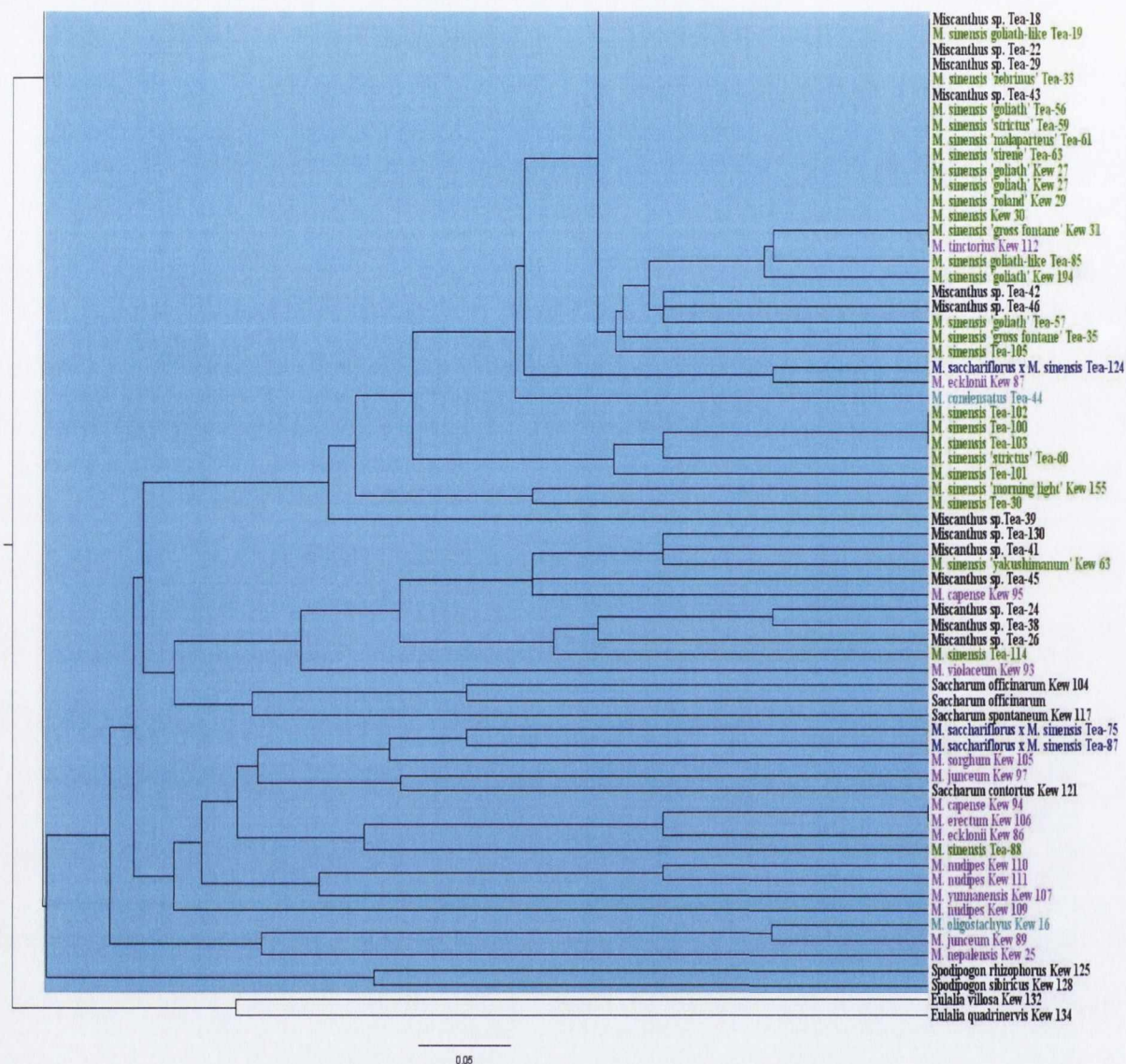
Figure 3.4.4 UPGMA tree showing inter-relationships of individuals using a combination of 6 cpSSR markers. Labels: Red = *M. ×giganteus;* Green = *M. sinensis;* Yellow = *M. sacchariflorus;* Blue = *M. sacchariflorus×M. sinensis* hybrids; Light blue = *Miscanthus s.s;* Violet = other *Miscanthus*; Black = non-*Miscanthus* species and unclassified *Miscanthus* accessions; Yellow box = Cluster I; Blue box = Cluster II.

3.4.5 AMOVA

An analysis of molecular variance was carried out to evaluate how the genetic variation is partitioned within and among populations. In this case, the two clusters obtained with

STRUCTURE 2.3.3 were used as supposed populations for the AMOVA. The output is summarised in Table 3.4.5 and Figure 3.4.5.

Table 3.4.5 Analysis of molecular variance (AMOVA) between clusters as shown by the cluster analysis with STRUCTURE 2.3.3. df= degrees of freedom; SS= sum of squares; MS= mean square; Est.Var.= Estimated variation; %= percentage of molecular variance.

Source	df	SS	MS	Est. Var.	%
Among Pops	1	46712.221	46712.221	302.561	2%
Within Pops	179	3460298.96	19331.279	19331.279	98%
Total	180	3507011.18		19633.84	100%
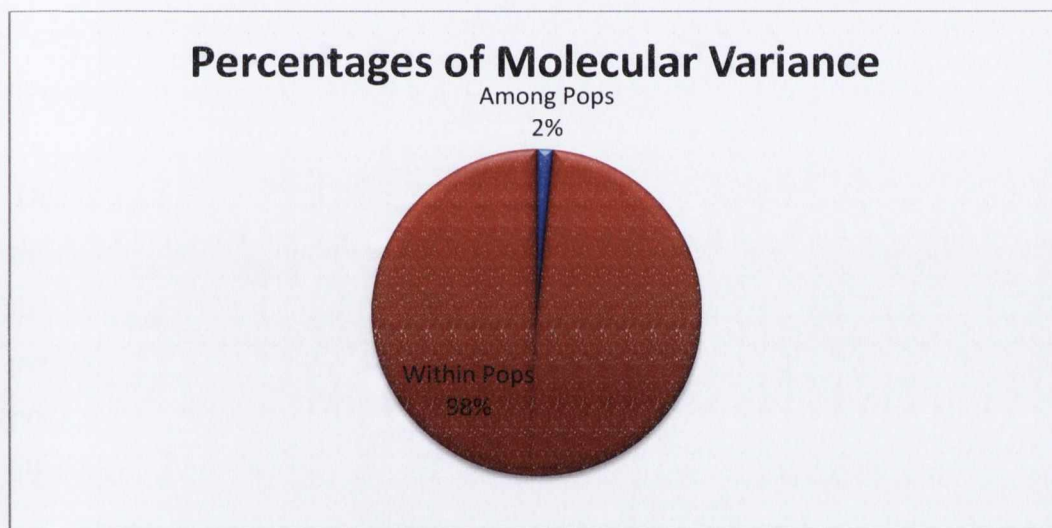
Figure 3.4.5 Graphical representation of the analysis of molecular variance (AMOVA) between using cluster I and II from STRUCTURE output as populations.

The AMOVA shows clearly that the genetic diversity within clusters accounts for the most of the diversity, with a percentage of 98%.

3.4.6 Principal coordinates analysis

In Table 3.4.6 are shown the percentages of variation explained by the first three axes of the PCA.

Table 3.4.6 Percentages of variation, for each axis and cumulative, explained by the first three axes.

Axis	1	2	3
%	47.21	24.4	14.57
Cum %	47.21	71.61	86.19

The first eigenvalue accounts for 47.21% of the total variation, the second 24.40% of the total variation (cumulative 71.61%) and the third 14.57% of the total variation (cumulative 86.19%). The eigenvectors were plotted in a two dimensional scatterplot (Figure 3.4.6).
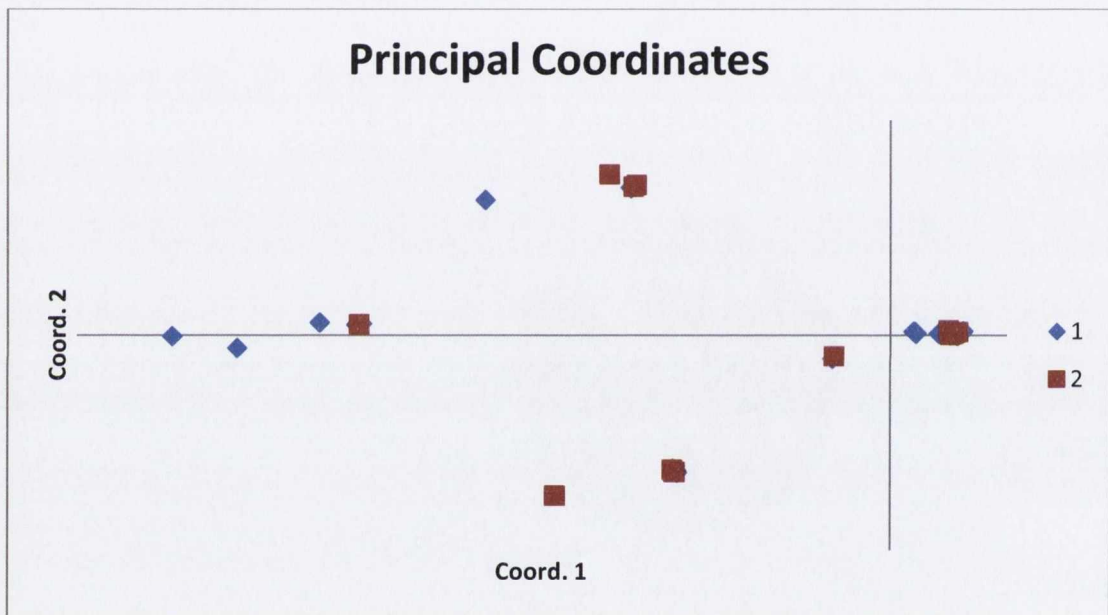
Figure 3.4.6 Principal coordinates analysis scatterplot for the cpSSRs data with the first coordinate as x-axis and the second as y-axis. Groups: ◊ Cluster I; □ Cluster II.

Four groups of accessions can be identified in the scatterplot: a core group of individuals (at the crossing of the axes in Figure 3.4.6), a smaller group at the opposite side of the plot along the first axis, and two additional small groups separated along the second axis.

The PCA does not show an obvious pattern of separation between the two clusters highlighted by previous analyses: all four groups include individuals from both clusters.

3.5 Discussion

Primer pairs used in this study were developed from non-coding regions, containing mononucleotide repeats, of the *Saccharum* chloroplast genome, a genus closely related to *Miscanthus* (Clayton and Renvoize 1986).

Thirty primer pairs were designed to target possible polymorphic regions in the chloroplast genome of *Miscanthus* and were tested on a small number of individuals belonging to *M. sinensis, M. sacchariflorus* and *M. ×giganteus* to check for transferability of primers from *Saccharum* to *Miscanthus*. With only one exception, all primer pairs amplified in *Miscanthus*. To confirm the presence of mononucleotide repeats, PCR products from twelve loci were sequenced on 24 accessions equally divided among *M. sinensis, M. sacchariflorus* and *M. ×giganteus* and sequences were aligned to check for length polymorphism in the simple sequence repeats (SSRs).

The alignment of sequences highlighted a species-specific polymorphism in six chloroplast microsatellite markers used (Figure 3.4.1), with *M. sacchariflorus* and *M. ×giganteus* sharing the same alleles for five out of six loci. This finding was confirmed by the genotyping performed on a large number of accessions belonging to the three species, where although polymorphism was found within species, it appeared to be a bias in the presence of certain alleles in *M. sinensis* compared to the alleles found in *M. sacchariflorus* and *M. ×giganteus*.

The six primer pairs amplified non-coding regions of the chloroplast genome, in detail four intergenic spacer (IGS) regions and two intronic regions of genes *trnK* and *atpF* respectively (Table 3.3.1). For each marker locus, size range and allele number was determined (Table 3.3.1) on 165 individuals belonging to 17 species of the genus *Miscanthus* and 13 related species. Between seven (Sac-26; *atpF* intron) and 16 alleles (Sac-10; IGS *rps11-rpl36*) were found per locus for all the grasses tested and between four (Sac-26) and ten (Sac-10) alleles for *Miscanthus* accessions.

The number of haplotypes ranges between 50 when only accessions that amplified in all six loci are taken into account, and 85 considering missing data as null alleles (Table E in

Appendix). Three haplotypes are the most frequent (shared between 36, 34 and 16 individuals respectively), whereas 76 are unique to only one accession.

It is clear that a combination of alleles results in a large number of haplotypes that can be used for comparative analyses. This compares to 511 found in *Lolium perenne* (McGrath et al. 2006) using twelve loci to genotype 1,575 individuals across 104 accessions. *Lolium* was found to be highly diverse in its plastid DNA variation. *Miscanthus* is less diverse but the variation detected is of use for genetic resource characterization.

The markers were shown to be transferable among *Miscanthus s.l.* species tested (*M. capense, M. condensatus, M. ecklonii, M. erectum, M. fusca, M. junceum, M. nepalensis, M. nudipes, M. oligostachyus, M. sorghum, M. teretifolium, M. transmorrisonensis, M. tinctorius* and *M. violaceum*) and also related genera (*Cymbopogon citrates, Eulalia quadrinervis, Eulalia tripsicata, Eulalia villosa, Pennisetum sp., Saccharum contortus, Saccharum officinarum, Saccharum porphyrocoma, Saccharum spontaneum, Sorghum halpense, Spodiopogon rhizophorus, Spodiopogon sibricus* and *Zea diploperennis*) (Table 3.4.2). Therefore the markers should be of value as 'universal' plastid DNA markers in grasses and especially in Panicoideae grasses the subfamily that *Miscanthus* belongs. Several other economically important crops are found in this group including *Saccharum* (sugarcane), *Zea* (maize) and *Sorghum*.

All the known *M. ×giganteus* share the same haplotype, with just an exception for locus Sac-10 in *M. ×giganteus* Tea-17 whereas more variation is present in the *M. sacchariflorus* accessions. This could be explained by the sterile nature of the allotriploid *M. ×giganteus* that has been mostly propagated vegetatively in Europe since it was first introduced from Japan in 1935 (Hodkinson et al. 2002). It is therefore encouraging that variation exists in the chloroplast genome of our genetic resource collection of *M. sacchariflorus*. It represents a novel source of genes for plant breeding purposes. It is not possible to determine the total number of *M. ×giganteus* genotypes from the cpDNA data presented here but the nuclear DNA variation (reported in chapter 4) shows that several genotypes exist and the material is not all clonal. However, they all share the same cpDNA halplotype which indicates that they are closely related and of limited cytoplasmic diversity in comparison to *Miscanthus* as a whole sampled in this study.

In the pioneering study by Adati and Shiotani (1962), it was theorized that many plants classified as *M. sacchariflorus* may be hybrids with a genome inherited by *M. sinensis* and one of unknown origin. Linde-Laursen (1993) demonstrate that so called *M. sinensis* 'Giganteus' are allotriploid with two genomes with high homology and one with lower homology. Hodkinson et al. (2002c) demonstrated with the use of ITS sequencing, that these allotriploids are indeed *M. ×giganteus* whose putative parents are *M. sinensis* and *M. sacchariflorus*. Moreover the sequencing of the plastid *trnL* intron and *trnL-F* intergenic spacer suggested that the maternally inherited cpDNA in *M. ×giganteus* originates from *M. sacchariflorus*.

Some artificial crosses of *M. sinensis* and *M. sacchariflorus* were included in our study. In several of these the hybrid has the plastid genome of *M. sinensis* showing that hybridisation is possible in both directions (with both species as maternal parent). There is no reason to believe that the formation of *M. ×giganteus* in the wild is unidirectional but our study suggests that this could be the case as all the *M. ×giganteus* accessions have *M. sacchariflorus* plastid DNA. This unidirectional hybridisation can be caused by several factors including nuclear cytoplasmic DNA incompatability (Anderson and Maan 1995) effects or by population factors. For example if *M. sinensis* was rare and *M. sacchariflorus* common (or if phenological differences created such a pattern), the vast number of seeds set would be from *M. sacchariflorus* ovule donors. However, a small number of *M. sinensis* plants can potentially father a large number of *M. ×giganteus* seed.

When additional *Miscanthus* and related species were introduced in the study, all six loci proved polymorphic both within *Miscanthus* and *Saccharum* and across all grass species analysed.

Among *Miscanthus* s.s. species, some authors have treated *M. condensatus* as a separate species and others have treated it as a subspecies of *M. sinensis*. The data presented here supports the latter hypothesis. Likewise, *M. transmorrisonensis* from Taiwan is clearly closely allied with *M. sinensis*. Both *M. condensatus* and *M. transmorrisonensis* shared the most common sinensis haplotype (coded as '1' in Table E) for cpSSRs. This is in contrast with the findings of Hodkinson et al. (2002b) on *M. transmorrisonensis* based on AFLP data, where this species showed significant divergence from *M. sinensis*.

As for the Japanese endemic species *M. oligostachyus* and *M. tinctorius*, the first shares the same alleles with most of *M. sinensis* in four loci and one with *M. sacchariflorus*. The Sac-10 allele in *M. oligostachyus* was not found in any other *Miscanthus* s.s. species, but only in some *Miscanthus* s.l. species (*M. junceum* and *M. nepalensis)* as well as in *Saccharum officinarum*. This is consistent with the classification of *M. oligostachyus* in the *Miscanthus* s.s. group in Hodkinson et al. (2002b) together with *M. sinensis* and *M. sacchariflorus*.

3.6 Conclusions

Six new plastid SSR markers were developed from the complete cpDNA sequence in *Saccharum officinarum* and tested on a collection of *Miscanthus* accessions belonging to 17 different species in the genus as well as on 13 species from related genera. All markers proved to be polymorphic within and between species, with a species-specific preference for certain alleles.

These are among the first cpSSR and SNP markers developed for *Miscanthus*. These new markers will be useful in breeding programs for *Miscanthus*, for testing maternal inheritance of the chloroplast genome and for species differentiation. The cross amplification of the markers in other species and genera of the subfamily Panicoideae was also proved.

All the *M. ×giganteus* accessions have *M. sacchariflorus* plastid DNA indicating that hybridisation might be unidirectional. However, more wild populations will need to be screened to confirm this phenomenon occurs in a general sense. Variation has been detected in the *M. sacchariflorus* germplasm collection and this will be of value to future breeding efforts that combine *M. sacchariflorus* and *M. sinensis* genomes.

Chapter 4

Characterisation of genetic diversity and population structure in a collection of *Miscanthus* and related species using newly developed nuclear DNA microsatellite markers

4.1 Introduction

4.1.1 Nuclear molecular markers

Molecular markers are useful tools to detect and analyse genetic variation in plants. An ideal molecular marker should be highly polymorphic and evenly distributed along the genome, easy to detect, inexpensive, with high reproducibility and no need for prior information about the genome of the organism to study (Agarwal et al. 2008). Several techniques to generate molecular markers have been developed and are now well established, such as restriction fragment length polymorphism, RFLP (Botstein et al. 1980), random amplified polymorphic DNA, RAPD (Williams et al. 1990), amplified fragment length polymorphism, AFLP (Vos et al., 1995) and microsatellite or simple sequence repeats, SSRs (Tautz and Renz 1984)

RFLP was among the first molecular markers developed to detect DNA polymorphism and works by hybridization of labelled probes with DNA previously digested with restriction endonucleases. Though highly informative, RFLP markers are not as widely used as PCR based methods because of the need for a high quantity of DNA and long reaction time. After the invention of PCR, most of the molecular markers were developed based on this technique, with two different approaches: (1) sequence non-specific markers and (2) sequence targeted markers (Agarwal et al. 2008). RAPD and AFLP belong to the first group, whereas SSR are sequence-specific markers.

RAPD uses short random oligonucleotides to amplify genomic DNA without prior knowledge of the genome needed to design primers. The polymorphism detected is due to rearrangements in the sequences at or between the primer binding sites. This technique is fast and produces a large number of markers, but the reproducibility of the results is deeply affected by the reaction conditions (Bardakci 2001).

AFLPs are generated by PCR amplification of selected fragments obtained through digestion of the genomic DNA using restriction enzymes. The amplification produces 50-100 bands per reaction, mostly uniquely positioned along the genome, thus allowing for these markers to be used for both physical and genetic mapping (Yin et al. 1999). Both RAPD and AFLP are dominant markers and are thus unable to distinguish between homo- and heterozygotes.

Microsatellites or SSRs, on the other hand, are codominant markers. A microsatellite is typically a short nucleotide sequence (1-5 bp) repeated in tandem (Tautz and Renz 1984). They are relatively abundant in all eukaryotic genomes. Polymorphism is due to variation in the number of repeats caused by slippage of the polymerase during replication that leads to a high number of alleles per locus. Each microsatellite can be detected through PCR using primers that recognize the flanking non-repetitive regions. The downside of this technique is the need for prior sequence information to design new primers. However, once the primers are available, its use is relatively inexpensive and straightforward. Depending on conservation of the flanking regions and the stability of the microsatellite, SSRs have proven to be transferable to species in the same genus or in related genera (Thomas and Scott 1993).

Microsatellites have found several applications in plants. SSR linkage maps are today available for a number of plant genomes (Röder et al. 1998; Temnykh et al. 2000) and interesting traits have already been tagged to help for marker-assisted selection (MAS) in plant breeding as a way of speeding up the selection of useful traits. Due to the high number of alleles per locus, SSRs have been shown to be more effective in discriminating among cultivars (Thomas and Scott 1993; Rongwen et al. 1995) compared to other molecular markers, and in assessing genetic variation in the genepool of crop plants, and thanks to their codominant nature, they also permit discrimination of parental contributions in hybrids (Powell et al. 1996).

4.1.2 Nuclear molecular markers in *Miscanthus*

An early attempt to characterise genetic diversity in *Miscanthus* and to clarify the taxonomy of the genus was conducted using AFLP fingerprinting on a collection of plants including clones of *M. ×giganteus*, *M. sacchariflorus* and *M. sinensis* sampled in botanic and market gardens in Europe (Greef et al. 1997). The taxonomic identity of some European *Miscanthus*

has been investigated further using AFLP in conjunction with additional molecular markers such as inter-simple sequence repeat, ISSR PCR (Hodkinson et al. 2002b) and DNA sequencing of nuclear (ITS) and chloroplast (*trnL-F*) regions (Hodkinson et al. 2002a) who also extended their analyses to other *Miscanthus* species. A preliminary genetic linkage map was built using RAPD markers (Atienza et al. 2002). RFLP and SSR markers from maize (Hernández et al. 2001) and more recently from *Brachypodium distachyon* (Zhao et al. 2011) have been successfully applied to *Miscanthus*. New nuclear SSR markers have been developed for *M. sinensis* and tested for cross-amplification on *M. floridulus* (Ho et al. 2011), *M. sacchariflorus* and *M. lutarioriparius* (Zhou et al. 2011). There is a need to develop more SSR markers for *Miscanthus* and to use these to characterise genetic diversity in a broad range of germplasm including the hybrid *M.* ×*giganteus*, and species outside the *Miscanthus s.s.* group and closely related genera.

4.2 Aims

The aims of the chapter were to develop new nuclear SSR markers for the genus *Miscanthus* and to determine genetic diversity in a collection of *Miscanthus* including *M. ×giganteus*, *M. sacchariflorus* and *M. sinensis* established in Teagasc, Oak Park, Carlow.

The detailed objectives were:

- To design and optimise new primer pairs to amplify regions containing microsatellites;
- To determine the informativeness of the newly developed SSRs by testing them on several species of the genus *Miscanthus* and on representative species of related genera;
- To assess the genetic variation in the *Miscanthus* collection in Teagasc;
- To clarify the taxonomic status of unknown accessions in the collection.

4.3 Materials and methods

4.3.1 Plant material and DNA isolation

Rhizomes of 33 *Miscanthus sinensis* were provided by Svalöf Weibull, Sweden; 80 individuals of *M. ×giganteus*, *M. sacchariflorus* and *M. sinensis*, including different ornamental varieties, were collected from TCD Botanic Gardens, Dublin, Ireland; 15 additional genotypes of the three species were made available by the University of Hohenheim, Germany (Clifton-Brown and Lewandowski 2002). Specimens for other *Miscanthus, Saccharum* and related grasses (subfamily Panicoideae) were collected from the living collections at the Royal Botanic Gardens, Kew, Surrey, UK and ADAS, Arthur Rickwood Research Station, Cambridge, UK. Fresh leaves were frozen in liquid nitrogen and ground manually to a fine powder. Total genomic DNA was extracted following the cetyltrimethyl ammonium bromide (CTAB) method (Doyle and Doyle 1987) or a modification of it (Hodkinson et al. 2002b).

4.3.2 Primer design

Total genomic DNA from the *Miscanthus sinensis* clone SW217 was isolated to build a nuclear microsatellite enriched library. A small amount of genomic DNA (>0.5 µg) was provided to *ATG Genetics* Inc., Canada. After digestion with multiple 4 cutter restriction enzymes, enrichment for SSRs containing fragments was obtained through biotinylated TC_n, TG_n and $GATA_n$ simple sequence motifs.

The selected fragments were cloned into the EcoRI site of the plasmid pUC19 and screened for positive clones using ^{32}P labelled TC_n, CA_n and $GATA_n$ simple sequence motifs. Two 96-well microtitre plates containing single positive bacterial colonies in 0.2 ml LB plus 15% glycerol, one selected for the presence of dinucleotide repeats and the second for tetranucleotides, were sent back for analysis. The 192 clones were sequenced by a commercial sequencing company (AGOWA GmbH, Germany) and SSRs were identified in the clones using 'find microsat Win32' (Salamin, unpublished).

80 primer pairs were designed manually or using Primer3 software (http://frodo.wi.mit.edu/primer3/) to amplify the SSR regions. The 80 sets were selected in order to have an equal ratio between di- and tetranucleotide SSRs (Table F in Appendix)

4.3.3 Amplification and SSRs detection

To select the most suitable set of primers among the total of 80, an amplification test was conducted on eight genotypes for each pair. The eight genotypes were chosen to have at least two representatives of each of the three species (*M. ×giganteus*, *M. sacchariflorus* and *M. sinensis*). A template DNA volume of 1 µl (40ngµl^{-1}) was amplified with an initial denaturation of 5 min at 95°C followed by 35 cycles each with a denaturation of 1 min at 95°C, 1 min at a primer-specific annealing temperature (Table 4.3.1), and an extension of 1 min at 72°C, followed by a final extension at 72°C for 10 min. The reaction mixture (final volume) contained 1× reaction buffer (New England Biolabs) containing 2mM MgSO$_4$, 0.125µM dNTPs, 0.25µM of each primer, 0.5U of *Taq* DNA polymerase (New England Biolabs).

The PCR products were loaded on 3% MetaPhor® Agarose (Lonza, Rockland, ME, USA) gels. For primers which showed weak amplification, PCR conditions were optimized using a gradient PCR, with temperatures ranging from 48°C to 60°C, and the amplification test (on eight samples) was repeated using the optimal annealing temperature. 30 primer pairs producing the best amplification were selected to be used on the full dataset.
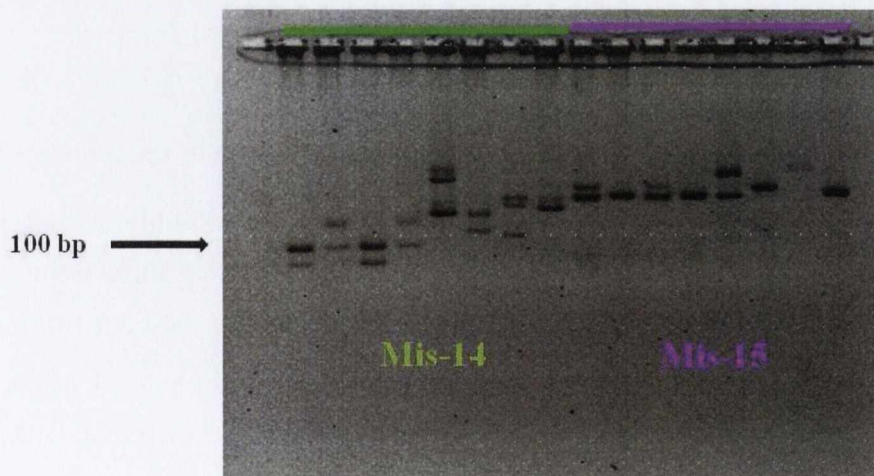
Figure 4.3.1 An example of amplification of the initial sample using Mis-14 and Mis-15 markers.

Table 4.3.1 List of 30 primer pairs developed for SSR amplification and genotyping. T (°C) indicates the annealing temperature used for PCR based on gradient. Shading highlights pools for genotyping runs on the genetic analyzer.

Primer name	Clone	SSR motif	Dye	Pool	Forward sequence 5' - 3'	Reverse sequence 5' – 3'	T (°C)	Sequence length (bp)	SSR size (bp)
Mis-01	SSR1A10	(TCTA)$_{20}$	FAM	1	CAGTCCTTGGAGCAGGCTAT	AAGATCTCAAACCTATAGTC	54	202	80
Mis-23	SSR1G9	(ATCT)$_{15}$	TAMRA	1	CACGAACTGAATCAGCATGC	GTAGCTGCAACTGCTAGTGT	60	240	60
Mis-22	SSR1G8	(TAGA)$_{17}$	VIC	1	CGAGCGAGCCTGCATGTGTG	TTGACGTCAGCAAGATATTG	54	173	68
Mis-37	SSR2D9	(TC)$_{34}$	FAM	2	GAATGCAGTCATCAGCAGCT	TGGACATCTCTAGGTTGATC	54	218	68
Mis-52	SSR2C11	(GA)$_{19}$	NED	2	TTATTGGTGCCCAAAGGTGT	AACAAGCCCTCAAGCTTCCT	60	370	38
Mis-50	SSR2H9	(GA)$_{21}$	ROX	2	TACGGACGATTAACCAAGCC	CGCAAGGTGCAGGACCATCA	54	230	42
Mis-66	SSR1D5	(CTAT)$_{13}$	VIC	2	CATGGCTACAGGCACCTAAAA	ATAACGAGAAATGGCCGATG	60	165	52
Mis-14	SSR1F12	(GATA)$_{15}$	FAM	3	GTAGCTGCAACTGCTAGTGT	ACTCGCATTGGTTGGTATGA	59	141	60
Mis-78	SSR2G11	(CT)$_{15}$	NED	3	TCTGCAGGTGACAAGGAAGA	GTCAACCGGCATAGTTCGAT	60	167	30
Mis-41	SSR2F5	(GA)$_{24}$	ROX	3	ATAATGCAGGTCAGTTCAAC	CGCAGCTAGCTGCTTGTCAG	54	226	48
Mis-67	SSR1E10	(TCTA)$_{13}$	VIC	3	CCTCTGCGGATATGAGGTGT	GAAGTGACAACATGCGATGG	60	175	52
Mis-15	SSR1F2	(ATCT)$_{16}$	FAM	4	ACTACTGCATGCATCATGATG	TGCTTCGCGGCGAAGTTTCA	59	195	64
Mis-20	SSR1G12	(TCTA)$_{17}$	TAMRA	4	TAGCTGAGCTGTCTATGGTA	TAGCCATTGAGGCTAAGGAT	54	249	68
Mis-24	SSR1H10	(AGAT)$_{15}$	VIC	4	ATACACGATCCAAACATGTC	ATGTGCTCACCCAAGAGATG	60	324	60
Mis-60	SSR2C3	(GA)$_{16}$	FAM	5	AGATGGCAGCTTGCTCTTGT	CCATTTGTTGAGCACGATGT	60	190	32
Mis-69	SSR1F4	(TCTA)$_{13}$	NED	5	CCTCTGCGGATATGAGGTGT	GAAGTGACAACATGCGATGG	60	175	52
Mis-63	SSR1G3	(TCTA)$_{14}$	VIC	5	AGGCTAGCACTTCCTCCAAA	CTGCCTGGTGACCCCTATAA	60	234	56
Mis-59	SSR2B3	(GA)$_{16}$	FAM	6	GAGCTGATCGCGTAGCAAG	TTCGATAAACAGGGGATTGG	60	152	32
Mis-54	SSR2A11	(CT)$_{18}$	NED	6	TAAGAAACGCAGCAGCAGAA	AGTCTCCGGCTTTCTCACAA	60	226	36
Mis-13	SSR1F10	(TAGA)$_{19}$	ROX	6	CGGACTAACTTGTGAATCTT	GTCCTTGGAGCAGGCTATGA	54	230	76
Mis-71	SSR1D3	(TAGA)$_{12}$	VIC	6	CAACCATGAGCACTTCTCCA	AACATAGGAGGCCAAGCAAA	60	179	48
Mis-51	SSR2G4	(TC)$_{20}$	FAM	7	GATCCATCACGGATTCATCA	ATCATAGGCAAAACGGATCG	60	164	40
Mis-70	SSR1B10	(TATC)$_{12}$	NED	7	TCGCACCTTTAATTTTTGCAT	TTATGAACCCGACAGGGAGA	60	249	48
Mis-79	SSR2G9	(CT)$_{15}$	VIC	7	GCCAACTCGTGGATTTGAGT	CGTAGCAAGAGGGGAACAAA	60	248	30
Mis-53	SSR2G10	(GA)$_{19}$	FAM	8	AGGCAGCACCTCACAAAACT	GGTGGAGATGCTCTTCTTGC	60	173	38
Mis-64	SSR1G6	(AGAT)$_{14}$	NED	8	TCCCCTTAGTGTCCGTGAAG	GAGGCAGGTGTAGTCGGAGA	60	236	56
Mis-55	SSR2B9	(GA)$_{18}$	VIC	8	CGGCTTCGAGTGATACCTTT	TACCGGATTTAAGGGGCTTT	60	250	36
Mis-42	SSR2F6	(AG)$_{31}$	FAM	9	GCCGCCAGGCTCCCAAGCCT	ATCCGAGCCATGTATGCACG	54	206	62
Mis-33	SSR2B7	(CT)$_{20}$	TAMRA	9	TGACATAGGGCTACACATAT	CGAGTGAGGCAGCTAGTTCA	48	242	40
Mis-16	SSR1F5	(TATC)$_{13}$/(TCTA)$_{16}$	VIC	9	ATCTTGCCTAGGATGCATTAG	TGGTCTATTACAACAAGGCT	60	264	52+64*

Mis-16 was a compound microsatellite with two repetitive sequences separated by a non-polymorphic region.

4.3.4 Genotyping

Five different fluorescent dyes were used for primer labelling to allow multiplexing, in pools, as shown by the shading in Table 4.3.1. A polyA treatment at 65°C was applied for 30 min to the PCR products. 0.5 µl from each pool was added to 9.5µl of a mix of 9.25µl formamide + 0.25µl LIZ500 internal sizing standard. The PCR products were then sized using an ABI 3130xl automated DNA sequencer and the resulting peaks were scored with GENEMAPPER ™ V4.0 software (Applied Biosystems). 11 loci were not consistently amplified across our

collection and were discarded from further analyses. Our final analysis therefore included 19 SSR markers (Table 4.3.1).

4.3.5 Data analyses

Allele number and size range were calculated for each locus. The polymorphism information content (PIC) values were calculated according to Röder et al. (1995) as $1 - \Sigma \, [p_i]$, where p_i is the frequency of the ith allele.

Genetic distance

Due to the extensive occurrence of polyploidy in the collection (as determined by flow cytometry; Chapter 2), many samples had more than two alleles at a particular locus. It was therefore necessary to transform the data matrix in to a binary matrix scoring 1 for presence of alleles and 0 for absence. Genetic similarity (GS) indices were calculated using the Jaccard's coefficient (S_j) for all possible pairwise comparisons. The Jaccard's coefficients were calculated as $Sj = a_{12} / (a_{12} + a_1 + a_2)$, where a_{12} is the number of alleles shared between two genotype, a_1 is the number of alleles unique to the first genotype, and a_2 the number of bands unique to the second genotype. Sj disregards the conjoint absence of alleles in the pairwise comparison, reducing the risk of over-estimating similarity. Jaccard's coefficients were calculated using the software FreeTree (Pavlícek et al. 1999) and used to cluster genotypes according to similarity. The UPGMA (unweighted pair group method using arithmetic means) clustering tree building approach was used, with internal support assessed using 1000 bootstrap replicates. The UPGMA tree was visualized using FIGTREE 1.2.1 (Rambaut 2007).

PCO

Principle coordinates analysis was performed on the data using NTSYSpc v2.2 software (Rohlf 2008) starting from the binary matrix. Sj coefficients were calculated using the SIMQUAL module and the resulting similarity matrix was transformed to scalar product form using the DCENTER module in order that eigenvalues and eigenvectors could be determined. This 'double centers' the distance matrix by first replacing the off-diagonal

97

element. The row and column means are then subtracted from each element and the grand mean is added on. Using the EIGEN module, this matrix is factored so that the elements of the eigenvectors corresponding to positive eigenvalues can be interpreted as the coordinates of each point in a Cartesian space. For a better interpretation of the results, a three dimensional graph of the eigenvectors and eigenvalues was construct using Minitab® 16.2.0 (2007) software.

Structure

The software STRUCTURE 2.3.3 (Pritchard et al. 2000) was used to infer the genetic structure of the collection. A series of simulations were run with the number of clusters K ranging from 1 to 8, with three independent runs for each K value. Each run consisted of a burn-in period of 10,000 steps and 10,000 MCMC (Monte Carlo Markov Chain) replicates, assuming an admixture model and uncorrelated allele frequencies. No prior information about the structure of the population was defined. The most likely value of K was chosen following Evanno et al. (2005) and used to run a simulation with a burn-in period of 10,000 steps and 100,000 MCMC replicates.

AMOVA

An analysis of molecular variance, AMOVA, (Excoffier et al. 1992) was carried out with GenAlEx 6 (Peakall and Smouse 2006) to estimate the components of genetic variation between and within groups as observed in the UPGMA dendrogram and in the PCO. 999 permutations were used to test for statistical significance.

4.4 Results

Polymorphism at 19 microsatellite loci was studied in a collection of 176 individual grasses, mostly belonging to the species *M. sinensis, M. sacchariflorus* and *M.* ×*giganteus*. 14 individuals belonging to closely related genera were also included. All markers revealed considerable length polymorphism, with the number of alleles ranging from 13 to 44, with an average of 27.5 (Table 4.4.1). PIC values ranged from 0.65 to 0.91, with an average of 0.83. Thirteen out of 19 primer pairs showed cross-amplification in non-*Miscanthus* species (Table 4.4.1).

Table 4.4.1 Expected heterozygosity (He) and PIC values for 19 nuclear SSR markers. Cross-amplification in *Miscanthus* species other than *M. sinensis, M. sacchariflorus* and *M.* ×*giganteus* and in non-Miscanthus species (in blue) is shown (v = yes; × = no).

	Alleles	Size	He	PIC	M. capense	M. condensatus	M. ecklonii	M. erectum	M. fusca	M. junceum	M. nepalensis	M. nudipes	M. oligostachyus	M. sorghum	M. teretifolium	M. tinctorius	M. violaceus	Eulalia	Narenga	S. contortus	Saccharum officinarum	Saccharum	Sorghum halpense	Spodiopogon	Pennisetum	Zea diploperennis	Cymbopongon citratus
Mis-1	20	125-256	0.86	0.85	×	v	×	×	×	×	×	×	×	×	×	×	v	×	×	×	×	×	×	×	×	×	×
Mis-14	33	71-208	0.91	0.90	×	v	×	×	×	×	v	v	×	×	×	×	×	×	×	×	×	v	v	v	×	×	×
Mis-15	21	144-205	0.78	0.75	×	v	×	×	×	×	v	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
Mis-20	33	197-300	0.83	0.82	×	v	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
Mis-22	16	103-174	0.68	0.66	×	v	×	×	×	×	×	×	×	×	×	v	×	×	×	×	×	×	×	×	×	×	×
Mis-23	30	176-314	0.91	0.90	×	v	×	×	×	×	×	×	×	×	×	v	×	×	×	×	×	×	×	×	×	×	×
Mis-24	23	248-361	0.85	0.84	×	v	×	×	×	×	v	×	×	×	×	×	×	×	×	×	v	v	v	×	×	×	×
Mis-37	33	169-226	0.89	0.88	×	v	×	×	×	×	v	×	×	×	×	×	×	v	×	×	×	×	×	×	×	×	v
Mis-41	44	131-512	0.90	0.89	×	v	×	×	×	×	v	×	×	×	×	×	×	×	×	×	v	v	×	×	×	×	×
Mis-42	29	121-247	0.91	0.90	×	v	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	v
Mis-50	30	199-260	0.82	0.81	×	v	×	×	×	×	v	×	×	×	×	×	×	×	×	×	×	v	×	×	×	×	×
Mis-51	27	132-176	0.82	0.81	×	v	v	v	v	v	v	×	×	v	×	v	v	×	×	×	v	×	×	×	×	×	×
Mis-52	22	132-207	0.85	0.83	×	v	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
Mis-54	20	207-244	0.87	0.86	×	v	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	v	v	v
Mis-59	13	123-162	0.76	0.72	×	v	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
Mis-64	40	177-286	0.91	0.91	×	v	×	×	×	×	v	×	×	×	×	v	×	v	×	v	×	×	×	×	×	×	×
Mis-69	24	105-220	0.85	0.83	×	v	×	×	×	v	×	×	×	×	×	v	v	×	v	×	v	×	×	×	×	×	×
Mis-70	31	211-328	0.82	0.80	×	v	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	v	×
Mis-79	34	224-276	0.92	0.91	×	v	×	×	v	×	v	×	×	×	×	v	×	×	×	×	×	v	×	v	v	×	v
Average	27.5			0.84																							

The UPGMA constructed from the matrix of the Jaccard's coefficients is shown in Figure 4.4.1.
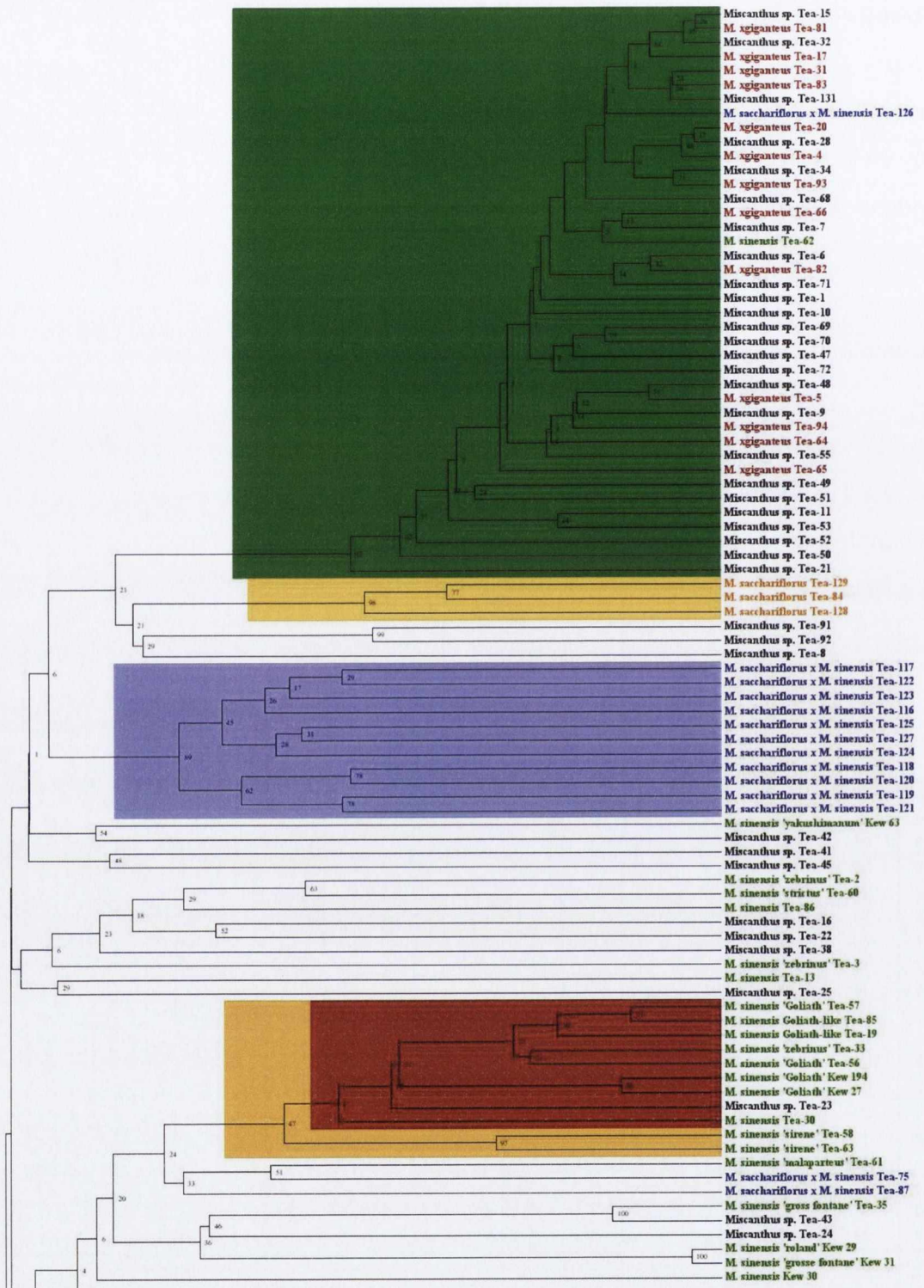
Figure 4.4.1 *(continued)*

100
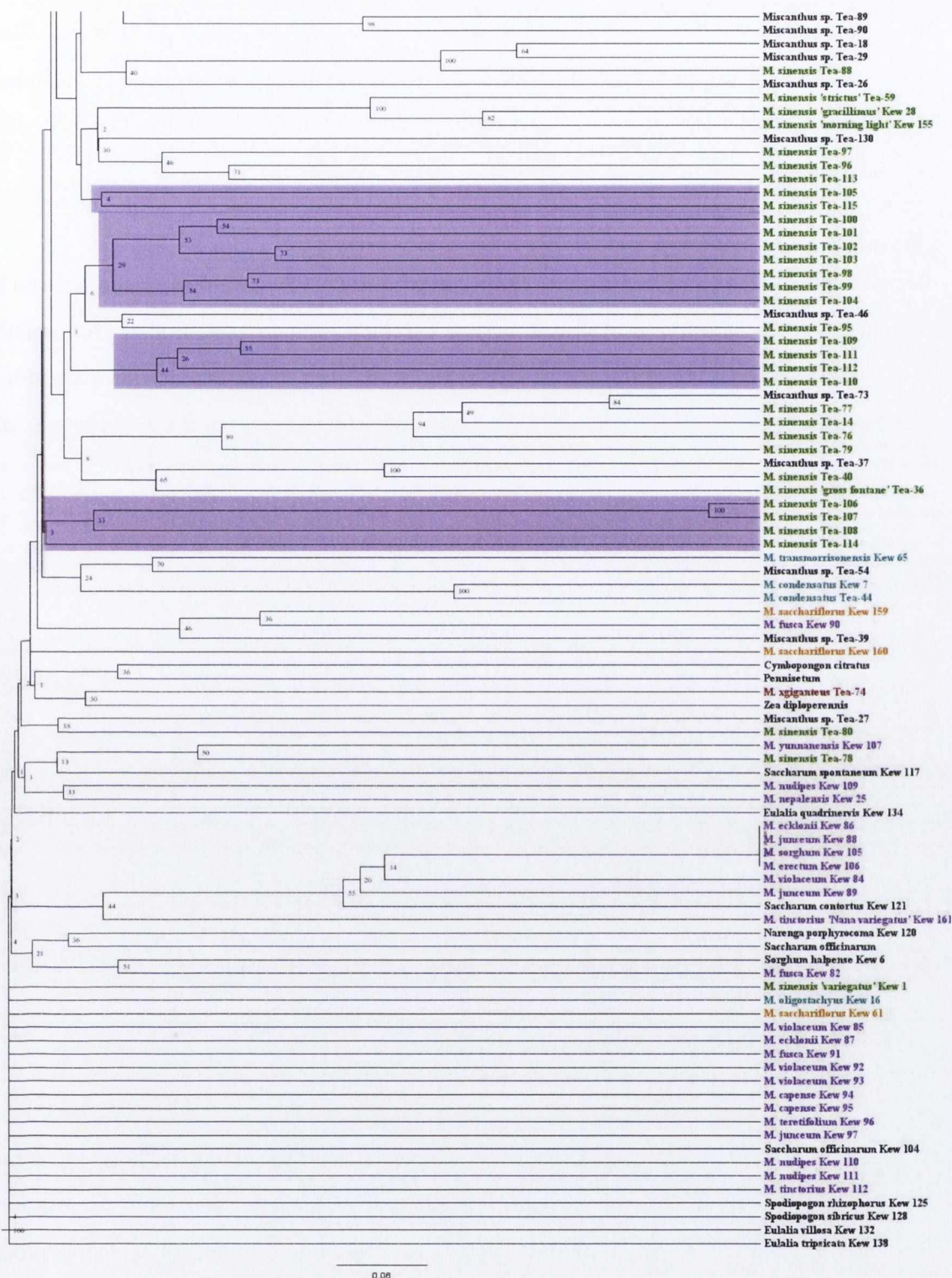
Figure 4.4.1 UPGMA tree showing inter-relationships of individuals using a combination of 19 SSR markers. Accessions names are coloured according to species: Red = *M. ×giganteus;* Green = *M. sinensis;* Yellow = *M. sacchariflorus;* Blue = *M. sacchariflorus×M. sinensis* hybrids; Light blue = *Miscanthus s.s;* Violet = other *Miscanthus*; Black = non-*Miscanthus* species and unclassified *Miscanthus* accessions.

The UPGMA tree shows a cluster of 42 *M. ×giganteus* accessions (highlighted in green) that are clearly separated from the remaining samples. The closest group to this cluster includes three individuals of the *M. sacchariflorus* genotype M11 MATREC11 (Tea-84, Tea-128 and Tea-129) and a group containing Tea-91, Tea-92 and Tea-8.

With the exception of the genotype Tea-126, all the triploid individuals from the Swedish germplasm collection group together (lilac) and are sister to the above mentioned group. The diploid Swedish genotypes are split into clearly separated clades, but they are not exclusive, as they include also other individuals of the species *M. sinensis*. Among the different *M. sinensis* varieties, only the individuals of the 'Goliath' type form a definite cluster.

4.4.1 Principal coordinates analysis

Table 4.4.2 shows four axes of the PCO analysis that cumulatively account for 23.97% of the variation.

Table 4.4.2 Eigenvalues and percentage of variation expressed by each axis for nSSRs dataset.

Axis	Eigenvalue	Percent	Cumulative
1	21.85993359	13.93	13.93
2	6.65572614	4.24	18.18
3	4.84701806	3.09	21.27
4	4.22192863	2.69	23.96

The first eigenvalue accounts for 13.94% of the total variation, the second 4.2% of the total variation (cumulative 18.18%) and the third 3.09% of the total variation (cumulative 21.27%). The eigenvectors were plotted in both two- and three-dimensional scatterplots (Figure 4.4.2). For a better comparison of the results, some of groups defined by the UPGMA analysis were labelled.
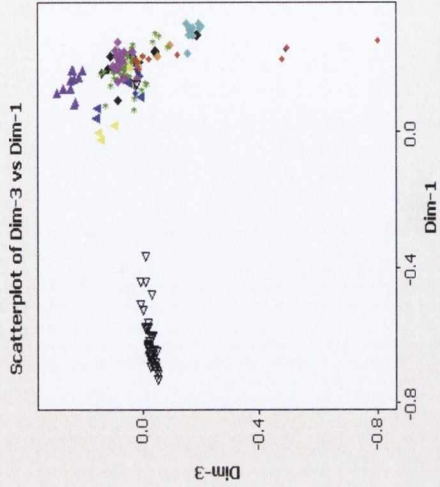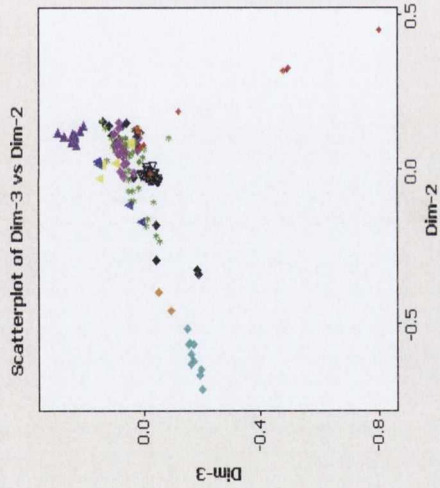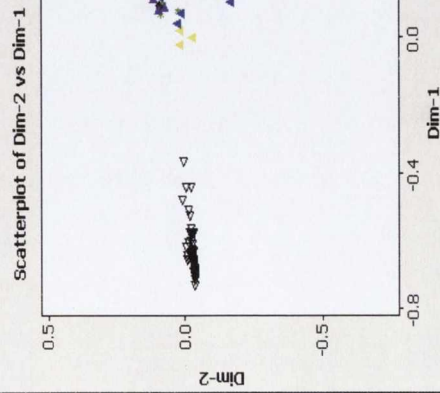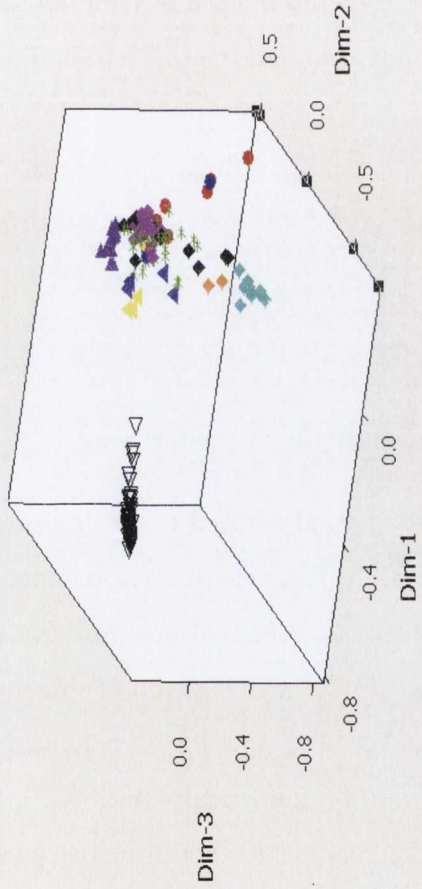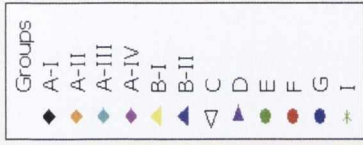
Figure 4.4.2 Principal coordinates analysis scatterplots in three dimensions and relative two-dimensional projections for the nuclear SSR data. Groups : A (◊) *M. sinensis* (A-II var. 'sirene'; A-III var. 'Goliath'; A-IV Swedish clones); B (▲) *M. sacchariflorus*; C (▽) *M. ×giganteus*; D (▶) Swedish *M. sinensis × M. sacchariflorus*; E (●) other *Miscanthus sensu stricto*; F (●) red) *Miscanthus sensu lato*; G (● blue) non-*Miscanthus* species; I (∗) *Miscanthus sp.*

103

A group of 42 accessions (group C in Figure 4.4.2) was clearly separated from the remaining individuals. This group correspond to the cluster of *M. ×giganteus* in the UPGMA tree. Along the first axis, the group closest to C is B-I, which includes the three M11 MATREC11 individuals of *M. sacchariflorus*. Along the second dimension what is noticeable in both scatterplot (b) and (c) is the spread of the individuals belonging to *M. sinensis* (group A). Two subgroups of *M. sinensis* separate from the core batch: group A-II and A-III. Group A-III corresponds to the Goliath-like group in the UPGMA tree, whereas A-II includes two individuals of *M. sinensis* var. *sirene* that are the most closely related to the Goliath group in the tree. The introduction of the third dimension allows the resolution of two further groups (tagged as F and D in Figure 4.4.2). Group D includes all the triploid hybrids of the Sweden genotypes, with the exception of Tea-126, which falls in group C. The diploid genotypes (tagged as A-IV), on the other hand, are not resolved from the core group. Group F includes individuals belonging to *Miscanthus* species that are not classified as *Miscanthus sensu stricto* (Clifton-Brown et al. 2008). As in the UPGMA tree, the genotypes Tea-89, Tea-90, Tea-91 and Tea-92 (group B-II) form two separated pairs, with one closely related to the M11 MATREC11 group. Species classified as *Miscanthus s.s* other than *M. sinensis, M. sacchariflorus* and *M. ×giganteus* (group E) and the non-*Miscanthus* species included in the study do not form any obvious groupings in the PCO.

4.4.2 Cluster analysis with STRUCTURE

The genetic structure of the population was detected using a model-based clustering method (Pritchard et al. 2000) as implemented in STRUCTURE 2.3.3. A series of three independent runs for each value of K (i.e. the number of populations in the collection) was run. For each run, the estimated log probability of data $Pr(X|K)$ for each value of K is given (Table 4.4.3). The mean likelihood, indicated as $L(K)$ afterwards, over the three runs for each K was first plotted against K (Figure 4.4.3). $L(K)$ could be seen increasing dramatically until K=3, after which it reaches a plateau. In order to harvest the true value for K, three additional steps were introduced, following (Evanno et al. 2005). In the second step, the mean difference between successive values of likelihood of K $L'(K) = L(K)-L(K-1)$ was calculated and in the third step, the absolute value of the difference between successive values of $L'(K)$, $|L''(K)| = |L'(K+1)-L'(K)|$ (Table 4.4.3).

Table 4.4.3 Evanno parameters calculated for each of the three runs for K values from one to eight.

Run #	K	Raw STRUCTURE output			Raw Evanno Data Table		
		Est. Ln Prob of Data	Mean value of Ln likelihood	Variance of ln likedlihood	LnP(K)	L'(K)	|L''(K)|
1	1	-14143.5	-13957.8	371.6	-14143.5	N.A.	N.A.
2	1	-14154.2	-13957.9	392.6	-14154.2	N.A.	N.A.
3	1	-14148.9	-13957.7	382.3	-14148.9	N.A.	N.A.
4	2	-12471.8	-12068.8	805.9	-12471.8	1671.7	663.1
5	2	-12172.6	-11787.8	769.7	-12172.6	1981.6	1260.2
6	2	-12129.6	-11759.6	740.1	-12129.6	2019.3	1361.7
7	3	-11463.2	-10943.1	1040.1	-11463.2	1008.6	510.6
8	3	-11451.2	-10903.6	1095.1	-11451.2	721.4	217.6
9	3	-11472	-10939.6	1064.8	-11472	657.6	174.9
10	4	-10965.2	-10287.3	1355.8	-10965.2	498	617.3
11	4	-10947.4	-10288.6	1317.7	-10947.4	503.8	349.1
12	4	-10989.3	-10290.1	1398.3	-10989.3	482.7	344.2
13	5	-11084.5	-10096.8	1975.4	-11084.5	-119.3	447.6
14	5	-10792.7	-10032.1	1521.2	-10792.7	154.7	102.2
15	5	-10850.8	-10042.3	1617	-10850.8	138.5	51.7
16	6	-10756.2	-9882.1	1748.1	-10756.2	328.3	338.7
17	6	-10740.2	-9871.8	1736.9	-10740.2	52.5	129.3
18	6	-10764	-9920.5	1686.8	-10764	86.8	55.1
19	7	-10766.6	-9835.6	1862	-10766.6	-10.4	1.4
20	7	-10817	-9820.6	1992.7	-10817	-76.8	182.5
21	7	-10732.3	-9835.3	1794.1	-10732.3	31.7	8.1
22	8	-10775.6	-9755.5	2040.1	-10775.6	-9	N.A.
23	8	-10711.3	-9751	1920.6	-10711.3	105.7	N.A.
24	8	-10708.7	-9762.8	1891.8	-10708.7	23.6	N.A.

Finally the value ΔK is estimated as the mean of $|L''(K)|$ averaged over the three runs divided by the standard deviation of $L(K)$, $\Delta K = m|L''(K)| / s[L(K)]$.
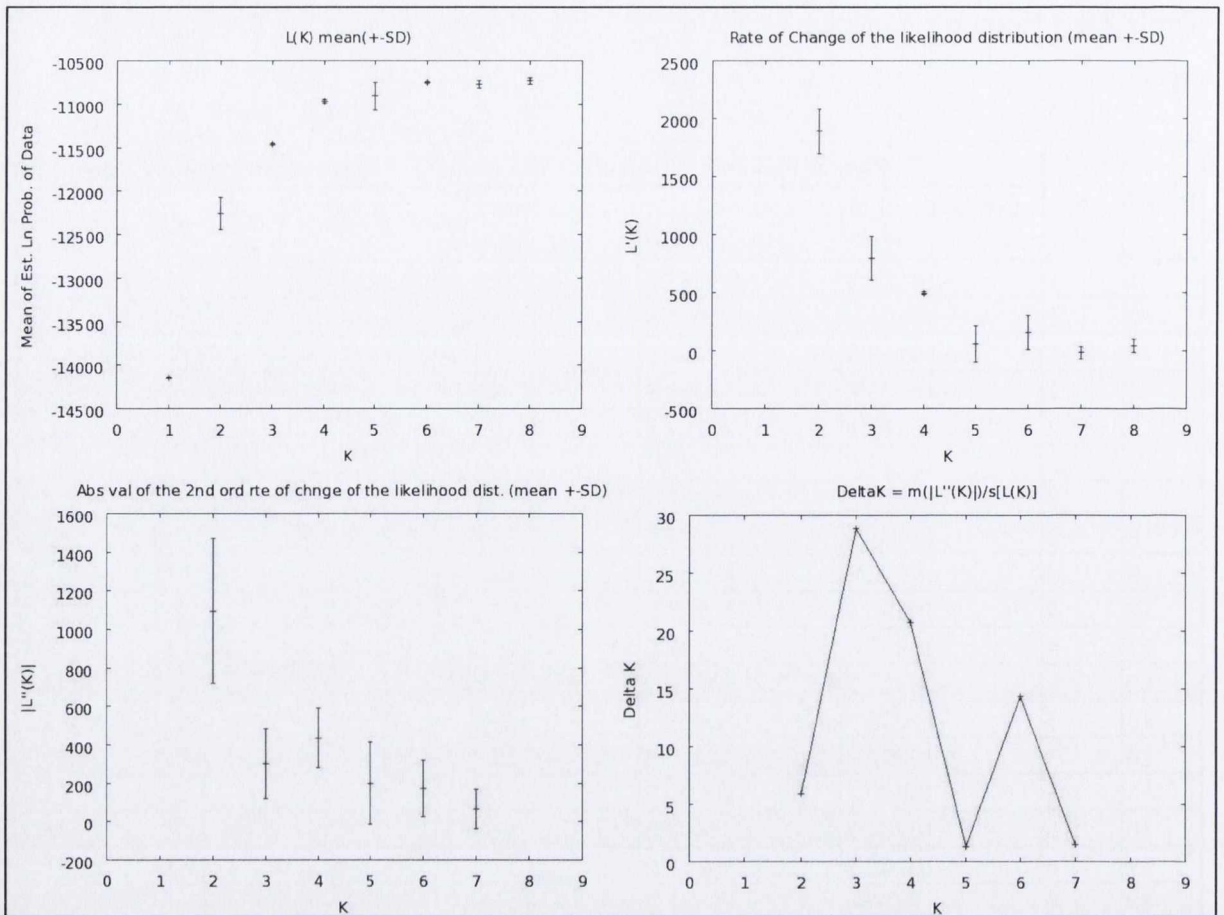
Figure 4.4.3 Graphical representation of the Evanno parameters for the estimation of the K value.

Plotting ΔK against the values of K, a peak is detected corresponding to the true value of K for the data, in our case at K=3. A simulation for this value of K was then run in order to assign the individuals to each of the three clusters. The results are summarized in Table 4.4.4 and in Figure 4.4.4.

Table 4.4.4 Accessions assigned to each cluster according to STRUCTURE analysis.

ID	Groups	%Missing	Inferred clusters		
			I	II	III
Miscanthus sp. Tea-1	C	28	0.998	0.001	0.001
M. sinensis 'zebrinus' Tea-2	I	57	0.011	0.004	0.986
M. sinensis 'zebrinus' Tea-3	I	84	0.004	0.007	0.989
M. xgiganteus Tea-4	C	25	0.998	0.001	0.001
M. xgiganteus Tea-5	C	28	0.998	0.001	0.001
Miscanthus sp. Tea-6	C	28	0.999	0.001	0.001
Miscanthus sp. Tea-7	C	32	0.998	0.001	0.001
Miscanthus sp. Tea-8	I	42	0.963	0.032	0.005

Table 4.4.4 *(continued)*

ID	Groups	%Missing	Inferred clusters		
			I	II	III
Miscanthus sp. Tea-8	I	94	0.814	0.041	0.145
Miscanthus sp. Tea-9	C	0	0.999	0.001	0.001
Miscanthus sp. Tea-10	C	28	0.999	0.001	0.001
Miscanthus sp. Tea-11	C	32	0.998	0.001	0.001
M. sinensis Tea-13	I	55	0.015	0.281	0.704
M. sinensis Tea-14	I	55	0.002	0.996	0.002
Miscanthus sp. Tea-15	C	25	0.999	0.001	0.001
Miscanthus sp. Tea-16	I	68	0.005	0.006	0.989
M. xgiganteus Tea-17	C	25	0.999	0.001	0.001
Miscanthus sp. Tea-18	I	57	0.002	0.002	0.996
M. sinensis goliath-like Tea-19	A-III	21	0.015	0.001	0.985
M. sinensis goliath-like Tea-19	A-III	36	0.001	0.001	0.998
M. xgiganteus Tea-20	C	25	0.999	0.001	0.001
Miscanthus sp. Tea-21	C	25	0.998	0.001	0.001
Miscanthus sp. Tea-21	I	52	0.007	0.002	0.991
Miscanthus sp. Tea-22	I	60	0.003	0.004	0.993
Miscanthus sp. Tea-23	I	78	0.35	0.003	0.646
Miscanthus sp. Tea-23	A-III	26	0.001	0.001	0.998
Miscanthus sp. Tea-24	I	57	0.002	0.009	0.988
Miscanthus sp. Tea-25	I	63	0.031	0.909	0.06
Miscanthus sp. Tea-26	I	60	0.004	0.125	0.871
Miscanthus sp. Tea-27	I	81	0.004	0.01	0.986
Miscanthus sp. Tea-28	C	25	0.998	0.001	0.001
Miscanthus sp. Tea-29	I	63	0.002	0.004	0.994
M. sinensis Tea-30	A-III	50	0.001	0.001	0.998
M. xgiganteus Tea-31	C	25	0.999	0.001	0.001
Miscanthus sp. Tea-32	C	25	0.999	0.001	0.001
M. sinensis 'zebrinus' Tea-33	A-III	15	0.001	0.001	0.998
Miscanthus sp. Tea-34	C	25	0.999	0.001	0.001
Miscanthus sp. Tea-34	C	28	0.999	0.001	0.001
M. sinensis 'gross fontane' Tea-35	I	55	0.004	0.002	0.994
M. sinensis 'gross fontane' Tea-36	I	57	0.004	0.974	0.022
Miscanthus sp.Tea-37	I	55	0.005	0.983	0.013
Miscanthus sp. Tea-38	I	57	0.072	0.059	0.869
Miscanthus sp.Tea-39	I	92	0.977	0.012	0.011
M. sinensis Tea-40	I	52	0.002	0.991	0.007
Miscanthus sp. Tea-41	I	84	0.842	0.067	0.091
Miscanthus sp. Tea-41	I	73	0.003	0.031	0.966
Miscanthus sp. Tea-42	I	57	0.004	0.94	0.057
Miscanthus sp. Tea-43	I	81	0.003	0.004	0.993
Miscanthus sp. Tea-43	I	60	0.001	0.002	0.996
M. condensatus Tea-44	E	57	0.01	0.037	0.953
Miscanthus sp. Tea-45	I	78	0.204	0.341	0.455
Miscanthus sp. Tea-46	I	63	0.002	0.948	0.05
Miscanthus sp. Tea-47	C	50	0.998	0.001	0.001
Miscanthus sp. Tea-48	C	25	0.999	0.001	0.001
Miscanthus sp. Tea-49	C	36	0.998	0.001	0.001
Miscanthus sp. Tea-50	C	48	0.998	0.001	0.001
Miscanthus sp. Tea-51	C	28	0.998	0.001	0.001
Miscanthus sp. Tea-52	C	44	0.998	0.001	0.001
Miscanthus sp. Tea-53	C	40	0.998	0.001	0.001
Miscanthus sp.Tea-54	I	57	0.002	0.23	0.768
Miscanthus sp.Tea-55	C	28	0.999	0.001	0.001
M. sinensis 'goliath'Tea-56	A-III	15	0.001	0.001	0.999
M. sinensis 'goliath' Tea-57	A-III	5	0.001	0.001	0.999
M. sinensis 'sirene' Tea-58	A-II	65	0.003	0.002	0.995
M. sinensis 'strictus' Tea-59	I	44	0.004	0.002	0.993
M. sinensis 'strictus' Tea-60	I	73	0.005	0.009	0.986

Table 4.4.4 *(continued)*

ID	Groups	%Missing	Inferred clusters		
			I	II	III
M. sinensis 'malaparteus' Tea-61	A-I	65	0.002	0.002	0.995
M. sinensis Tea-62	C	32	0.998	0.001	0.001
M. sinensis 'sirene' Tea-63	A-II	63	0.003	0.003	0.994
M. xgiganteus Tea-64	C	52	0.998	0.001	0.001
M. xgiganteus Tea-64	C	25	0.999	0.001	0.001
M. xgiganteus Tea-65	C	32	0.998	0.001	0.001
M. xgiganteus Tea-66	C	25	0.999	0.001	0.001
Miscanthus sp. Tea-68	C	28	0.998	0.001	0.001
Miscanthus sp. Tea-69	C	28	0.998	0.001	0.001
Miscanthus sp. Tea-70	C	28	0.999	0.001	0.001
Miscanthus sp. Tea-71	C	28	0.999	0.001	0.001
Miscanthus sp. Tea-72	C	28	0.999	0.001	0.001
Miscanthus sp. Tea-73	I	63	0.002	0.996	0.002
M. xgiganteus Tea-74	C	84	0.924	0.069	0.007
M. sacchariflorus x M. sinensis Tea-75	I	52	0.018	0.008	0.973
M. sinensis Tea-76	A-I	57	0.002	0.995	0.003
M. sinensis Tea-77	A-I	55	0.002	0.997	0.002
M. sinensis Tea-78	A-I	92	0.064	0.159	0.776
M. sinensis Tea-79	A-I	55	0.001	0.996	0.003
M. sinensis Tea-80	A-I	60	0.004	0.953	0.043
M. xgiganteus Tea-81	C	25	0.999	0.001	0.001
M. xgiganteus Tea-82	C	28	0.999	0.001	0.001
M. xgiganteus Tea-83	C	25	0.999	0.001	0.001
M. sacchariflorus Tea-84	B-I	10	0.988	0.005	0.007
M. sinensis goliath-like Tea-85	A-III	5	0.001	0.001	0.999
M. sinensis Tea-86	I	57	0.004	0.006	0.989
M. sacchariflorus x M. sinensis Tea-87	I	55	0.099	0.041	0.86
M. sinensis Tea-88	I	55	0.002	0.002	0.996
Miscanthus sp. Tea-89	B-II	57	0.314	0.006	0.681
Miscanthus sp. Tea-90	B-II	31	0.472	0.002	0.526
Miscanthus sp. Tea-91	B-II	31	0.984	0.007	0.009
Miscanthus sp. Tea-92	B-II	21	0.929	0.011	0.06
M. xgiganteus Tea-93	C	25	0.999	0.001	0.001
M. xgiganteus Tea-94	C	28	0.999	0.001	0.001
M. sinensis Tea-95	A-IV	57	0.002	0.931	0.068
M. sinensis Tea-96	A-IV	57	0.002	0.01	0.988
M. sinensis Tea-97	A-IV	60	0.002	0.124	0.875
M. sinensis Tea-98	A-IV	50	0.002	0.778	0.22
M. sinensis Tea-99	A-IV	55	0.008	0.916	0.077
M. sinensis Tea-100	A-IV	55	0.001	0.914	0.085
M. sinensis Tea-101	A-IV	55	0.013	0.847	0.141
M. sinensis Tea-102	A-IV	52	0.011	0.788	0.202
M. sinensis Tea-103	A-IV	52	0.002	0.917	0.081
M. sinensis Tea-104	A-IV	63	0.002	0.92	0.078
M. sinensis Tea-105	A-IV	52	0.002	0.399	0.599
M. sinensis Tea-106	A-IV	52	0.01	0.98	0.01
M. sinensis Tea-107	A-IV	52	0.01	0.98	0.01
M. sinensis Tea-108	A-IV	55	0.002	0.963	0.035
M. sinensis Tea-109	A-IV	57	0.002	0.795	0.203
M. sinensis Tea-110	A-IV	63	0.003	0.993	0.004
M. sinensis Tea-111	A-IV	55	0.002	0.98	0.017
M. sinensis Tea-112	A-IV	57	0.007	0.99	0.003
M. sinensis Tea-113	A-IV	55	0.001	0.022	0.977
M. sinensis Tea-114	A-IV	55	0.042	0.921	0.037
M. sinensis Tea-115	A-IV	55	0.003	0.765	0.232
M. sacchariflorus x M. sinensis Tea-116	D	25	0.132	0.867	0.001
M. sacchariflorus x M. sinensis Tea-117	D	30	0.019	0.98	0.001
M. sacchariflorus x M. sinensis Tea-118	D	32	0.002	0.997	0.001

Table 4.4.4 *(continued)*

ID	Groups	%Missing	Inferred clusters		
			I	II	III
M. sacchariflorus x M. sinensis Tea-119	D	28	0.001	0.998	0.001
M. sacchariflorus x M. sinensis Tea-120	D	28	0.002	0.997	0.001
M. sacchariflorus x M. sinensis Tea-121	D	36	0.002	0.997	0.001
M. sacchariflorus x M. sinensis Tea-122	D	32	0.005	0.994	0.001
M. sacchariflorus x M. sinensis Tea-123	D	44	0.092	0.906	0.001
M. sacchariflorus x M. sinensis Tea-124	D	40	0.007	0.992	0.001
M. sacchariflorus x M. sinensis Tea-125	D	28	0.003	0.996	0.001
M. sacchariflorus x M. sinensis Tea-126	C	28	0.999	0.001	0.001
M. sacchariflorus x M. sinensis Tea-127	D	28	0.008	0.99	0.002
M. sacchariflorus Tea-128	B-I	47	0.987	0.003	0.01
M. sacchariflorus Tea-129	B-I	44	0.994	0.002	0.004
Miscanthus sp. Tea-130	I	60	0.006	0.775	0.219
Miscanthus sp. Tea-131	C	25	0.999	0.001	0.001
M. sinensis 'variegatus' Kew 1	A-I	97	0.022	0.945	0.033
Sorghum alpense nKew 6	G	92	0.107	0.78	0.112
M. condensatus Kew 7	E	68	0.009	0.009	0.982
M. oligostachyus Kew 16	E	100	0.334	0.334	0.331
M. nepalensis Kew 25	F	60	0.002	0.002	0.996
M. sinensis 'goliath' Kew 27	A-III	48	0.001	0.001	0.998
M. sinensis 'gracillimus' Kew 28	A-I	76	0.004	0.008	0.988
M. sinensis 'roland' Kew 29	A-I	63	0.002	0.002	0.996
M. sinensis Kew 30	A-I	73	0.004	0.024	0.972
M. sinensis 'gross fontane' Kew 31	A-I	63	0.002	0.002	0.995
M. sacchariflorus Kew 61	B-I	100	0.334	0.335	0.331
M. sinensis 'yakushimanum' Kew 63	A-I	63	0.023	0.973	0.004
M. transmorrisonensis Kew 65	E	65	0.003	0.173	0.824
M. fusca Kew 82	F	94	0.046	0.104	0.849
M. violaceum Kew 84	F	94	0.052	0.756	0.192
M. violaceum Kew 85	F	100	0.334	0.333	0.333
M. ecklonii Kew 86	F	97	0.019	0.598	0.383
M. ecklonii Kew 87	F	100	0.333	0.334	0.332
M. junceum Kew 88	F	97	0.019	0.583	0.398
M. junceum Kew 89	F	94	0.009	0.56	0.43
M. fusca Kew 90	F	97	0.787	0.037	0.176
M. fusca Kew 91	F	100	0.332	0.335	0.332
M. violaceum Kew 92	F	100	0.333	0.335	0.333
M. violaceum Kew 93	F	100	0.333	0.332	0.334
M. capense Kew 94	F	100	0.334	0.334	0.333
M. capense Kew 95	F	100	0.332	0.334	0.333
M. teretifolium Kew 96	F	100	0.333	0.331	0.336
M. junceum Kew 97	F	100	0.334	0.334	0.332
Saccharum officinarum Kew 104	G	100	0.333	0.336	0.331
M. sorghum Kew 105	F	97	0.018	0.586	0.396
M. erectum Kew 106	F	97	0.02	0.59	0.39
M. yunnanensis Kew 107	F	97	0.016	0.027	0.957
M. nudipes Kew 109	F	97	0.065	0.126	0.809
M. nudipes Kew 110	F	100	0.334	0.334	0.332
M. nudipes Kew 111	F	100	0.335	0.335	0.33
M. tinctorius Kew 112	E	100	0.333	0.337	0.33
Saccharum spontaneum Kew 117	G	92	0.308	0.497	0.195
Narenga porphyrocoma Kew 120	G	97	0.233	0.335	0.432
Saccharum contortus Kew 121	G	94	0.013	0.623	0.364
Spodipogon rhizophorus Kew 125	G	98	0.27	0.392	0.338
Spodipogon sibiricus Kew 128	G	97	0.223	0.465	0.311
Eulalia villosa Kew 132	G	100	0.332	0.332	0.335
Eulalia quadrinervis Kew 134	G	97	0.968	0.016	0.016
Eulalia tripsicata Kew 138	G	100	0.332	0.335	0.333
M. sinensis 'morning light' Kew 155	A-I	76	0.005	0.009	0.987

Table 4.4.4 *(continued)*

ID	Groups	%Missing	Inferred clusters		
			I	II	III
M. Sacchariflorus Kew 159	B-I	94	0.687	0.277	0.035
M. sacchariflorus Kew 160	B-I	92	0.498	0.329	0.173
M. tinctorius 'nana variegata' Kew 161	F	81	0.003	0.474	0.523
M. sinensis 'goliath' Kew 194	A-III	73	0.002	0.002	0.996
Cymbopogon citratus	G	88	0.394	0.199	0.407
Pennisetum sp.	G	96	0.117	0.61	0.273
Saccharum officinarum	G	76	0.613	0.339	0.048
Zea diploperennis	G	88	0.946	0.005	0.049
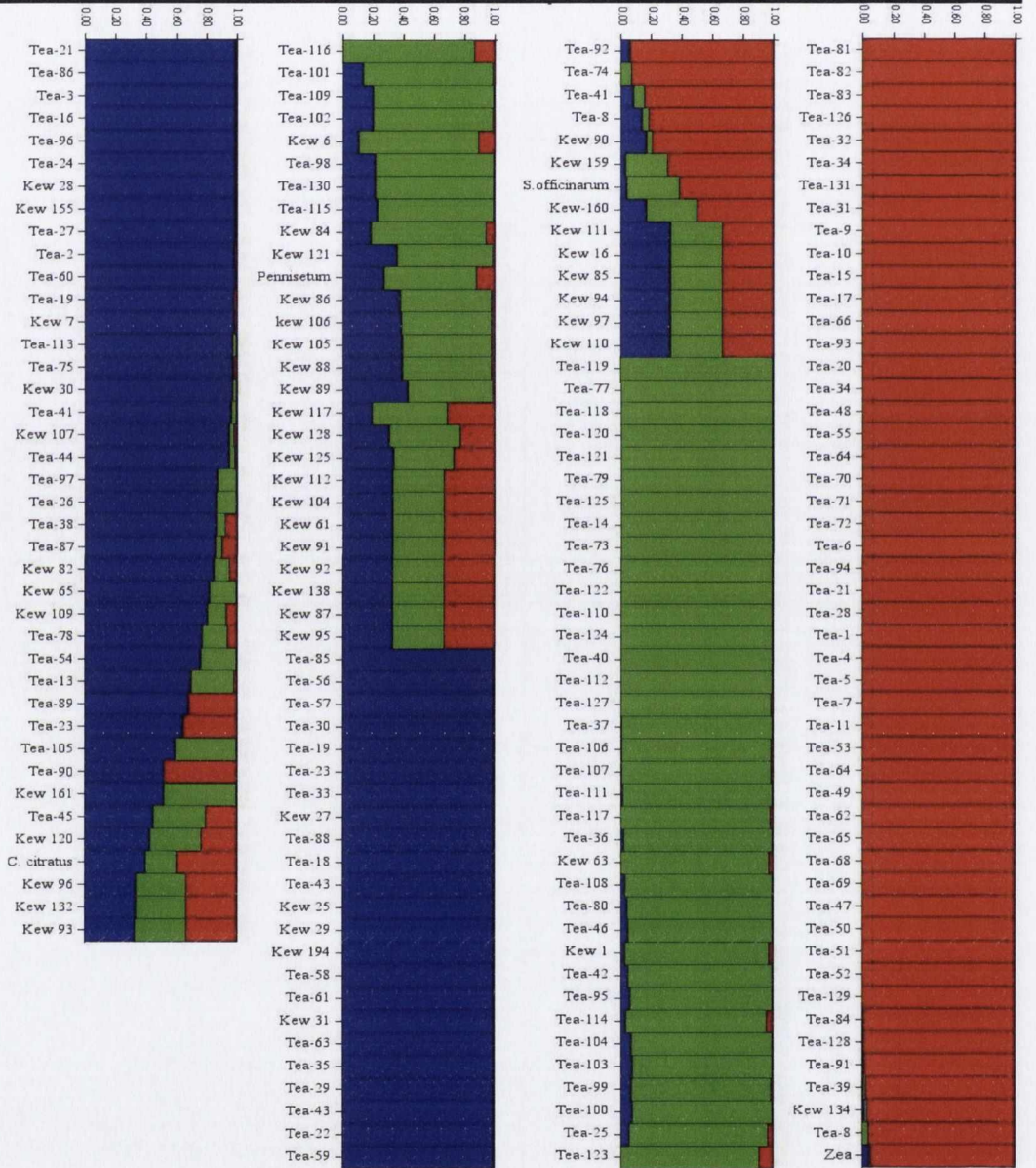
Figure 4.4.4 Structure barplot assigning each accession to cluster I (red), cluster II (green) and cluster III (blue).

110

All individuals belonging to group C fall in cluster I (red) , together with group B-I. Consistently with the UPGMA dendrogram and the PCO, the group B-II has two individuals (Tea-91 and Tea-92) in the same cluster with B-I, whereas the other two individuals belongs to cluster III. Cluster I is completed by the presence of Tea-8 and Tea-39 from group I, *Saccharum officinarum, Zea diploperennis, Eulalia quadrinervis* (Kew 134) and *M. fusca* (Kew 90).

Unlike the UPGMA tree, the cluster analysis groups together both the diploid and the triploid genotypes from Sweden in Cluster II, with few exceptions: Tea-126, already assigned to group C, and four diploid individuals (Tea-96, Tea-97, Tea-105 and Tea-113) that fall in the third cluster. Cluster III includes all the individuals in groups A-II, A-III, *M. transmorrisonensis* (Kew 65) and *M. condensatus* (Kew 7 and Tea-44) from group E. Apart from Tea-62, all other know *M. sinensis* are evenly distributed between Cluster II and III, as well as the *Miscanthus* species in group F, with the only exception of *M. fusca* (Kew 90).

4.4.3 AMOVA

A pairwise comparison between some of the groups revealed by PCO and by the UPGMA tree was carried out. When comparing group C with other groups, the variation among groups accounted for the most of the diversity, from 61% with the closest group B-I to 77% in the comparison with A-III genotypes (Table 4.4.5).

Table 4.4.5 Analysis of molecular variance (AMOVA) between groups as shown by the principal coordinates analysis

Source of variation	df	SS	MS	Est. Var.	%
D vs. A-IV					
Among Pops	1	130.747	130.747	7.626	27%
Within Pops	30	619.472	20.649	20.649	73%
Total	**31**	**750.219**		**28.275**	**100%**
D vs. B-I					
Among Pops	1	58.086	58.086	5.703	29%
Within Pops	15	207.091	13.806	13.806	71%
Total	**16**	**265.176**		**19.509**	**100%**
C vs. D					
Among Pops	1	256.467	256.467	14.224	66%
Within Pops	52	378.533	7.279	7.279	34%
Total	**53**	**635**		**21.504**	**100%**
C vs. B-I					
Among Pops	1	101.089	101.089	9.059	61%
Within Pops	47	267.442	5.69	5.69	39%
Total	**48**	**368.531**		**14.749**	**100%**
C vs. A-III					
Among Pops	1	305.868	305.868	18.512	77%
Within Pops	51	279.642	5.483	5.483	23%
Total	**52**	**585.509**		**23.995**	**100%**
A-III vs. B-I					
Among Pops	1	80.925	80.925	9.76	56%
Within Pops	14	108.2	7.729	7.729	44%
Total	**15**	**189.125**		**17.488**	**100%**

Within groups variation was lower also between Goliath-like individuals when compared with *M. sacchariflorus*. In contrast, the group of Swedish triploid showed higher variation within group when compared with either *M. sacchariflorus* or the diploid ones.

4.5 Discussion

4.5.1 Nuclear molecular markers development

The genetic diversity of 176 individuals, mostly belonging to the species *M. sinensis, M. sacchariflorus* and *M.* ×*giganteus,* and 14 individuals belonging to closely related genera was characterised using 19 newly developed nuclear SSR markers. The loci amplified included a tetranucleotide repetition in nine cases and a dinucleotide repetition in the remaining ten. No bias was observed between di- and tetranucleotide microsatellite in their ability to detect variation (Table 4.4.1).

Despite the presence of triploid and tetraploid plants in the collection of *Miscanthus* analysed, some markers did not show more than two alleles in all individuals genotyped. For markers where more than two alleles were observed, the additional alleles were not present in all tri- and tetraploid accessions, thus not allowing an estimation of the ploidy based on nSSRs.

The presence of more than two alleles per marker for some genotypes made it necessary to convert the data in a presence/absence matrix for further analyses, due to the lack of suitable software that allows the analysis of more than two alleles per locus.

A high level of polymorphism was observed at all loci, with an average allele number of 27.5 per locus (Table 4.4.1) and PIC values of 0.84. A higher level of variation was detected within some species, in particular in *M. sinensis,* compared to other species like *M.* ×*giganteus.*

Average allele number was higher than the value of 12 found by Hernández et al. (2001) in a previous study using SSR from maize. The higher number of clones used in our study (190 against sixteen clones) and the introduction of other *Miscanthus* species other than *M. sinensis, M. sacchariflorus* and *M.* ×*giganteus* could account for the difference in allele number found. However, the average PIC value of 0.836 was consistent with the value of 0.830 in Hernández et al. (2001), both higher than the average PIC value recently found by Zhao et al. (2011) in a study about transferability of 49 microsatellite markers from *Brachypodium distachyon* to *M. sinensis.*

113

In the last few years the first nuclear microsatellite markers for *Miscanthus* have been developed (Hung et al. 2009; Ho et al. 2011; Zhou et al. 2011). Both studies from Zhao et al. (2011) on transferability from *Brachypodium* and Hung et al. (2009) on nine new microsatellite loci specific for *Miscanthus,* were limited to *M. sinensis,* thus explaining the low level of polymorphism (informativeness) found compared to the markers in this study.

Zhou et al. (2011) extended the test for their 14 newly developed markers to *M. floridulus, M. sacchariflorus* and *M. lutarioriparius,* increasing the average number of allele found to 16.1 and the PIC value to 0.757. *Miscanthus lutarioriparius* is a recently described giant *Miscanthus* from China and has until now been understudied genetically (Chen and Renvoize 2006).

A different approach was used by Ho et al. (2011) in the development of 12 new primer pairs, where genic microsatellite loci (EST-SSRs) were obtained through transcriptome sequencing and tested on *M. sinensis* and *M. floridulus,* with a number of alleles of 7.9 on average.

SSRs from *Shaccarum officinarum* ESTs have been recently used by Kim et al. (2012) to generate a genetic map of *M. sacchariflorus* Robustus and *M. sinensis* with a genome coverage of 72.7% and 84.9% respectively. The numbers of linkage groups found for the two maps (40 for *M. sacchariflorus* and 23 for *M.* sinensis) were still higher than the basal chromosome number for *Miscanthus* (x=19), and additional markers will be required to saturate the map, especially from non-coding regions that are underrepresented in the current maps.

The newly developed primers in the study presented here were found to cross amplify not only within *Miscanthus sensu stricto* species but also in other members of the Saccharinae, Andropogoneae and even Paniceae. They amplified DNA in *Zea* (Andropogoneae; Tripsacinae) and *Pennisetum* (Paniceae). The primers are clearly of high value for genetic characterisation and genetic mapping of *Miscanthus* species (Kim et al. 2012) but they could be applied to other closely related genera including *Saccharum* and *Erianthus.*

<u>4.5.2 Genetic characterisation of the germplasm collection</u>

The relationship among *M. sinensis, M. sacchariflorus* and *M. ×giganteus* using molecular markers had been previously investigated by Greef et al. (1997), Hernández et al. (2001) and Hodkinson et al. (2002b). In the first study 48 samples were screened using AFLP fingerprinting. The results of the three studies were controversial. The cluster analysis based on the 17 SSR markers derived from maize (Hernández et al. 2001) indicated a closer clustering of the *M. ×giganteus* clone with the *M. sinensis* cluster than with the *M. sacchariflorus* cluster. This result is in accordance with the origin of the allotriploid *M. ×giganteus* postulated by Linde-Laursen (1993), whose cytogenetic analysis of *M. ×giganteus* revealed the presence of two genomes with high homology to *M. sinensis* and a third with low homology derived from *M. sacchariflorus*. However the two cluster analyses based on AFLP revealed a clear association of *M. ×giganteus* with *M. sacchariflorus* in Greef et al. (1997) but an equal distance between M. ×*giganteus* and both putative parents in Hodkinson et al. (2002b).

Both the UPGMA cluster analysis and the PCO obtained with the 19 SSR markers indicated a cluster of three *M. sacchariflorus* individuals belonging to the MATEREC 11 genotype as the closest to the cluster encompassing all the known *M. ×giganteus* clones. Nevertheless, *M. sacchariflorus* named Kew 159 and Kew 160 showed a higher distance from *M. ×giganteus* clones than *M. sinensis* clones.

Using the model-based clustering method as implemented in STRUCTURE the relationship among the three species is clearer. The entire collection was reduced to three clusters. Both *M. ×giganteus* and *M. sacchariflorus* clones belong to the same cluster, whereas *M. sinensis* individuals are spread between the remaining two clusters, thus confirming the findings of previous studies that used AFLP fingerprinting to evaluate genetic diversity. However, the estimated membership to cluster I for Kew 159 and Kew 160 is lower compared to other *M. sacchariflorus* (0.687 and 0.498 respectively against an average of 0.989 for the MATEREC 11 clones).

The groups revealed by the UPGMA and PCO analyses were compared using AMOVA analysis to account for variation within and between groups. When the *M. ×giganteus* cluster

115

was included in the pairwise comparison, among group variation was found higher than within group variation. A lower level of genetic diversity in *M. ×giganteus* would be expected for a triploid sterile clone that has been mostly propagated vegetatively in European breeding programs. Nevertheless a certain level of variation was found in the species, probably due to multiple origin of *M. ×giganteus* in parts of Asia where the distribution areas of the two putative parents overlap. Similar results were found for the cluster of the triploid variety 'Goliath' of *M. sinensis*. Due to the lack of information about the origin of this clone, it could only be speculated that the triploid genetic set and the consequent sterility of this variety played a role in reducing polymorphisms in the genome.

4.6 Conclusions

Nineteen new nuclear SSR markers were developed starting from a microsatellite enriched library of *M. sinensis* and tested on a collection of *Miscanthus* accessions belonging to 16 different species in the genus as well as on ten species from related genera. The markers proved to be highly polymorphic in *Miscanthus* and transferable to other genera, including *Saccharum*. As part of the study, the genetic diversity in the *Miscanthus* collection established in Teagasc was assessed by UPGMA, PCO and AMOVA, demonstrating a high level of variation among the three species *M. sinensis, M. sacchariflorus* and *M. ×giganteus* and within species. Possible association between the markers and valuable biomass traits should be evaluated in further studies.

Chapter 5

General discussion on the characterisation of genetic and morphological diversity of a collection of *Miscanthus*

5.1 Introduction

The main objectives of this work were to develop new molecular markers for the genus *Miscanthus* and to characterize morphological and molecular diversity in a collection of *Miscanthus* established in Oak Park, Carlow. Such characterisation is essential pre-breeding work necessary to define gene pools, identify taxa, establish inter-relationships of the accessions and develop markers suitable for association studies, quantitative trait loci mapping and marker-aided selection.

The plant material used was from three different sources: Svalöf Weibull (Sweden), University of Hohenheim (Germany) and TCD Botanic Gardens (Ireland), and was mostly composed of individuals of *M. sinensis, M. sacchariflorus* and *M. ×giganteus*, but with some accession not assigned to any of the three species. The morphological characterization was evaluated by mean of scoring important vegetative and reproductive traits and by comparing the results with a set of herbarium specimens belonging to several species of *Miscanthus sensu stricto* (*s.s.*). Molecular diversity of the collection and taxon identity was evaluated using a combination of newly designed DNA markers (chloroplast and nuclear SSRs), sequencing (internal transcribed spacer of the nrDNA) and ploidy estimation through flow cytometry.

5.1.1 Morphological and cytological characterization of a collection of *Miscanthus*

A selection of traits were scored during the second growing season for a newly established collection of *Miscanthus* in Oak Park, Carlow. A high level of variation was found for all the traits, with only a few of them showing a normal distribution in the multi-species dataset. When the same traits were measured in herbarium specimens from nine different species of *Miscanthus,* it was observed that mean values and standard deviation among species varied considerably, thus explaining the non-normal behaviour of a collection with mixed species.

Among all morphological characters, two, both in the inflorescence, are known to be crucial to distinguish *M. sinensis* from *M. sacchariflorus* and *M. ×giganteus*: the presence of an awned lemma in the spikelets and the length of the spikelet callus hairs. *Miscanthus sacchariflorus* and *M. ×giganteus* have long callus hairs (much longer than the length of the spikelet) and lack an awn. *Miscanthus sinensis* is awned and has shorter callus hairs. *Miscanthus sacchariflorus* and M. ×giganteus are difficult to differentiate as both have gigantic cane-like stature, are awnless and have long callus hairs. The attempt to use the data collected in the field for the classification of unidentified accessions was hampered by the unavailability of inflorescences for approximately half of the plants in the collection, and for the plants that did flower, all had *sinensis*-like spikelets except for two accessions. It could be argued that only plants of *M. sinensis* can flower in cold regions (Lewandowski and Clifton Brown 2000). Certainly, only these were able or were mature enough to flower in their second growing season in Carlow. Flowering is influenced by day length and temperature in *Miscanthus* (Lewandowski et al. 2000) and cold late summer and autumn temperature precludes flowering (if plants are transferred in to a sheltered but unheated glasshouse in Ireland they do flower; personal observation). Morphological identification of these accessions as *M. sinensis* is in accordance with the data on their ploidy and the DNA sequences of the ITS region.

The ploidy level estimated in the collection ranged from diploid to tetraploid, with the genotypes almost equally divided between di- and triploid, with a few tetraploid. All the *M. sacchariflorus* were found to be tetraploid, and, as expected, the individuals classified as *M. ×giganteus* were triploid. Diploids were only recorded in *M. sinensis* and *M. condensatus*. Diploid *M. sacchariflorus* are known (Hodkinson et al. 2002c) but were not found in our collection. Two groups were recognised among the triploids: 1) *M. ×giganteus* and some new *M. sacchariflorus× M. sinensis* hybrids and 2) a few individuals of *M. sinensis* 'Goliath'; the latter showing a higher DNA content compared to the other triploids. This higher DNA content was probably due to the different content in haploid sets between the autotriploid M. sinensis 'Goliath' with three *M. sinensis* genomes, and the allotriploid M. ×giganteus, which is likely to have two *M. sinensis* genomes and one from the other putative parent *M. sacchariflorus*, that is known to have a lower DNA content per haploid genome, as reported by Rayburn et al. (2008). In a study on genome size in *Miscanthus*, they estimated the DNA content of a diploid *M. sacchariflorus* in 4.5 pg, around 22% lower than the value of 5.5 pg

found in diploid *M. sinensis*. The genome size of 7.0 pg for the triploid *M. ×giganteus* is in accordance with the presence of two genomes from *M. sinensis* and one from *M. sacchariflorus* (Rayburn et al. 2008).

The comparison between the ploidy levels and the data from the inflorescences showed that all the diploids that produced inflorescences had *sinensis*-like spikelets, while in the triploid group, the accessions identified as *M. sinensis* 'Goliath' had *sinensis*-like spikelets, whereas among the *M. sacchariflorus × M. sinensis* hybrids, some carried *sinensis*-like inflorescences and some the *sacchariflorus*-like ones. None of the plants identified as *M. ×giganteus* or *M. sacchariflorus* did flower.

Where available, the sequencing of the ITS-1 region of the nrDNA confirmed that all the diploid plants had a *sinensis* DNA profile for crucial nucleotide positions that differs between species. This profile was shared with the triploid plants identified as *M. sinensis* 'Goliath', confirming their autotriploid status. All the remaining triploids, both *M. ×giganteus* and *M. sacchariflorus × M. sinensis* hybrids, showed what appeared to be a mixed sequence of the *sacchariflorus* and the *sinensis* DNA profiles. This is not surprising because they are triploid sterile hybrids that have not undergone unequal crossing over and gene conversion (concerted evolution) that would homogenise ITS repeat type (Wendel et al. 1995). The detection of polymorphic sites in the DNA sequence could be useful to support both morphological and ploidy information to define species and distinguish between auto- and allotriploids.

5.1.2 Characterization of genetic diversity using newly developed cpSSRs markers

Starting from the complete sequence of the *Saccharum officinarum* chloroplast genome, a close ally to *Miscanthus,* 30 primer pairs were designed to amplify regions containing SSRs. With one exception, they all proved to be transferable to the genus *Miscanthus*. PCR products for the twelve primer pairs that performed better were sequenced to verify the presence in *Miscanthus* of microsatellite regions and possible polymorphism. Six markers showed length polymorphism of the repeats, with a species-specific preference in alleles.

The six newly developed cpSSR primers were used to genotype the collection of *Miscanthus*. Their cross-amplification was also tested in closely related taxa. The results of the cpSSR genotyping revealed a high number of different haplotypes (85 in 181 accessions tested), but

120

with a clear bias in allele composition between *M. sinensis* and the two species *M. sacchariflorus* and *M. ×giganteus*, thus confirming *M. sacchariflorus* as the maternal lineage of the hybrid *M. ×giganteus*. The newly bred *M. sacchariflorus×M. sinensis* hybrids on the contrary shared their haplotype with *M. sinensis* with the only exception of the genotype Tea-126, whose haplotype is more similar to that of *M. ×giganteus*.

Both the Bayesian analysis with STRUCTURE and the UPGMA tree obtained with the cpSSR data confirmed the presence in the screened population of two clusters, one for individuals with haplotypes typical of *M. sinensis* and one for the *M. sacchariflorus* and *M. ×giganteus* haplotypes. Nevertheless a certain amount of variation was found within clusters, as underlined by the AMOVA performed using the two clusters as populations, which showed that 98% of the variation in the dataset was due to the within population component. When other *Miscanthus* species were considered, it could be observed that other *Miscanthus* s.s. species tend to cluster with *M. sinensis,* whereas the remaining *Miscanthus* s.l. species are more closely related to *M. sacchariflorus*.

The markers reported here are among the first cpSSR and SNP markers developed for *Miscanthus*. These new markers will be useful in *Miscanthus* breeding programmes, for testing maternal inheritance of the chloroplast genome, for population genetic applications and for species differentiation.

5.1.3 Characterization of genetic diversity using newly developed nSSRs markers

New primer pairs for the amplification of nineteen nuclear SSRs loci were developed from the sequences of 192 clones from a microsatellite enriched library. The enrichment of the library was obtained by screening clones for sequences of TC_n, TG_n and $GATA_n$ simple sequence motifs. The newly developed primers were used to characterise the genetic diversity in a collection of *Miscanthus* collection and test their cross-amplification in closely related taxa. All nineteen markers showed high levels of polymorphism with an average number of alleles of 27.5 per locus.

In order to reduce the number of variables accounting for the genetic diversity in the data set, a PCO analysis was performed. The first three axes expressed only 21.3% of variation, but it was possible to identify some groups of accessions. Individuals belonging to *M. ×giganteus*

were clearly separated from the rest of the plants, due to the lower level of polymorphism observed in this species, as expected from a sterile hybrid that can only be propagated only vegetatively. Another group including the 'Goliath' variety of *M. sinensis* could be distinguished from the main core of *M. sinensis* genotypes, as well as a group of *M. sacchariflorus×M. sinensis* hybrids.

The same clustering was observed in the UPGMA tree, where *M. ×giganteus* genotypes cluster together and seemed to be more closely related to a group of *M. sacchariflorus* accessions than to *M. sinensis*. Two accessions of *M. sacchariflorus* appeared to be closer to other *Miscanthus* s.s. species. Among *M. sinensis,* the variety 'Goliath' formed a defined cluster, while the relationship of the other accessions could not be resolved.

When groups defined by the PCO analysis and the UPGMA tree were compared using an analysis of molecular variance (AMOVA), variation among populations was higher than within population variation every time *M. ×giganteus* was included in the calculation, confirming the low level of diversity in this hybrid and the clear separation from other species.

The cluster analysis of the nuclear dataset using STRUCTURE identified three major clusters. One of the clusters included all the *M. ×giganteus,* together with the *M. sacchariflorus* accessions that grouped closer to *M. ×giganteus* in the UPGMA tree, and the non *Miscanthus* individuals that showed cross-amplification with the nSSRs markers. The other two clusters included *M. sinensis* accessions, with both the *M. sinensis* and the *M. sacchariflorus×M. sinensis* hybrids from Sweden in a different cluster from *M. sinensis* 'Goliath'.

A high level of variation within and among species was demonstrated in *Miscanthus*. The newly developed markers will be useful to further explore the diversity of the existing collection and the diversity of newly collected or created accessions. They will be used in association mapping of useful plant breeding traits for biomass production, QTL mapping and MAS.

5.2 Overview of the findings and future work

The morphological, ploidy, sequence and microsatellite results have highlighted the high level of diversity still unexplored in *Miscanthus*. The new molecular tools developed in this study, together with the morphological observation, can be used to establish taxon identity of many accessions in the collection.

Among the *Miscanthus* sp., all the triploids with a flow-cytometry fluorescence ratio comparable with the *M. ×giganteus* standard shared their chloroplast haplotype with *M. ×giganteus* and cluster with *M. ×giganteus* when nSSRs were analysed. The ITS sequence, where available, confirm a *M. ×giganteus* profile for these accessions, suggesting that they probably belong to *M. ×giganteus*. The only exception was Tea-1, which showed a *sinensis*-like chloroplast haplotype.

More complicated is the taxonomical position of the diploid *Miscanthus* sp., due to the high level of variation observed in *M. sinensis*. The genotype Tea-47 was peculiar, since it clustered together with *M. ×giganteus* with both plastid and nuclear SSR markers; however it has a diploid genome. The accessions Tea-16, Tea-24, Tea-26 and Tea-38 shared their chloroplast haplotype with *M. ×giganteus* and *M. sacchariflorus*, but they cluster with *M. sinensis* when nSSRs were taken into account. This might suggest that hybridisation and/or introgression have occurred in the past with these lineages and that they have retained a *sacchariflorus* cpDNA-type (chloroplast capture). The ITS sequence, available for all but Tea-16, showed a *sinensis* profile. The remaining unidentified accessions cluster with *M. sinensis* in all analysis performed. Future work should involve the acquisition of further data from the ITS region and obtaining inflorescences for all accessions (via greenhouse induction) to help clarify the taxonomic status of these genotypes.

Both morphological and molecular characterization highlighted a high level of variation in the genus *Miscanthus,* in particular in *M. sinensis*. Variation was observed also in *M. sacchariflorus,* but the lower number of accessions for this species limited comparisons with *M. ×giganteus*. New genotypes of *M. sacchariflorus* could be analysed with the chloroplast and nuclear markers developed in this study. New accessions are being collected in the wild

in China and Russia by TCD researchers as part of an EU FP7 Grass Margins Project coordinated by Teagasc. It will be interesting to use the new markers on those collections.

One paper has already been published from this thesis on the cpDNA markers (de Cesare et al. 2010) and the following are in preparation:

de Cesare et al. The application of a new set of nuclear SSR markers for pre-breeding and diversity studies in *Miscanthus* (Poaceae). *Theoretical and Applied Genetics*.

de Cesare et al. Genome size and polyploid evolution in the bioenergy grass *Miscanthus*. *Global Change Biology Bioenergy*.

de Cesare et al. Taxon identity and differentiation in *Miscanthus* based on morphology, genome size and nrDNA sequences.

Reference

2007. Minitab 15 Statistical Software. *State College, PA: Minitab, Inc.*

Adati S. 1958. Studies on the genus *Miscanthus* with special reference to the Japanese species suitable for breeding purpose as fodder crops. *Bulletin Faculty Agricultural Mie University* **17**(1-112).

Adati S, Mitsuishi S. 1956. Wild growing forage plants of the Far East, especially Japan, suitable for breeding purposes, part 1, Karyological study in *Miscanthus* (1). *Bulletin of the Faculty of Agriculture Mie University* **12**: 1-14.

Adati S, Shiotani I. 1962. The cytotaxonomy of the genus *Miscanthus* and its phylogenetic status. *Bulletin of the Faculty of Agriculture Mie University* **25**: 1-24.

Agarwal M, Shrivastava N, Padh H. 2008. Advances in molecular marker techniques and their applications in plant sciences. *Plant Cell Reports* **27**: 617-631.

Anderson JA, Maan SS. 1995. Interspecific nuclear–cytoplasmic compatibility controlled by genes on group 1 chromosomes in durum wheat. *Genome* **38**: 803-808.

Andersson NJ. 1855. Om de med *Saccharum* beslägtade genera. *Öfvers Kungl Vet Adad Förh Stockholm* **12**: 151-168.

Arumuganathan K, Earle E. 1991. Estimation of nuclear DNA content of plants by flow cytometry. *Plant Molecular Biology Reporter* **9**: 229-241.

Atienza G, Satovic Z, Petersen K, Dolstra O, Martín A. 2002. Preliminary genetic linkage map of *Miscanthus sinensis* with RAPD markers. *TAG Theoretical and Applied Genetics Theoretische Und Angewandte Genetik* **105**: 946-952.

Baldwin BG, Sanderson MJ, Porter JM, Wojciechowski MF, Campbell CS, Donoghue MJ. 1995. The its Region of Nuclear Ribosomal DNA: A Valuable Source of Evidence on Angiosperm Phylogeny. *Annals of the Missouri Botanical Garden* **82**: 247-277.

Bardakci F. 2001. Random Amplified Polymorphic DNA (RAPD) Markers. *Turkish Journal of Biology* **25**: 185-196.

Botstein D, White RL, Skolnick M, Davis RW. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics* **32**: 314-331.

Brown JWS, Shaw PJ. 1998. Small Nucleolar RNAs and Pre-rRNA Processing in Plants. *The Plant Cell Online* **10**: 649-658.

Chen SL, Renvoize SA. 2006. *Miscanthus*. In *Flora of China*, Vol 22 (ed. RP Wu ZY, Hong DY), pp. 581-583. Science Press, Missouri Botanical Garden Press, Beijing, St. Louis.

Clayton WD, Renvoize SA. 1986. Genera *Graminum*, grasses of the world. In *Kew Bulletin Additional Series XIII*, Springer, New York.

Clifton-Brown J, Chiang YC, Hodkinson TR. 2008. *Miscanthus*: Genetic resources and breeding potential to enhance bioenergy production. In *Vermerris W (ed) Genetic improvement of bioenergy crops*, pp. 295-308. Springer, New York.

Clifton-Brown JC, Lewandowski I. 2000. Overwintering problems of newly established *Miscanthus* plantations can be overcome by identifying genotypes with improved rhizome cold tolerance. *New Phytologist* **148**: 287-294.

Clifton-Brown JC, Lewandowski I. 2002. Screening *Miscanthus* genotypes in field trials to optimise biomass yield and quality in Southern Germany. *European Journal of Agronomy* **16**: 97-110.

de Cesare M, Hodkinson TR, Barth S. 2010. Chloroplast DNA markers (cpSSRs, SNPs) for *Miscanthus*, *Saccharum* and related grasses (Panicoideae, Poaceae). *Molecular Breeding* **26**: 539-544.

Diekmann K, Hodkinson TR, Wolfe KH, van den Bekerom R, Dix PJ, Barth S. 2009. Complete chloroplast genome sequence of a major allogamous forage species,

perennial ryegrass (*Lolium perenne L.*). *DNA Research: An International Journal for Rapid Publication of Reports on Genes and Genomes* **16**: 165-176.

Doležel J, Greilhuber J, Suda J. 2007. Estimation of nuclear DNA content in plants using flow cytometry. *Nature Protocols* **2**: 2233-2244.

Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* **19**: 11-15.

Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**: 2611-2620.

Excoffier L, Smouse PE, Quattro JM. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial dna restriction data. *Genetics* **131**: 479-491.

Flannery ML, Mitchell FJG, Coyne S, Kavanagh TA, Burke JI, Salamin N, Dowding P, Hodkinson TR. 2006. Plastid genome characterisation in *Brassica* and Brassicaceae using a new set of nine SSRs. *Theoretical and Applied Genetics* **113**: 1221-1231.

Głowacka K, Jeżowski S, Kaczmarek Z. 2010. Impact of colchicine application during callus induction and shoot regeneration on micropropagation and polyploidisation rates in two *Miscanthus* species. *In Vitro Cellular & Developmental Biology - Plant* **46**: 161-171.

GPWG. 2001. Phylogeny and Subfamilial Classification of the Grasses (Poaceae). *Annals of the Missouri Botanical Garden* **88**: 373-457.

Grass Phylogeny Working Group II. 2012. New grass phylogeny resolves deep evolutionary relationships and discovers C4 origins. *New Phytologist* **193**:304-312.

Greef JM, Deuter M, Jung C, Schondelmaier J. 1997. Genetic diversity of European *Miscanthus* species revealed by AFLP fingerprinting. *Genetic Resources and Crop Evolution* **44**: 185-195.

Green BR. 2011. Chloroplast genomes of photosynthetic eukaryotes. *The Plant Journal* **66**: 34-44.

Hernández P, Dorado G, Laurie DA, Martín A, Snape JW. 2001. Microsatellites and RFLP probes from maize are efficient sources of molecular markers for the biomass energy crop Miscanthus. *TAG Theoretical and Applied Genetics* **102**: 616-622.

Hirayoshi I, Nishikawa K, Kato R, Kitagawa M. 1955. Cytogenetical forage studies on forage plants (III): chromosome numbers in *Miscanthus*. *Japanese Journal of Breeding* **5**: 49-50.

Ho CW, Wu TH, Hsu TW, Huang JC, Huang CC, Chiang TY. 2011. Development of 12 genic microsatellite loci for a biofuel grass, *Miscanthus sinensis* (Poaceae). *American Journal of Botany* **98**: 201-203.

Hodkinson T, Chase M, Lledò MD, Salamin N, Renvoize S. 2002a. Phylogenetics of *Miscanthus*, *Saccharum* and related genera (Saccharinae, Andropogoneae, Poaceae) based on DNA sequences from ITS nuclear ribosomal DNA and plastid trnL intron and trnL-F intergenic spacers. *Journal of Plant Research* **115**: 381-392.

Hodkinson TR, Chase MW, Renvoize SA. 2002b. Characterization of a genetic resource collection for *Miscanthus* (Saccharinae, Andropogoneae, Poaceae) using AFLP and ISSR PCR. *Annals of Botany* **89**: 627-636.

Hodkinson TR, Chase MW, Takahashi C, Leitch IJ, Bennett MD, Renvoize SA. 2002c. The use of DNA sequencing (ITS and trnL-F), AFLP, and fluorescent in situ hybridization to study allopolyploid *Miscanthus* (Poaceae). *American Journal of Botany* **89**: 279-286.

Hodkinson TR, Ní Chonghaile G, Sungkaew S, Chase MW, Salamin N, Stapleton CMA. 2010. Phylogenetic analyses of plastid and nuclear DNA sequences indicate a rapid late Miocene radiation of the temperate bamboo tribe Arundinarieae (Poaceae, Bambusoideae). *Plant Ecology & Diversity* **3**: 109-120.

Honda M. 1930. Monographia Poacearum Japonicarum, Bambusoides exclusis. *Journal of the Faculty of Science, University of Tokyo, Section III. Botany* **H1**: 484.

Hung K-H, Chiang T-Y, Chiu C-T, Hsu T-W, Ho C-W. 2009. Isolation and characterization of microsatellite loci from a potential biofuel plant: *Miscanthus sinensis* (Poaceae). *Conservation Genetics* **10**: 1377-1380.

James G. 2004. *Sugarcane*. Blackwell, Oxford.

Jensen E, Farrar K, Thomas-Jones S, Hastings A, Donnison I, Clifton-Brown J. 2011. Characterization of flowering time diversity in *Miscanthus* species. *GCB Bioenergy* **3**: 387-400.

Jones MB, Walsh M. 2001. *Miscanthus - for energy and fibre*. James and James (Science Publishers), London.

Kelchner SA. 2000. The evolution of non-coding chloroplast DNA and its application in plant systematics. *Annals of the Missouri Botanical Garden* **87**: 482-498.

Keng YL. 1957. *Claves Generum et Speciarum Graminearum Primarum Sinicarum Appendice Nomenclatione Systematica*. Scientific Book Publishers, Shanghai.

Kim C, Zhang D, Auckland S, Rainville L, Jakob K, Kronmiller B, Sacks E, Deuter M, Paterson A. 2012. SSR-based genetic maps of *Miscanthus sinensis* and *M. sacchariflorus*, and their comparison to sorghum. *TAG Theoretical and Applied Genetics*: 1-14.

Kleine T, Maier UG, Leister D. 2009. DNA transfer from organelles to the nucleus: the idiosyncratic genetics of endosymbiosis. *Annual Review of Plant Biology* **60**: 115-138.

Kumar S, Tamura K, Jakobsen IB, Nei M. 2001. MEGA2: molecular evolutionary genetics analysis software. *Bioinformatics* **17**: 1244-1245.

Lafferty J, Lelley T. 1994. Cytogenetic Studies of Different Miscanthus Species with Potential for Agricultural Use. *Plant Breeding* **113**: 246-249.

Lee YL. 1964. Taxonomic studies on the genus *Miscanthus*. (3). Relationship among the section, subsection and species. *Journal of Japanese Botanic* **38**: 197-205.

Lewandowski I, Clifton-Brown JC, Scurlock JMO, Huisman W. 2000. *Miscanthus*: European experience with a novel energy crop. *Biomass and Bioenergy* **19**(4): 209-227.

Lewandowski I, Clifton-Brown J. 2000. European *Miscanthus* Improvement (FAIR3 CT- 96-1392). Final Report. p. 260.

Linde-Laursen I. 1993. Cytogenetic Analysis of *Miscanthus* 'Giganteus', an Interspecific Hybrid. *Hereditas* **119**: 297-300.

Martin W, Rujan T, Richly E, Hansen A, Cornelsen S, Lins T, Leister D, Stoebe B, Hasegawa M, Penny D. 2002. Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. *Proceedings of the National Academy of Sciences of the United States of America* **99**: 12246-12251.

McCauley DE. 1995. The use of chloroplast DNA polymorphism in studies of gene flow in plants. *Trends in Ecology & Evolution* **10**: 198-202.

McGrath S, Hodkinson TR, Barth S. 2007. Extremely high cytoplasmic diversity in natural and breeding populations of *Lolium* (Poaceae). *Heredity* **99**(5): 531-544.

McGrath S, Hodkinson TR, Salamin N, Barth S. 2006. Development and testing of novel chloroplast microsatellite markers for *Lolium perenne* and other grasses (Poaceae) from de novo sequencing and in silico sequences. *Molecular Ecology Notes* **6**: 449-452.

Mereschkowsky C. 1905. Über Natur und Ursprung der Chromatophoren im Pflanzenreiche. *Biologisches Centralblatt* **25**: 593-604.

Monteith JL. 1978. Reassessment of maximum growth rates for C3 and C4 crops. *Experimental Agriculture* **14**: 1-5.

Naidu SL, Moose SP, AL-Shoaibi AK, Raines CA, Long SP. 2003. Cold Tolerance of C4 photosynthesis in *Miscanthus ×giganteus*: adaptation in amounts and sequence of C4 photosynthetic enzymes. *Plant Physiology* **132**: 1688-1697.

Nielsen PN. 1987. Vegetativ formering af elefantgraes, *Miscanthus sinensis* 'Giganteus' (Vegetative propagation of *Miscanthus sinensis* 'Giganteus'). *Tidsskr Planteaul* **91**: 361-368.

Ohwi J. 1942. Gramina Japonica IV. *Acta Phytotaxonomica Geobotanica* **11**: 145-193.

Ohwi J, Meyer FG, Walker EH. 1965. *Flora of Japan '(in English)' by Jisaburo Ohwi: A combined much rev. and extended transl. by the author of his Flora of Japan ' (1953) ' and Flora of Japan, Pteridophyta ' (1957) '.* Smithsonian Institution.

Ohyama K, Fukuzawa H, Kohchi T, Shirai H, Sano T, Sano S, Umesono K, Shiki Y, Takeuchi M, Chang Z et al. 1986. Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. *Nature* **322**: 572-574.

Olmstead RG, Palmer JD. 1994. Chloroplast DNA systematics: a review of methods and data analysis. *American Journal of Botany* **81**: 1205-1224.

Osada T. 1989. *Illustrated grasses of Japan.*Heibonsha Ltd., Tokyo.

Palmer JD, Soltis DE, Chase MW. 2004. The plant tree of life: an overview and some points of view. *American Journal of Botany* **91**: 1437-1445.

Pavlícek A, Hrdá S, Flegr J. 1999. Free-Tree--freeware program for construction of phylogenetic trees on the basis of distance data and bootstrap/jackknife analysis of the tree robustness. Application in the RAPD analysis of genus Frenkelia. *Folia Biologica* **45**: 97-99.

Peakall R, Smouse PE. 2006. genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6**: 288-295.

Powell W, Machray GC, Provan J. 1996. Polymorphism revealed by simple sequence repeats. *Trends in Plant Science* **1**: 215-222.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155**: 945-959.

Provan J. 2001. Chloroplast microsatellites: new tools for studies in plant ecology and evolution. *Trends in Ecology & Evolution* **16**: 142-147.

Pude R, Franken H, Diepenbrock W, Greef JM. 1997. Ursachen der Auswinterung von einjährigen *Miscanthus*-Beständen. *Pflanzenbauwiss* **4**: 171-176.

Rambaut A. 2007. FigTree, a graphical viewer of phylogenetic trees. Latest version 1.2.2, March 2009.

Ravi V, Khurana JP, Tyagi AK, Khurana P. 2007. An update on chloroplast genomes. *Plant Systematics and Evolution* **271**: 101-122.

Rayburn AL, Crawford J, Rayburn CM, Juvik JA. 2008. Genome size of three *Miscanthus* species. *Plant Molecular Biology Reporter* **27**: 184-188.

Ribaut J-M, Hoisington D. 1998. Marker-assisted selection: new tools and strategies. *Trends in Plant Science* **3**: 236-239.

Ris H, Plaut W. 1962. Ultrastructure of DNA-containing areas in the chloroplast of *Chlamydomonas*. *The Journal of Cell Biology* **13**: 383-391.

Röder MS, Korzun V, Wendehake K, Plaschke J, Tixier MH, Leroy P, Ganal MW. 1998. A microsatellite map of wheat. *Genetics* **149**: 2007-2023.

Röder MS, Plaschke J, König SU, Börner A, Sorrells ME, Tanksley SD, Ganal MW. 1995. Abundance, variability and chromosomal location of microsatellites in wheat. *Molecular & General Genetics: MGG* **246**: 327-333.

Rohlf FJ. 2008. NTSYSpc: Numerical Taxonomy System, ver. 2.20. *Exeter Publishing, Ltd: Setauket, NY.*

Rongwen J, Akkaya MS, Bhagwat AA, Lavi U, Cregan PB. 1995. The use of microsatellite DNA markers for soybean genotype identification. *Theoretical and Applied Genetics* **90**: 43-48.

Shimper AFW. 1885. Untersuchungen uber die Chlorophyllkorner und die ihnen homologen Gebilde. *Jahrb Wiss Bot* **16**: 1-247.

Sun Y, Skinner DZ, Liang GH, Hulbert SH. 1994. Phylogenetic analysis of <i>Sorghum and related taxa using internal transcribed spacers of nuclear ribosomal DNA. *TAG Theoretical and Applied Genetics* **89**(1): 26-32.

Swallen JR. 1961. The Grasses of Burma, Ceylon, India and Pakistan (Excluding Bambuseae). N. L. Bor. Pergamon, New York, 1960. xviii + 767 pp. Illus. $25. *Science* **133**: 1125.

Tanaka M, Wakasugi T, Sugita M, Shinozaki K, Sugiura M. 1986. Genes for the eight ribosomal proteins are clustered on the chloroplast genome of tobacco (*Nicotiana tabacum*): similarity to the S10 and spc operons of *Escherichia coli. Proceedings of the National Academy of Sciences* **83**: 6030-6034.

Takizawa S. 1952. Chromosome constitution of the genus *Miscanthus* Anderss. *La Kromosoma* **14**: 509-517.

Tautz D, Renz M. 1984. Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acids Research* **12**: 4127-4138.

Temnykh S, Park WD, Ayres N, Cartinhour S, Hauck N, Lipovich L, Cho YG, Ishii T, McCouch SR. 2000. Mapping and genome organization of microsatellite sequences in rice (*Oryza sativa L.*). *TAG Theoretical and Applied Genetics* **100**: 697-712.

Thomas MR, Scott NS. 1993. Microsatellite repeats in grapevine reveal DNA polymorphisms when analysed as sequence-tagged sites (STSs). *Theoretical and Applied Genetics* **86**: 985-990.

Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**: 4673-4680.

Vermerris W. 2008. *Genetic Improvement of Bioenergy Crops*. Springer Science+Business Media, LLC. pp. 467.

Visser P, Pignatelli V. 2001. Utilisation of *Miscanthus*. In *Miscanthus for energy and fibre*, (ed. MB Jones, M Walsh), p. 192. James & James, London.

Vos PR, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Kuiper M, Zabeau M. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acid Research* **23**: 4407-4414.

Wendel JF, Schnabel A, Seelanan T. 1995. Bidirectional interlocus concerted evolution following allopolyploid speciation in cotton (*Gossypium*). *Proceedings of the National Academy of Science* **92**: 280-284.

White TJ, Bruns S.L., J. T. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *MA INNIS, DH GELFAND, JJ SNINSKY and TJ WHITE [eds] PCR protocols a guide to methods and applications, Academic Press, San Diego*

Williams JG, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research* **18**: 6531-6535.

Wilson EB. 1928. *The Cell in Development and Heredity, 3rd revised edtion.* Reprinted (1987) by Garland Publishing, New York.

Yan J, Chen W, Luo FAN, Ma H, Meng A, Li X, Zhu M, Li S, Zhou H, Zhu W et al. 2012. Variability and adaptability of *Miscanthus* species evaluated for energy crop domestication. *GCB Bioenergy* **4**: 49-60.

Yin X, Stam P, Dourleijn CJ, Kropff MJ. 1999. AFLP mapping of quantitative trait loci for yield-determining physiological characters in spring barley. *TAG Theoretical and Applied Genetics* **99**: 244-253.

Zhao H, Yu J, You FM, Luo M, Peng J. 2011. Transferability of Microsatellite Markers from Brachypodium distachyon to *Miscanthus sinensis*, a Potential Biomass Crop F. *Journal of Integrative Plant Biology* **53**: 232-245.

Zhou H-F, Li S-S, Ge S. 2011. Development of microsatellite markers for Miscanthus sinensis (Poaceae) and cross-amplification in other related species. *American Journal of Botany* **98**: 195-197.

Appendices

Table A List of accessions used in this study. Genotypes highlighted in grey were part of a field trial in Teagasc, Oak Park, Carlow.

ID	Acquired from	Accession	Notes
Miscanthus sp. Tea-1	TCD Bot.Gardens	TCDP15	M. sacchariflorus
M. sinensis 'zebrinus' Tea-2	TCD Bot.Gardens	TCDP20	M. sinensis-Zebrinus
M. sinensis 'zebrinus' Tea-3	TCD Bot.Gardens	TCDP21	M. sinensis-Zebrinus
M. xgiganteus Tea-4	TCD Bot.Gardens	TCDP34	M. giganteus
M. xgiganteus Tea-5	TCD Bot.Gardens	TCDP36	M. giganteus
Miscanthus sp. Tea-6	TCD Bot.Gardens		
Miscanthus sp. Tea-7	TCD Bot.Gardens	TCDP48	M. sinensis
Miscanthus sp. Tea-8	TCD Bot.Gardens	TCDP50	
Miscanthus sp. Tea-9	TCD Bot.Gardens	TCDP51	M. sinensis
Miscanthus sp. Tea-10	TCD Bot.Gardens	TCDP58	M. sacchariflorus
Miscanthus sp. Tea-11	TCD Bot.Gardens		
M. sinensis Tea-13	TCD Bot.Gardens	TCDP73	M. sinensis
M. sinensis Tea-14	TCD Bot.Gardens	TCDP75	M. sinensis
Miscanthus sp. Tea-15	TCD Bot.Gardens	TCDP104	M. sp
Miscanthus sp. Tea-16	TCD Bot.Gardens	TCDP105	M.transmorrisonensis
M. xgiganteus Tea-17	TCD Bot.Gardens	TCDP108	M. giganteus
Miscanthus sp. Tea-18	TCD Bot.Gardens		
M. sinensis goliath-like Tea-19	TCD Bot.Gardens	TCDP110	SIN-H6
M. xgiganteus Tea-20	TCD Bot.Gardens	TCDP114	M. giganteus
Miscanthus sp. Tea-21	TCD Bot.Gardens	Unknown	
Miscanthus sp. Tea-22	TCD Bot.Gardens	Unknown	
Miscanthus sp. Tea-23	TCD Bot.Gardens	Unknown	
Miscanthus sp. Tea-24	TCD Bot.Gardens	TCDP11	M. sinensis
Miscanthus sp. Tea-25	TCD Bot.Gardens	TCDP11	M. sinensis
Miscanthus sp. Tea-26	TCD Bot.Gardens	TCDP11	M. sinensis
Miscanthus sp. Tea-27	TCD Bot.Gardens	Unknown	
Miscanthus sp. Tea-28	TCD Bot.Gardens	Unknown	
Miscanthus sp. Tea-29	TCD Bot.Gardens	Unknown	
M. sinensis Tea-30	TCD Bot.Gardens		M. sinensis
M. xgiganteus Tea-31	TCD Bot.Gardens		M. giganteus
Miscanthus sp. Tea-32	TCD Bot.Gardens	Unknown	M. giganteus
M. sinensis 'zebrinus' Tea-33	TCD Bot.Gardens	TCDP20 (?)	M. sinensis-Zebrinus
Miscanthus sp. Tea-34	TCD Bot.Gardens	Unknown	
M. sinensis 'gross fontane' Tea-35	TCD Bot.Gardens	TCDP30	M. sinensis-Grosse Fontane
M. sinensis 'gross fontane' Tea-36	TCD Bot.Gardens		M. sinensis-Grosse Fontane

Table A (*continued*)

ID	Acquired from	Accession	Notes
Miscanthus sp.Tea-37	TCD Bot.Gardens	Unknown	
Miscanthus sp. Tea-38	TCD Bot.Gardens	Unknown	
Miscanthus sp.Tea-39	TCD Bot.Gardens	Unknown	
M. sinensis Tea-40	TCD Bot.Gardens	TCDP62	M. sinensis
Miscanthus sp. Tea-42	TCD Bot.Gardens	Unknown	
Miscanthus sp. Tea-43	TCD Bot.Gardens	Unknown	
M. condensatus Tea-44	TCD Bot.Gardens	TCDP94	M. condensatus
Miscanthus sp. Tea-45	TCD Bot.Gardens		
Miscanthus sp. Tea-46	TCD Bot.Gardens		
Miscanthus sp. Tea-47	TCD Bot.Gardens		
Miscanthus sp. Tea-48	TCD Bot.Gardens		
Miscanthus sp. Tea-49	TCD Bot.Gardens		
Miscanthus sp. Tea-50	TCD Bot.Gardens		
Miscanthus sp. Tea-51	TCD Bot.Gardens		
Miscanthus sp. Tea-52	TCD Bot.Gardens		
Miscanthus sp. Tea-53	TCD Bot.Gardens		
Miscanthus sp.Tea-54	TCD Bot.Gardens		
Miscanthus sp.Tea-55	TCD Bot.Gardens		
M. sinensis 'goliath'Tea-56	Teagasc Oak Park field		
M. sinensis 'goliath' Tea-57	TCD Bot.Gardens		
M. sinensis 'sirene' Tea-58	Teagasc Oak Park field		
M. sinensis 'strictus' Tea-59	TRH garden		
M. sinensis 'strictus' Tea-60	TCD Bot.Gardens		
M. sinensis 'malaparteus' Tea-61	TRH Garden		
M. sinensis Tea-62	TRH Garden		
M. sinensis 'sirene' Tea-63	TCD Bot.Gardens		
M. xgiganteus Tea-64	TCD Bot.Gardens		
M. xgiganteus Tea-65	TCD Bot.Gardens		
M. xgiganteus Tea-66	TRH Garden, used for naming species		
Miscanthus sp. Tea-68	TCD Bot.Gardens		
Miscanthus sp. Tea-69	TCD Bot.Gardens		
Miscanthus sp. Tea-70	TCD Bot.Gardens		
Miscanthus sp. Tea-71	TCD Bot.Gardens		
Miscanthus sp. Tea-72	TCD Bot.Gardens		
Miscanthus sp. Tea-73	TCD Bot.Gardens		
M. xgiganteus Tea-74	Germany - from Denmark	M1 Lasei 1	M. giganteus
M. sacchariflorus x M. sinensis Tea-75	Germany	M81 RH 81	M.sacchariflorusxM.sinensis
M. sinensis Tea-76	Germany - From Japan	88-110	M. sinensis
M. sinensis Tea-77	Germany - From Japan	88-111	M. sinensis
M. sinensis Tea-78	Germany - From Japan	90-5	M. sinensis

Table A (*continued*)

ID	Acquired from	Accession	Notes
M. sinensis Tea-79	Germany - From Japan	90-6	M. sinensis
M. sinensis Tea-80	Germany - From Sweden	SW 217	M. sinensis
M. xgiganteus Tea-81	Germany - from Denmark	M53 IPL 53	M. giganteus
M. xgiganteus Tea-82	Germany	M56 HAGA 56	M. giganteus
M. xgiganteus Tea-83	Germany	M63 GREIF 63	M. giganteus
M. sacchariflorus Tea-84	Germany - from Japan	M11 MATEREC 11	M. sacchariflorus
M. sinensis goliath-like Tea-85	Germany	M7 GOFAL 7	Goliath-like M. sinensis Hybrid
M. sinensis Tea-86	Germany	M42 BERBO 42	M. sinensis Hybrid
M. sacchariflorus x M. sinensis Tea-87	Germany	M43RH43	M.sacchariflorusx M.sinensis
M. sinensis Tea-88	Germany	M78 JESEL 78	M. sinensis Hybrid
Miscanthus sp. Tea-89	Oak Park		
Miscanthus sp. Tea-90	Oak Park		
Miscanthus sp. Tea-91	Oak Park		
Miscanthus sp. Tea-92	Oak Park		
M. xgiganteus Tea-93	IGER/JCB TinPlant trial Oak Park, Discard plot		
M. xgiganteus Tea-94	Old Trial - Oak Park		
M. sinensis Tea-95	Sweden		
M. sinensis Tea-96	Sweden		
M. sinensis Tea-97	Sweden		
M. sinensis Tea-98	Sweden		
M. sinensis Tea-99	Sweden		
M. sinensis Tea-100	Sweden		
M. sinensis Tea-101	Sweden		
M. sinensis Tea-102	Sweden		
M. sinensis Tea-103	Sweden		
M. sinensis Tea-104	Sweden		
M. sinensis Tea-105	Sweden		
M. sinensis Tea-106	Sweden		
M. sinensis Tea-107	Sweden		
M. sinensis Tea-108	Sweden		
M. sinensis Tea-109	Sweden		
M. sinensis Tea-110	Sweden		
M. sinensis Tea-111	Sweden		
M. sinensis Tea-112	Sweden		
M. sinensis Tea-113	Sweden		
M. sinensis Tea-114	Sweden		
M. sinensis Tea-115	Sweden		
M. sacchariflorus x M. sinensis Tea-116	Sweden		
M. sacchariflorus x M. sinensis Tea-117	Sweden		
M. sacchariflorus x M. sinensis Tea-118	Sweden		
M. sacchariflorus x M. sinensis Tea-119	Sweden		
M. sacchariflorus x M. sinensis Tea-120	Sweden		

Table A (*continued*)

ID	Acquired from	Accession	Notes
M. sacchariflorus x M. sinensis Tea-121	Sweden		
M. sacchariflorus x M. sinensis Tea-122	Sweden		
M. sacchariflorus x M. sinensis Tea-123	Sweden		
M. sacchariflorus x M. sinensis Tea-124	Sweden		
M. sacchariflorus x M. sinensis Tea-125	Sweden		
M. sacchariflorus x M. sinensis Tea-126	Sweden		
M. sacchariflorus x M. sinensis Tea-127	Sweden		
M. sacchariflorus Tea-128	TCD Bot.Gardens		
M. sacchariflorus Tea-129	TCD Bot.Gardens		
Miscanthus sp. Tea-130	TCD Bot.Gardens	Unknown	
Miscanthus sp. Tea-131	TCD Bot.Gardens	Unknown	
Cymbopogon citratus	TCD Bot.Gardens		
Pennisetum sp.	TCD Bot.Gardens		
Saccharum officinarum	TCD Bot.Gardens		
Zea diploperennis	TCD Bot.Gardens		
M. sinensis 'variegatus' Kew 1	RBG Kew 154 04	1969-19093	M. sinensis var. variegatus
Sorghum alpense Kew 6	RBG Kew 151 01	1966-54209	Sorghum halpense
M. condensatus Kew 7	RBG Kew 151	1969-19091	M. condensatus
M. oligostachyus Kew 16	RBG Kew 151 (pot)	1995-1864	M. oligostachyus
M. nepalensis Kew 25	RBG Kew TH 4	1985-8388	M. nepalensis
M. sinensis 'goliath' Kew 27	ADAS Steinmann nurseries	MB93/02	M. sinensis 'Goliath'
M. sinensis 'gracillimus' Kew 28	ADAS Piccoplants, Germany	MB94/05	M. sinensis 'Gracillimus'
M. sinensis 'roland' Kew 29	ADAS Piccoplants, Germany	MB94/06	M. sinensis 'Roland'
M. sinensis Kew 30	ADAS Wye College	MB94/07	M. chinensis (Sinensis)
M. sinensis 'gross fontane' Kew 31	ADAS Genft Dogels, Germany	PN95/01	M. sinensis 'Grobe rontane'
M. sacchariflorus Kew 61	RBG Kew 751 MB	1987-2727	M. purpurascens=M.sacchariflorus
M. sinensis 'yakushimanum' Kew 63	RBG Kew 751	1987-1148	M. sinensis 'yakushimanum'
M. transmorrisonensis Kew 65	RBG Kew 732	1990-2748	M. sinensis 'transmorrisonensis'
M. fusca Kew 82	RBG Kew Herbarium	590	M. fusca
M. violaceum Kew 84	RBG Kew Herbarium	7437	M. violaceus
M. violaceum Kew 85	RBG Kew Herbarium	468	M. violaceus
M. ecklonii Kew 86	RBG Kew Herbarium	2347	M. ecklonii
M. ecklonii Kew 87	RBG Kew Herbarium	3228	M. ecklonii
M. junceum Kew 88	RBG Kew Herbarium	1060	M. junceum
M. junceum Kew 89	RBG Kew Herbarium	2309	M. junceum
M. fusca Kew 91	RBG Kew Herbarium	US 56-5-5b	M. fusca
M. violaceum Kew 92	RBG Kew Herbarium	7437b	M. violaceum
M. violaceum Kew 93	RBG Kew Herbarium	468b	M. violaceum
M. capense Kew 94	RBG Kew Herbarium	2347b	M. capense
M. capense Kew 95	RBG Kew Herbarium	3228b	M. capense
M. teretifolium Kew 96	RBG Kew Herbarium	1060b	M. teretifolium
M. junceum Kew 97	RBG Kew Herbarium	2309b	M. junceum

Table A (*continued*)

ID	Acquired from	Accession	Notes
M. junceum Kew 97	RBG Kew Herbarium	2309b	M. junceum
Saccharum officinarum Kew 104	Palm House 4	1973-12242	Saccharum officinarum
M. sorghum Kew 105	Herbarium, RGG, Kew	2929	M. sorghum
M. erectum Kew 106	Herbarium, RBG, Kew	247	M. erectum
M. yunnanensis Kew 107	Herbarium, RBG, Kew	30689	M. yunnanensis
M. nudipes Kew 109	Herbarium, RBG, Kew	2007	M. nudipes
M. nudipes Kew 110	Herbarium, RBG, Kew	2517	M. nudipes
M. nudipes Kew 111	Herbarium, RBG, Kew	522	M. nudipes
M. tinctorius Kew 112	Herbarium, RBG, Kew	1466	M. tinctorius
Saccharum spontaneum Kew 117	Herbarium, RBG, Kew	Butt, 1977	Saccharum spontaneum
Narenga porphyrocoma Kew 120	Herbarium, RBG, Kew	2092	Narenga porphyrocoma
Saccharum contortus Kew 121	Herbarium, RBG, Kew	3797	S. contortus (E. contortus)
Spodipogon rhizophorus Kew 125	Herbarium, RBG, Kew	283	Spodiopogon rhizophorus
Spodipogon sibiricus Kew 128	Herbarium, RBG, Kew	210	Spodiopogon sibricus
Eulalia villosa Kew 132	Herbarium, RBG, Kew	1882	Eulalia villosa
Eulalia quadrinervis Kew 134	Herbarium, RBG, Kew	3294	Eulalia quadrinervis
Eulalia tripsicata Kew 138	Herbarium, RBG, Kew	10062	Eulalia tripsicata
M. sinensis 'morning light' Kew 155	RBGKew Living	1996 821	M. sinensis 'Morning Light'
M. Sacchariflorus Kew 159	Herbarium, RBG, Kew	3598 1935	M. sacchariflorus
M. sacchariflorus Kew 160	Herbarium, RBG, Kew	1984	M. sacchariflorus Japan
M. tinctorius 'nana variegata' Kew 161	RBGKew Living	1996 1065	M. tinctorius 'Nana variegatus'
M. sinensis 'goliath' Kew 194	ADAS	PN96/30	M. 'goliath'

Table B Characters and relative scores for herbarium specimens of *Miscanthus*.

Specimens	Location	Culm Height	Culm width	Leaf blade length	Leaf blade width	Leaf hair	Inflorescence length	Inflorescence axis length	Raceme number	Raceme length	Peduncle hairs	Axis hairs	Raceme axis hairs	Raceme internode	Upper pedicel length	Lower pedicel	Pedicel hairs	Spikelet length	Awn length	Spikelet hairs length
M. purpurescence E00257696	Corea (Edinburgh)	85	0.6	57	1.3	N	15	0	8	12	N	N	N	1	0	0	N	0.6	0.3	0.7
M. sinensis 32656(04/33)	Borneo (Harvard)	*80	0.5	64	2.3	N	30	12.5	50	15	N	N	N	0.8	0.3	0.5	N	0.5	0.5	0.5
M.sinensis var. transmorrisonensis E00257699	Yilan- Taiwan (Edinburgh)	48	0.2	27	0.4	N	13	0	8	8	N	N	N	0.6	0.2	0.6	N	0.6	0.7	0.5
M.sinensis var. formosanus E00129454	Taiwan (Edinburgh)	*100	0.4	*46	0.8	N	24	4	25	16	N	N	N	0.5	0.1	0.2	N	0.4	0.7	0.5
M.sinensis E00257711	Tonkin (Edinburgh)	*112	0.5	50	0.8	N	28	14.5	30	16.5	N	N	N	0.8	0.1	0.2	N	0.5	0.5	0.6
M. japonicus No. 490	Tonkin (Kew)	116	0.8	59	0.6	Y	24	5.5	20	13.5	N	N	N	0.5	0.1	0.2	N	0.3	0.4	0.5
M. sinensis And. 000938881	China-Mt.Bani (BM)	*71	0.7	34	0.5	N	47	33	30	16	N	N	N	0.5	0.15	0.4	N	0.5	0.6	0.4
M.sinensis 000938880	China-Mt.Bani (BM)	*60	0.4		0.6	N	50	30	30	17	N	N	N	0.8	0.2	0.6	N	0.5	0.4	0.5
M.sinensis 65/2008-37	(Copenhagen)	0	0.5	75	1	N	*34	*18	25	18	N	N	N	0.9	0.2	0.4	N	0.6	0.5	0.5
M.sinensis 65/2008-38	Indochina (Copenhagen)	*128	0.4	60	1	N	31	17	50	17	N	N	N	0.5	0.1	0.4	N	0.4	0.5	0.5
M.sacchariflorus E00129472	Tsa sie lou (Edinburgh)	40	0.3	26	1.5	Y	24	7	30	18	N	Y	N	0.9	0.2	0.4	N	0.5		0.8
M.sacchariflorus E00129474	China (Edinburgh)		0.2	39	1.6	Y	24	11.5	38	16.5				0.5	0.2	0.4		0.4		0.9
M.sacchariflorus E00129470	(Edinburgh)		0.5	57	0.8	Y	30	5	50	18				0.4	0.3	0.6		0.4		1.3
M.sacchariflorus e00129473	Houshu Japan (edinburgh)	163	0.5	53	2	Y	26	5.8	30	13				0.5	0.2	0.6		0.3		0.9

Table B (*continued*)

Specimens	Location	Culm Height	Culm width	Leaf blade length	Leaf blade width	Leaf hair	Inflorescence length	Inflorescence axis length	Raceme number	Raceme length	Peduncle hairs	Axis hairs	Raceme axis hairs	Raceme internode	Upper pedicel length	Lower pedicel	Pedicel hairs	Spikelet length	Awn length	Spikelet hairs length
M.sacchariflorus 000938877	P. Sinensis (BM)	75	0.3	24	0.4	N	20	2	6	11	N	N	N	0.6	0.2	0.4		0.5		0.6
M.sacchariflorus 000938876	P. Sinensis (BM)	88	0.2	32	0.6	N	20	7	20	9	N	N	N	0.5	0.1	0.3		0.4		0.7
M.sacchariflorus 000938878	P. Sinensis (BM)	92	0.2	52	0.7	N	16	11	11	12		N		0.7	0.15	0.5		0.4		1.6
M.sacchariflorus 65/2008-39	China (Copenhagen)	119	0.4	54	1.4	N	25	5	25	12	N	N		0.5	0.2	0.4		0.5		0.7
M.sacchariflorus 65/2008-35	USA (Copenhagen)	144	0.4	37	1	N	16	5	8	14	N	N	N	0.7	0.2	0.6		0.4		1.3
M.sacchariflorus 65/2008-36	sinense (Copenhagen)	58	0.2		0.3	N	18	9	10	16	N	N	N	0.5	0.2	0.4		0.5		1
M. floridulus E00257684	Philippines (Edinburgh)	50	0.9	70	2.6	N	33		50	19				0.4	0.2	0.3		0.4	0.6	0.5
M. japonicus E00257683	Philippines (Edinburgh)	70	0.4	36	1.1	N	35		20	19				0.4	0.2	0.3		0.3	0.7	0.3
M. japonicus E00257682	Philippines (Edinburgh)		0.4	41	0.8	N	30	9.5	20	14				0.5	0.1	0.3		0.4	0.5	0.4
M. japonicus E00257681	(Edinburgh)	66	0.5	41	0.8	N	36	20	50	12				0.5	0.1	0.3		0.5	0.7	0.6
M. floridulus E00257680	(Edinburgh)		0.7	20	0.8	N	41	22	100	23				0.5	0.1	0.3		0.3	0.5	0.7
M. floridulus E00257679	Guinea(Edinburgh)	67	0.2	32	0.3	N	18	6	20	7	Y			0.7	0.2	0.5		0.6	0.8	0.4
M. floridulus E00257678	Papua (Edinburgh)	75	0.5	20	0.8	N	23	12	50	8				0.5	0.2	0.5		0.6	0.8	0.4
M. japonicus E00257702	Japan (Edinburgh)	56	0.7	71	2.7	N	40	34	100	20				0.5	0.1	0.3		0.3	0.6	0.4

Table B (*continued*)

Specimens	Location	Culm Height	Culm width	Leaf blade length	Leaf blade width	Leaf hair	Inflorescence length	Inflorescence axis length	Raceme number	Raceme length	Peduncle hairs	Axis hairs	Raceme axis hairs	Raceme internode	Upper pedicel length	Lower pedicel	Pedicel hairs	Spikelet length	Awn length	Spikelet hairs length
M. sinensis H2008/00313-98	Kwangtung-tonkin (Kew)	79	0.4	47	0.8	N	23	11	20	13				0.7	0.2	0.4		0.3	0.4	0.5
E. Floridulus 1509	N. Caledonia (TCD)	62	0.3	20	0.8	N	35	23	60	17				0.4	0.2	0.4		0.4	0.9	0.6
E. Floridulus TCD8	N. Caledonia (TCD)	48	0.4	40	1.1	N	17	11	30	12				0.3	0.1	0.2		0.3	0.4	0.3
M. condensatus E00257708	Japan (Edinburgh)	44	0.7	41	1.2	N	21	7	25	18				0.2	0.1	0.2		0.4	1	0.4
M. condensatus E00129460	Japan (Edinburgh)	76	0.6	50	2	N	24	8	20	13	N	N		0.4	0.2	0.4		0.6	1.2	0.5
M. condensatus H2008/00313-104	Japan (Kew)	80	0.5	42	1.5	N	21	10.5	25	16	N	N		0.4	0.2	0.3		0.5	0.7	0.5
M. condensatus H2008/00313-103	Japan (Kew)	35	0.7	59	2.2	N														
M. condensatus H2008/00313-102	Japan (Kew)	81	0.5	45	1.5	N	35	7	30	18				0.3	0.1	0.4		0.5	1	0.6
M. condensatus H2008/00313-101	Japan (Kew)		0.7	47	1.2	Y														
M. condensatus H2008/00313-100	Japan (Kew)						31		50	18	N	N		0.4	0.2	0.5		0.5	0.7	0.5
M. condensatus H2008/00313-99	Japan (Kew)	132	1.7	67	3	N	38	20	80	15	N	N		0.2	0.15	0.3		0.4	0.8	0.4
M. condensatus 65/2008-29	Japan (Copenhagen)		0.8		1.4	N	20		25	16	N	N		0.2	0.1	0.2		0.5	1.1	0.5
M. condensatus 65/2008-30	Corea (Copenhagen)	105	0.9	32	1.3	N	18	9.5	35	8	Y	Y		0.3	0.1	0.3		0.6	0.5	0.5
M. JAPONICUS 65/2008-31	(Copenhagen)	183	0.5	87	0.6	Y	33	22.5	60	11	Y	Y		0.5	0.2	0.5		0.4	0.7	0.6

Table B (*continued*)

Specimens	Location	Culm Height	Culm width	Leaf blade length	Leaf blade width	Leaf hair	Inflorescence length	Inflorescence axis length	Raceme number	Raceme length	Peduncle hairs	Axis hairs	Raceme axis hairs	Raceme internode	Upper pedicel length	Lower pedicel	Pedicel hairs	Spikelet length	Awn length	Spikelet hairs length
M. oligostachyus E00257703	Honshu (Edinburgh)	72	0.2	19.5	0.8	Y	5	1	3	5	N	N	N	1	0.3	0.8		0.8	0.9	0.5
M. oligostachyus E00257704	Japan (Edinburgh)	92	0.2	23	1.2	Y	11	2.1	4	11	N	N	N	1.4	0.3	0.8		1	1	0.6
M. oligostachyus E00257705	Japan (Edinburgh)	74	0.2	40	1.2	Y	18	2.5	3	11	N	N	N	1.1	0.3	0.8		0.7	1	0.5
M. oligostachyus H2008/00313-92	Honshu (Kew)	72	0.2	30	1.2	N	16	5.5	4	10	N	N	N	0.9	0.2	0.7		0.9	1	0.5
M. oligostachyus H2008/00313-91	Honshu (Kew)	37	0.1	12	0.6	Y	11	1	2	9				0.6	0.2	0.5		0.8	1.1	0.4
M. oligostachyus H2008/00313-88	Japan (Kew)	122	0.2	30	1	N	9		3	8	N	N		0.7	0.1	0.5		0.6	0.5	0.3
M. oligostachyus H2008/00313-89	Japan (Kew)	72	0.2	27	1.2	Y	11	1	2	11	N	N		1	0.2	0.6		1	0.6	0.5
M. oligostachyus H2008/00313-90	Japan (Kew)	43	0.2	20	1	Y	12		1	12	N	N	N	0.8	0.1	0.5		0.9	0.9	0.6
M. oligostachyus 65/2008-33	Japan (Copenhagen)	59	0.2	23	1.1	Y	14	1.5	3	13	N	N		1	0.2	0.8		1	1.3	0.5
M. tinctorius E00257706	Japan (Edinburgh)	74	0.4	33	1.4		23	1.5	7	18			N	0.6	-0.1	0.3		0.6		0.3
M. tinctorius E00257707	Honshu (Edinburgh)	137.5	0.3	46	1.7	N	18.5	1	11	13.5	N	N		0.5	0.2	0.2		0.5		0.2
M. tinctorius 65/2008-39	Japan (Copenhagen)	155	0.3	48	1	N	16	1.5	8	14	N	N	N	0.6	0	0.4		0.6	0.9	0.3
M. tinctorius 65/2008-40	Japan (Copenhagen)	104	0.2	32	0.8		14		2	14	Y	Y		0.7	0.1	0.5		0.6		0.4
M. sorghum H2008/00313 96	South Rhodesia (Kew)	120	0.5	52	0.6		38	21	60	11				0.4	0.15	0.25		0.4	0.3	0.3

Table B (*continued*)

Specimens	Location	Culm Height	Culm width	Leaf blade length	Leaf blade width	Leaf hair	Inflorescence length	Inflorescence axis length	Raceme number	Raceme length	Peduncle hairs	Axis hairs	Raceme axis hairs	Raceme internode	Upper pedicel length	Lower pedicel	Pedicel hairs	Spikelet length	Awn length	Spikelet hairs length
M. sorghum H2008/00313 95	Rhodesia (Kew)	188	0.5	60			57	43	70	7				0.8	0.3	0.6		0.4	0.4	0.2
M. sorghum H2008/00313 94	Victoria-Rhodesia (Kew)	92	0.4	74	0.3		45	35	70	6				0.6	0.25	0.6		0.4	0.3	0.3
M.yunnanensis E00257700	2800m (Edinburgh)	54	0.3	30	0.4	Y	16	4	15	9				0.7	0.3	0.6		0.5	0.8	0.4
M.yunnanensis 000085221	Plantae sinensis (BM)	103	0.2	40	0.5	Y	9	2	6	6				0.6	0.2	0.3		0.4	0.5	0.5
M. nudipes E00257709	cameroun (Edinburgh)	96	0.2	22	0.6	Y	13	5	25	6	N	N	N	0.3	0.2	0.4		0.2	1.3	0.8
M. nudipes E00257695	Tibet- Birmania (Edinburgh)	106	0.4	25	0.6	Y	19	12	20	4	Y	Y	Y	0.5	0.15	0.5		0.5	0.6	0.4
Erianthus Herb. Ind. Or. Hook.	9-13,000	25	0.2	14	0.4	Y	8	1.5	4	8	Y	Y	Y	0.6	0.15	0.35	Y	0.5	1.3	0.5
M. nudipes 7/2008 / 00313-87	Sikkin (Kew)	25	0.15	7	0.4	Y	10.5	0.8	5	9.5	Y	Y	N	0.6	0.2	0.5		0.5	0.8	0.4
M. nudipes 65/2008-32	Copenhagen	25	0.2	10	0.4	Y	8	7	4	7	Y	Y		0.6	0.15	0.5		0.5	0.8	0.4
M. nepalensis E00257698	Tibet-SW China (Edinburgh)	64	0.3	43	0.8	N	21	7	100	13	Y	Y		0.5	0.1	0.3		0.3	1.1	0.9
M. nepalensis 20286	East Himalaia (Oakes Ames)	65	0.3	16	0.6	N	16	3.5	20	10	Y	Y		0.4	0.15	0.3		0.3	1.3	0.5
Erianthus 8000-1	Himalaia (TCD)	45	0.1			Y	7	0.9	6	4	Y							0.4	1.2	0.5
Erianthus 8000-2	(TCD)	43	0.3	34	1.4	N	18.5	7.2	20	6	N	N		0.3	0	0.2		0.5	1.2	0.6
Erianthus 8000-3	Himalaia (TCD)	30	0.1	3	0.1	Y	5	0.5	5	4	Y	Y		0.3	0	0.3		0.3	1	0.5

Table B (*continued*)

Specimens	Location	Culm Height	Culm width	Leaf blade length	Leaf blade width	Leaf hair	Inflorescence length	Inflorescence axis length	Raceme number	Raceme length	Peduncle hairs	Axis hairs	Raceme axis hairs	Raceme internode	Upper pedicel length	Lower pedicel	Pedicel hairs	Spikelet length	Awn length	Spikelet hairs length
M. nepalensis E00044025	Yuman, China (EDINBURGH)	97	0.3	38	0.8	Y	14	5	15	12	N	N		0.4	0.2	0.3		0.3	1.3	1
M. nepalensis E00161213	Yuman, China (EDINBURGH)	96	0.2	27	1	N	14	4	20	5	Y	Y		0.3	0.2	0.3		0.2	0.5	0.6
M. nepalensis H2008/00313-86	Garhwal (Kew)	74	0.2	22	0.2	N	12	4	16	10.5	N	N		0.4	0.2	0.3		0.3	1.1	0.7
M. nepalensis H2008/00313-85	Nepal (Kew)	59	0.1	6	0.1	N	8	1.5	6	7				0.2	0.1	0.2		0.2	1	0.9
M. nepalensis H2008/00313-84	Malaya (Kew)	126	0.5	46	1		15	3	30	10				0.3	0.15	0.3		0.2	0.9	0.6
M. nepalensis H2008/00313-83	Tibet-SW China (Kew)	84	0.2	32	0.4	Y	14	6	20	11	Y	Y		0.2	0.1	0.2		0.2	1	0.6
M. nepalensis H2008/00313-82	Szechwan, China (Kew)	62	0.3	24	0.7	N	17	10	25	10	N	N		0.4	0.2	0.3		0.2	0.8	0.5
M. nepalensis H2008/00313-81	Yuman (Kew)	60	0.2	36	0.8	N	13	3	10	10	N	N		0.3	0.1	0.2		0.3	1.3	0.7
Eulalia nepalensis 7998 1	Mont Khasia (TCD)		0.4	11	0.4	N	16	8	40	16	Y	Y		0.3	0.15	0.3		0.3	1	0.7
Eulalia nepalensis 7998 2	Sikkin (TCD)		0.2	50	0.4	Y	16	7.5	30	12	Y	Y		0.3	0.15	0.3		0.3	0.9	0.8
Erianthus longisetosus 000938874	Siam (BM)		0.5	50	3.5	N	31	19	100	7	N	N		0.6	0.2	0.4		0.5	0.9	0.7

Table C Morphological characters scored in the Oak Park collection for the first replicate.

ID	Habit (clumped, spreading)	Space between	Culm wax (Y/N)	Max culm width	Culm buds or branching	Plant height	Leaf number	Leaf variegation	Internode length (cm)	Max leaf length (cm)	Max leaf width (cm)	Leaf hairs	INFLORESCENCE	Inflorescence length	Inflorescence axis length	Axis hairs	Raceme length	Raceme number	Raceme axis hairs	Raceme internode	Upper pedicel	Lower pedicel	Pedicel hairs	Spikelet length	Spikelet hairs length	Awn length
Miscanthus sp. Tea-1	Clumped	5	N	1.3	N	2.2	440	N	14	75 / 80 / 90 / 78	2.3 / 2.3 / 2.8 / 3	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. sinensis 'zebrinus' Tea-3	Clumped	1.5	N	0.9	N	50	240	Y	6	57 / 51 / 34 / 49	1.3 / 1.5 / 1.4 / 1.2	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. xgiganteus Tea-4	Clumped	3.5	N	1.3	Y	190	264	N	16	71 / 65 / 75 / 74	2.8 / 2.7 / 2.8 / 2.7	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. xgiganteus Tea-5	Clumped	4	N	1.3	Y	200	336	N	20	69 / 67 / 62 / 70	2.4 / 2.6 / 2.4 / 2.6	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-7	Clumped	5.5	N	1.2	Y	200	600	N	14	66 / 61 / 65 / 63	2.6 / 2 / 2.5 / 2.3	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-8	Spread	13	N	1	Y	170	180	N	9.5	77 / 57 / 74 / 72	2.2 / 2.3 / 2.4 / 2.5	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-9	Clumped	5	N	1	N	200	440	N	16	53 / 70 / 59 / 71	2.1 / 2.4 / 1.7 / 2.3	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-10	Clumped	7	N	1.1	Y	200	370	N	15	68 / 77 / 73 / 68	2.5 / 2.4 / 2.4 / 2.5	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-11	Clumped	4	N	1.2	Y	200	500	N	16	55 / 60 / 61 / 66	1.9 / 1.9 / 1.9 / 2	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. sinensis Tea-13	Spread	16	N	0.5	N	120	200	N	8.5	55 / 52 / 50 / 41	1.9 / 1.8 / 1.7 / 1.6	Y	Y	24	4	Y	14	26	N	0.5	0.4	0.2	N	0.5	0.7	0.4
M. sinensis Tea-14	Clumped	2	N	0.5	N	140	280	N	22	35 / 44 / 39 / 28	1.4 / 1.5 / 1.5 / 1.4	Y	Y	24	8	Y	15	37	N	0.5	0.4	0.2	Y	0.5	0.7	0.6
Miscanthus sp. Tea-15	Clumped	8	N	1.1	Y	210	440	N	11	58 / 57 / 74 / 74	2.2 / 2 / 2.6 / 2	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-16	Clumped	4	Y	0.6	N	110	280	N	9	52 / 49 / 59 / 52	1.2 / 1.2 / 1 / 1.1	Y	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. xgiganteus Tea-17	Clumped	4	N	1.2	Y	230	440	N	15	57 / 71 / 64 / 74	1.9 / 2.6 / 2.4 / 2.1	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-18	Clumped	7	Y	1.2	N	140	225	N	7	44 / 48 / 46 / 48	1.9 / 2.1 / 2 / 1.9	Y	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. sinensis goliath-like Tea-19	Clumped	4	Y	0.9	N	140	315	N	11	53 / 54 / 45 / 59	1.2 / 1.3 / 1.1 / 1.4	Y	Y	36	19	N	20	50	N	0.5	0.7	0.3	N	0.5	0.7	0.3
M. xgiganteus Tea-20	Clumped	6	N	1.2	Y	160	275	N	10	71 / 72 / 70 / 72	2.3 / 2.3 / 2.2 / 2.2	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-21	Clumped	5	N	1.4	Y	220	450	N	14	74 / 74 / 76 / 64	2.7 / 2.4 / 2.1 / 2.2	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-22	Clumped	2	Y	0.5	N	140	300	N	7	63 / 64 / 70 / 72	1.9 / 1.8 / 1.8 / 2.1	Y	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-23	Clumped	3	Y	1.1	N	120	140	N	10	80 / 58 / 73 / 78	1.9 / 1 / 1.4 / 1.5	Y	Y	37	19	N	21	40	N	0.6	0.7	0.3	N	0.6	0.7	0.5
Miscanthus sp. Tea-24	Spread	10	N	0.6	N	120	80	N	13	58 / 47 / 50 / 40	1.5 / 1.5 / 1.4 / 1.5	Y	Y	26	6.5	N	12	32	N	0.5	0.4	0.1	N	0.4	0.6	0.5
Miscanthus sp. Tea-25	Clumped	5	N	0.4	N	130	200	N	8	41 / 42 / 58 / 38	1 / 1.1 / 1 / 1.1	Y	Y	28	9	N	17	28	N	1.1	0.7	0.2	N	0.4	0.6	0.4
Miscanthus sp. Tea-26	Clumped	5	N	0.5	N	140	200	N	18	44 / 40 / 44 / 49	1 / 1 / 0.8 / 0.9	Y	Y	27	10.5	Y	12	25	N	0.6	0.4	0.2	N	0.4	0.6	0.7
Miscanthus sp. Tea-27	Spread	13	N	0.6	N	130	126	N	8	39 / 49 / 41 / 43	1.6 / 1.8 / 1.5 / 1.5	N	Y	22	6	N	16	22	N	0.5	0.3	0.1	N	0.5	0.6	0.7

Table C (*continued*)

ID	Habit (clumped, spreading)	Space between culms (cm)	Culm wax (Y/N)	Max culm width	Culm buds or branching	Plant height	Leaf number	Leaf variegation	Internode length (cm)	Max leaf length (cm)	Max leaf width (cm)	Leaf hairs	INFLORESCENCE	Inflorescence length	Inflorescence axis length	Axis hairs	Raceme length	Raceme number	Raceme axis hairs	Raceme internode	Upper pedicel length	Lower pedicel length	Pedicel hairs	Spikelet length	Spikelet hairs length	Awn length
Miscanthus sp. Tea-28	Clumped	5	Y	1.5	Y	220	385	N	10	80 / 69 / 77 / 64	2.3 / 2.6 / 2.7 / 2	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. sinensis Tea-30	Clumped	2	N	0.5	N	120	175	N	7	56 / 58 / 59 / 62	1.9 / 2.1 / 1.7 / 2.1	Y	Y	21	8	N	14	30	N	0.4	0.6	0.2	N	4	0.6	0.3
M. xgiganteus Tea-31	Clumped	5	Y	1.1	Y	180	225	N	13	76 / 60 / 77 / 77	2.6 / 2 / 2.9 / 3	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-32	Clumped	15	N	1.1	Y	170	168	N	16	77 / 77 / 75 / 82	2.8 / 2.5 / 2.7 / 2.7	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. sinensis 'zebrinus' Tea-33	Spread	14	N	0.3	N	70	75	N	8	50 / 57 / 48 / 48	1 / 1.1 / 1 / 1	Y	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-34	Clumped	5	Y	1.3	N	150	128	N	12	65 / 62 / 65 / 72	1.8 / 1.6 / 1.5 / 2	Y	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. sinensis 'gross fontane' Tea-35	Clumped	10	N	0.5	N	80	160	N	9	57 / 61 / 63 / 65	1.4 / 1.4 / 1.5 / 1.6	Y	Y	28	10	N	18	30	N	0.9	0.5	0.2	N	0.4	0.7	0.4
M. sinensis 'gross fontane' Tea-36	Spread	8	N	0.5	N	80	240	N	6	50 / 49 / 49 / 48	1.9 / 1.8 / 1.9 / 1.7	Y	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-37	Spread	20	Y	0.5	N	50	60	N	5	39 / 38 / 38 / 44	1.3 / 1.3 / 1.2 / 1.2	Y	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-38	Spread	4	N	0.5	N	70	18	N	8	47 / 37 / 42 / 42	1.2 / 1 / 1.3 / 1.3	Y	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-39	Clumped	5	N	0.5	N	110	240	N	8	55 / 51 / 43 / 52	1.6 / 1.6 / 1.3 / 1.5	Y	Y	26	6	Y	19	30	N	0.7	0.6	0.3	N	0.5	0.6	0.5
M. sinensis Tea-40	Spread	10	Y	0.8	N	90	250	N	8	57 / 57 / 56 / 56	2 / 2 / 1.8 / 2	Y	Y	28	12	N	13	40	N	0.5	0.5	0.2	N	0.4	0.5	0.5
Miscanthus sp. Tea-41	Spread	30	N	0.5	N	130	90	N	8	41 / 39 / 42 / 49	1.5 / 1.6 / 1.5 / 1.7	Y	Y	25	11	N	17	23	N	0.7	0.5	0.2	N	0.5	0.6	
Miscanthus sp. Tea-42	Clumped	6	N	0.4	N	100	110	N	12	36 / 37 / 37 / 47	1.4 / 1.5 / 1.4 / 1.4	Y	N	25	4	Y	18	14	N	0.6	0.5	0.2	N	0.4	0.6	0.8
Miscanthus sp. Tea-43	Clumped	4	Y	0.9	N	100	84	N	5	56 / 57 / 50 / 61	1.6 / 1.6 / 1.7 / 2	Y	Y	30	13	N	16	35	N	0.5	0.6	0.2	N	0.5	0.6	0.5
M. condensatus Tea-44	Spread	10	Y	10	N	110	140	N	6	47 / 54 / 62 / 55	1.6 / 1.7 / 1.6 / 1.7	Y	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-45	Spread	30	Y	0.7	N	60	65	N	7	30 / 31 / 36 / 32	2 / 3 / 2.2 / 2.8	N	Y	33	17	N	19	17	N	0.5	0.6	0.3	N	0.5	0.6	0.8
Miscanthus sp. Tea-48	Clumped	3	N	1	N	140	200	N	12	47 / 50 / 57 / 57	1.2 / 1.4 / 1.5 / 1.6	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-49	Spread	18	N	1	N	110	150	N	8	61 / 58 / 63 / 61	1.6 / 1.5 / 1.6 / 1.4	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-51	Clumped	5	N	0.6	N	60	20	*	4	* / * / * / *	* / * / * / *	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-52	Clumped	6	N	1.3	N	110	30	N	7	48 / 55 / 57 / 41	1.2 / 1.2 / 1.3 / 1.3	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-54	*	*	*	*	*	*	*	N	*	47 / 44 / 34 / 34	1.3 / 1.1 / 1 / 1	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-55	Clumped	4	N	0.9	Y	190	360	N	13	51 / 67 / 64 / 65	1.8 / 2.2 / 1.9 / 1.8	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. sinensis 'goliath' Tea-56	Clumped	2	Y	0.6	N	120	140	N	4	53 / 65 / 61 / 59	1.4 / 1.3 / 1.3 / 1.3	Y	Y	36	19	N	22	28	N	0.6	0.5	0.2	N	0.5	0.6	0.4
M. sinensis 'goliath' Tea-57	Clumped	5	N	0.5	N	130	240	N	12	65 / 54 / 65 / 56	1.1 / 1 / 0.9 / 0.9	Y	N	31	19	N	18	35	N	0.5	0.6	0.3	N	0.5	0.6	0.5

Table C (*continued*)

ID	Habit (clumped, spreading)	Space between culms (cm)	Culm wax (Y/N)	Max culm width	Culm buds or branching	Plant height	Leaf number	Leaf variegation	Internode length (cm)	Max leaf length (cm)	Max leaf width (cm)	Leaf hairs (Y/N)	INFLORESCENCE (Y/N)	Inflorescence length	Inflorescence axis length	Axis hairs	Raceme length	Raceme number	Raceme axis hairs	Raceme internode	Upper pedicel length	Lower pedicel length	Pedicel hairs	Spikelet length	Spikelet hairs length	Awn length
M. sinensis 'sirene' Tea-58	Clumped	6	N	0.7	N	100	270	N	10	57 50 50 57	0.9 0.8 0.9 0.9	Y	2	*	*	*	*	*	*	*	*	*	*	*	*	*
M. sinensis 'strictus' Tea-59	Clumped	1	N	0.4	N	100	150	N	7	56 55 50 59	0.3 0.3 0.25 0.3	Y	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. sinensis 'strictus' Tea-60	Clumped	3	N	0.8		90	140	B	5	46 44 40 45	1.5 1.5 1.1 1.3	Y	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. sinensis 'malaparteus' Tea-61	Clumped	2	N	1	N	110	90	N	6	45 37 38 38	0.8 0.6 0.7 0.8	Y	Y	31	13	N	12	45	N	0.6	0.4	0.2	N	0.5	0.7	0.3
M. sinensis Tea-62	*	*	*	*	*	*	*	N	*	43 44 42 35	1 1.1 1 1.1	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. xgiganteus Tea-64	Clumped	3	N	1	N	160	220	N	11	76 77 74 70	2.4 2.4 2.5 2.5	Y	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. xgiganteus Tea-65	Clumped	5	N	1	N	170	250	N	16	74 68 74 64	1.8 1.8 1.7 1.6	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. xgiganteus Tea-66	Clumped	13	N	1	Y	200	320	N	15	67 66 64 65	2.2 2 2.5 2.4	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-68	Clumped	4	N	0.9	Y	210	400	N	15	69 70 64 67	2.4 2.4 2.3 1.9	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-69	Clumped	11	Y	1	N	120	63	N	16	69 59 71 66	2.1 1.8 1.8 2	Y	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-70	*	*	*	*	*	*	*	N	*	57 64 61 65	1.6 1.5 1.4 1.5	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-71	Clumped	11	N	1.3	N	170	200	N	13	66 64 69 72	1.9 2.3 2.4 2.2	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-72	Clumped	5	N	0.8	Y	140	120	N	12	75 71 71 80	2 1.9 1.3 2.1	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-73	Clumped	2	N	0.5	N	100	240	N	7	32 29 29 26	1.9 1.2 1.1 1	Y	Y	17	4	Y	13	24	N	0.5	0.4	0.2	N	0.4	0.8	0.6
M. xgiganteus Tea-74	Clumped	6	Y	0.9	Y	140	120	N	14	64 63 65 72	2 2 1.8 2	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. sacchariflorus x M. sinensis Tea-75	Clumped	9	N	0.7	N	120	440	N	10	50 48 49 34	1.7 1.9 1.7 1.5	N	Y	26	11	N	19	38	N	0.5	0.6	0.3	N	0.4	0.7	0.2
M. sinensis Tea-76	Clumped	3	N	0.4	N	90	160	N	6	38 31 35 27	1.3 1.2 1.3 1.2	Y	Y	15	4	N	11	15	N	0.5	0.3	0.1	N	0.4	0.6	0.7
M. sinensis Tea-77	Clumped	3	N	0.4	N	70	140	N	2.8	38 32 36 25	1 1.1 1.1 1.2	Y	Y	14	3.5	N	11	10	N	0.5	0.4	0.2	N	0.4	0.5	0.6
M. sinensis Tea-78	Clumped	3	N	0.6	N	90	120	N	10	47 46 44 40	1.8 1.8 1.7 1.7	Y	Y	21	3	Y	14	28	N	0.6	0.4	0.1	N	0.4	0.5	0.5
M. sinensis Tea-79	Clumped	4	N	0.5	N	100	90	N	6	39 41 42 38	1.5 1.4 1.4 1.2	Y	Y	17.5	1	N	13.5	8	N	0.6	0.4	0.2	N	0.4	0.6	0.5
M. sinensis Tea-80	Spread	18	N	0.4	N	90	40	N	7	33 39 42 30	1.2 1.2 1.1 1	Y	Y	19	0.5	Y	14	20	N	0.7	0.6	0.3	N	0.4	0.5	0.5
M. xgiganteus Tea-81	Clumped	4	N	1	N	130	160	N	14	62 63 66 71	2.2 2.2 2.1 2.2	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. xgiganteus Tea-82	Clumped	8	Y	1	N	140	90	N	12	35 65 72 71	2 1.9 2.3 2.1	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. xgiganteus Tea-83	Clumped	14	Y	1	N	160	180	N	11	68 69 65 75	2.4 2.4 2.6 2.3	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. sacchariflorus Tea-84	Clumped	7	Y	0.7	Y	140	90	N	9	68 71 80 68	3 3.4 3 2.9	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*

149

Table C (*continued*)

ID	Habit (clumped, spreading)	Space between culms (cm)	Culm wax (Y/N)	Max culm width	Culm buds or branching	Plant height	Leaf number	Leaf variegation	Internode length (cm)	Max leaf length (cm)	Max leaf width (cm)	Leaf hairs	INFLORESCENCE	Inflorescence length	Inflorescence axis length	Axis hairs	Raceme length	Raceme number	Raceme axis hairs	Raceme internode	Upper pedicel length	Lower pedicel length	Pedicel hairs	Spikelet length	Spikelet hairs length	Awn length
M. sinensis goliath-like Tea-85										63	1.2															
										79	1.4															
										63	1.5															
	Clumped	5	N	0.6	N	100	160	N	9	75	1.4	Y	Y	32	13	N	18	19	N	0.5	0.6	0.3	N	0.5	0.8	0.5
M. sinensis Tea-86										57	1.4															
										57	1.3															
										56	1.6															
	Clumped	5	N	0.8	N	90	180	N	4	55	1.5	Y	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. sacchariflorus x M. sinensis Tea-87										38	2.1															
										43	2															
										42	2															
	Clumped	5	N	0.7	N	120	360	N	12	38	1.8	N	Y	28	7	N	17	25	N	0.7	0.5	0.2	N	0.5	0.8	N
M. sinensis Tea-88										55	2.1															
										56	2.3															
										57	2.2															
	Clumped	4	Y	1	N	120	210	N	5	50	2.2	Y	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-89										*	*															
										*	*															
										*	*															
	Spread	4	N	1	N	60	60	*	3	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-90										75	1.8															
										75	1.7															
										82	1.9															
	Clumped	4	N	0.9	N	120	160	N	4	77	1.8	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-91										58	1.9															
										50	1.9															
										52	1.8															
	Clumped	10	N	1	N	130	150	N	7	59	1.9	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-92										54	1.8															
										51	1.8															
										50	1.8															
	Clumped	4	N	1.1	N	100	160	N	4	52	1.6	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. xgiganteus Tea-93										64	2.2															
										57	2															
										55	2.1															
	Clumped	6	N	1.1	N	170	210	N	10	65	1.8	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. xgiganteus Tea-94										70	2.3															
										70	2.5															
										74	2.3															
	Clumped	5	N	1	Y	150	210	N	12	60	1.8	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. sinensis Tea-95										45	1.7															
										53	1.8															
										53	1.5															
	Clumped	4	N	0.7	N	150	120	N	13	47	1.5	Y	Y	26	11	N	18	23	N	0.6	0.6	0.2	N	0.4	0.8	0.5
M. sinensis Tea-96										62	2.1															
										58	1.9															
										60	2.1															
	Spread	14	N	0.9	N	170	80	N	13	61	2.1	Y	Y	38	18	N	19	40	N	0.5	0.4	0.2	N	0.4	0.6	0.4
M. sinensis Tea-97										48	1.3															
										48	1.3															
										41	1.4															
	Clumped	4	N	0.6	N	140	90	N	12	52	1.4	Y	Y	27	9	Y	17	23	N	0.6	0.7	0.3	N	0.5	1	0.7
M. sinensis Tea-98										41	1.9															
										47	1.7															
										36	1.7															
	Clumped	12	N	0.7	N	130	140	N	9	45	1.7	Y	Y	27	9	N	18	25	N	0.6	0.5	0.3	N	0.5	0.8	0.7
M. sinensis Tea-99										39	1.7															
										40	1.8															
										40	1.7															
	Clumped	5	N	0.7	N	150	180	N	10	37	1.8	Y	Y	29	15	N	16	33	N	0.6	0.5	0.2	N	0.4	0.6	0.5
M. sinensis Tea-100										35	1.5															
										31	1.5															
										38	1.4															
	Clumped	10	N	0.7	N	150	120	N	7	42	1.5	Y	4	27	15	N	13	45	N	0.5	0.5	0.2	N	0.5	0.6	0.3
M. sinensis Tea-101										31	1.8															
										31	2															
										35	1.8															
	Clumped	4	N	0.6	N	130	120	N	8	34	1.7	Y	8	26	13	N	15	25	N	0.6	0.5	0.2	N	0.5	0.8	0.4
M. sinensis Tea-102										29	1.9															
										26	1.5															
										30	1.9															
	Spread	8	N	0.9	N	170	90	N	11	35	1.7	Y	8	25	13	N	13	34	N	0.6	0.6	0.3	N	0.4	0.5	0.5
M. sinensis Tea-103										38	1.4															
										39	1.4															
										36	1.4															
	Clumped	8	N	0.5	N	150	90	N	20	34	1.6	Y	2	22	8	N	14	32	N	0.5	0.4	0.1	N	0.5	0.8	0.6
M. sinensis Tea-104										42	2.1															
										49	2.4															
										45	2.2															
	Spread	36	N	0.7	N	150	70	N	9	54	2	N	N	24	10	N	15	40	N	0.5	0.5	0.1	N	0.5	0.8	0.7
M. sinensis Tea-105										45	1.8															
										52	1.9															
										45	1.9															
	Spread	9	N	1	N	150	35	N	8	42	1.4	Y	N	25	6	N	16	25	N	0.4	0.4	0.1	N	0.5	0.8	0.6
M. sinensis Tea-106										49	1.6															
										50	1.7															
										52	1.6															
	Spread	11	N	0.6	N	170	80	N	9	47	1.6	Y	1	24	11	N	12	32	N	0.6	0.4	0.1	N	0.5	0.7	0.5
M. sinensis Tea-107										47	2.3															
										45	2.1															
										52	2.1															
	Clumped	8	N	0.7	N	180	140	N	10	48	2	Y	2.5	23	9	N	13	43	N	0.4	0.3	0.1	N	0.4	0.6	0.5
M. sinensis Tea-108										59	1.6															
										51	1.4															
										51	1.7															
	Clumped	10	N	0.6	N	130	60	N	9	43	1.1	Y	N	30	11	N	19	20	N	0.6	0.4	0.2	N	0.6	0.9	0.6
M. sinensis Tea-109										39	1															
										42	1															
										41	1.2															
	Spread	15	N	0.7	N	160	28	N	8	46	1.2	Y	N	31	10	N	17	21	N	0.5	0.5	0.2	N	0.5	0.6	0.5

Table C (*continued*)

ID	Habit (clumped, spreading)	Space between culms (cm)	Culm wax (Y/N)	Max culm width	Culm buds or branching	Plant height	Leaf number	Leaf variegation	Internode length (cm)	Max leaf length (cm)	Max leaf width (cm)	Leaf hairs (Y/N)	INFLORESCENCE	Inflorescence length	Inflorescence axis length	Axis hairs	Raceme length	Raceme number	Raceme axis hairs	Raceme internode	Upper pedicel length	Lower pedicel length	Pedicel hairs	Spikelet length	Spikelet hairs length	Awn length
M. sinensis Tea-110										36	1.1															
										38	1.2															
										36	1.1															
	Clumped	7	N	0.6	N	120	120	N	10	36	1.1	Y	1.5	25	15	N	17	29	N	0.5	0.5	0.2	N	0.5	0.7	0.5
M. sinensis Tea-111										39	1.6															
										33	1.5															
										44	1.5															
	Spread	23	N	0.7	N	150	120	N	7	38	1.4	Y	3	21	10	N	15	35	N	0.7	0.5	0.2	N	0.5	0.8	0.5
M. sinensis Tea-112										40	1.5															
										38	1.5															
										30	1.4															
	Spread	25	N	0.6		150	60	N	12	32	1.6	Y	3	24	11	N	13	23	N	0.7	0.5	0.2	N	0.5	0.8	0.5
M. sinensis Tea-113										42	2															
										47	2.3															
										55	2.2															
	Clumped	6	N	0.6	N	170	90	N	11	50	2	Y	5	22	6	Y	17	25	N	0.7	0.5	0.2	N	0.4	0.6	0.4
M. sinensis Tea-114										34	1.1															
										34	1.2															
										32	1.2															
	Spread	30	N	0.5	N	120	24	N	9	27	1	Y	1	24	10	N	20	53	N	0.6	0.4	0.2	N	0.4	0.7	0.4
M. sinensis Tea-115										46	1.7															
										54	2.1															
										40	2															
	Clumped	10	N	0.7	N	120	70	N	7	48	2	Y	1	21	11	N	15	35	N	0.5	0.4	0.2	N	0.4	0.6	0.4
M. sacchariflorus x M. sinensis Tea-116										41	2															
										36	1.8															
										40	2															
	Clumped	11	N	0.5	N	100	200	N	5	41	1.7	N	N	23	11	N	11	18	N	0.6	0.4	0.1	N	0.5	0.8	N
M. sacchariflorus x M. sinensis Tea-117										44	1.8															
										44	2.2															
										46	2.1															
	Clumped	4	N	0.4	N	100	160	N	6	43	1.8	N	2.5	23	11	N	11	20	N	0.6	0.5	0.2	N	0.4	1	0.2
M. sacchariflorus x M. sinensis Tea-118										29	1.5															
										29	1.6															
										25	1.5															
	Clumped	5	N	0.5	N	90	100	N	6	37	1.2	N	1	32	19	N	16	25	N	0.7	0.4	0.2	N	0.4	0.8	N
M. sacchariflorus x M. sinensis Tea-119										40	1.6															
										38	1.8															
										40	1.8															
	Spread	22	N	0.9	N	140	300	N	5	37	1.4	N	8	29	16	N	16	30	N	0.5	0.6	0.3	N	0.4	0.7	N
M. sacchariflorus x M. sinensis Tea-120										33	1.4															
										32	1.6															
										38	1.5															
	Clumped	3	N	0.6	N	110	310	N	10	29	1.7	N	5	27	15	N	12	32	N	0.5	0.4	0.2	N	0.4	0.7	N
M. sacchariflorus x M. sinensis Tea-121										28	1.5															
										38	1.5															
										35	1.5															
	Clumped	6	N	0.4	N	90	180	N	7	34	1.6	N	3	24	13	N	15	22	N	0.6	0.5	0.2	N	0.4	0.9	N
M. sacchariflorus x M. sinensis Tea-122										50	1.9															
										48	2															
										50	2															
	Spread	16	N	0.6	N	100	160	N	7	46	1.9	N	N	24	12	N	12	18	N	0.6	0.5	0.2	N	0.5	1	0.2
M. sacchariflorus x M. sinensis Tea-123										41	1.6															
										41	1.7															
										34	1.6															
	Clumped		N	0.5	N	70	8	N	7	40	1.4	N	1	*	*	*	*	*	*	*	*	*	*	*	*	*
M. sacchariflorus x M. sinensis Tea-124										58	2															
										52	2.1															
										60	2.1															
	Spread	6	N	0.7	N	100	160	N	11	54	1.8	N	3	28	14	N	14	24	N	0.5	0.6	0.3	N	0.5	1	0.2
M. sacchariflorus x M. sinensis Tea-125										47	1.4															
										44	1.3															
										44	1.4															
	Spread	15	N	0.5	N	130	300	N	12	44	1.5	N	5	24	10	N	14	22	N	0.6	0.5	0.2	N	0.5	0.9	0.3
M. sacchariflorus x M. sinensis Tea-126										52	1.4															
										50	1.6															
										59	1.5															
	Spread	6	N	1.2	Y	140	210	N	12	59	1.6	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. sacchariflorus x M. sinensis Tea-127										43	2.1															
										44	1.5															
										48	2															
	Clumped	2	N	0.5	N	90	60	N	5	42	1.5	N	N	30	13	N	18	19	N	0.7	0.6	0.2	N	0.4	1	0.2
M. sacchariflorus Tea-128										57	2.2															
										61	2.1															
										69	2.1															
	Spread	18	N	0.7	N	160	140	N	10	64	2.5	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. sacchariflorus Tea-129										59	2.3															
										56	2.5															
										54	2.3															
	Spread	27	N	0.7	N	160	90	N	12	63	2.2	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*

Table D Morphological characters scored in the Oak Park collection for the second replicate.

ID	Habit (clumped, Spread)	Space between culms (cm)	Culm wax (Y/N)	Max culm width	Culm buds or Vanching (Y/N)	Plant height	Leaf number	Leaf variegation	Internode length (cm)	Max leaf length (cm)	Max leaf width (cm)	Leaf hairs	INFLORESCENCE	Inflorescence length	Inflorescence axis length	Axis hairs	Raceme length	Raceme number	Raceme axis hairs	Raceme internode	Upper pedicel length	Lower pedicel length	Pedicel hairs	Spikelet length	Spikelet hairs length	Awn length
M. xgiganteus Tea-4	Clumped	10	N	1.1	N	150	211	N	10	58 / 59 / 61 / 67	1.8 / 1.7 / 1.9 / 1.8	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. xgiganteus Tea-5	Clumped	10	Y	1.1	Y	220	311	N	12	62 / 67 / 63 / 67	2.5 / 2.5 / 2.2 / 2.6	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. xgiganteus Tea-17	Clumped	6	Y	1.1	Y	220	260	N	20	61 / 55 / 68 / 70	2.5 / 2.2 / 2.3 / 2.3	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. xgiganteus Tea-20	Clumped	4	Y	1.2	Y	220	310	N	14	60 / 63 / 66 / 65	2.4 / 2.4 / 2.4 / 2.2	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. xgiganteus Tea-31	Clumped	5	Y	1	Y	220	300	N	14	65 / 70 / 61 / 68	2.4 / 2 / 2.3 / 2.3	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-32	Clumped	5	Y	1.1	Y	220	260	N	13	68 / 66 / 62 / 67	2 / 2 / 2.4 / 2.3	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. sinensis 'sirene' Tea-63	Clumped	10	Y	1.1	N	120	110	N	6	61 / 70 / 67 / 70	1.3 / 1.6 / 1.9 / 1.7	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. xgiganteus Tea-64	Clumped	5	Y	1	Y	140	200	N	10	60 / 60 / 70 / 70	1.6 / 1.7 / 1.8 / 1.7	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. xgiganteus Tea-65	Clumped	15	Y	1	N	170	210	N	10	65 / 62 / 66 / 67	2 / 2 / 2.2 / 1.9	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. xgiganteus Tea-74	Clumped	10	N	1.2	N	170	240	N	12	69 / 76 / 73 / 64	1.9 / 2.1 / 2 / 2	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. xgiganteus Tea-81	Clumped	7	Y	1	Y	170	210	N	12	60 / 70 / 70 / 71	2.1 / 2.1 / 2.1 / 2.2	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. xgiganteus Tea-82	Clumped	10	N	1.2	Y	170	240	N	12	64 / 55 / 69 / 66	1.9 / 1.6 / 2 / 2.2		N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. xgiganteus Tea-83	Clumped	8	Y	1.3	Y	210	300	N	13	56 / 60 / 68 / 78	2.7 / 2.2 / 2.4 / 2.6	N	CLOSE	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-90	Clumped	4	N	0.9	N	130	160	N	17	67 / 73 / 77 / 69	1.7 / 1.6 / 1.7 / 1.5	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-91	Clumped	6	Y	0.8	N	140	160	N	9	58 / 59 / 59 / 55	2 / 1.9 / 2 / 1.8	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-92	Clumped	4	Y	1	N	110	160	N	5	49 / 53 / 49 / 49	1.6 / 1.7 / 1.7 / 1.7	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. xgiganteus Tea-93	Clumped	21	N	1.1	N	150	240	N	9	65 / 72 / 64 / 65	2.5 / 2.4 / 2.2 / 2.1	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. xgiganteus Tea-94	Clumped	6	Y	1	N	140	100	N	11	52 / 57 / 55 / 61	1.6 / 1.7 / 1.7 / 1.7	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. sinensis Tea-95	Clumped	13	N	0.7	N	180	160	N	9	55 / 55 / 51 / 55	1.8 / 1.8 / 1.9 / 1.8	Y	7	29	18	N	13	50	N	0.5	0.5	0.2	N	0.4	0.6	0.5
M. sinensis Tea-96	Clumped	8	Y	1	N	170	120	N	12	49 / 60 / 45 / 42	1.7 / 1.8 / 1.7 / 1.5	Y	2	39	23	N	19	41	N	0.6	0.5	0.2	N	0.4	0.7	0.3
M. sinensis Tea-97	Clumped	5	Y	0.6	N	150	210	N	12	34 / 38 / 37 / 35	1.3 / 1.3 / 1.3 / 1.7	Y	6	20	11	N	16	36	N	0.6	0.6	0.3	N	0.4	0.8	0.5
M. sinensis Tea-98	Clumped	16	N	0.7	N	150	120	N	9	52 / 52 / 48 / 50	1.8 / 1.8 / 1.7 / 1.7	Y	1	24	11	N	13	40	N	0.5	0.4	0.1	N	0.5	0.7	0.6
M. sinensis Tea-99	Clumped	3	N	0.7	N	120	120	N	6	43 / 39 / 50 / 42	2 / 1.9 / 1.6 / 2	Y	2	19	7	N	12	21	N	0.6	0.4	0.1	N	0.4	0.6	0.4
M. sinensis Tea-100	Spread	10	N	0.5	N	140	35	N	10	65 / 43 / 58 / 50	1.5 / 1.9 / 1.6 / 1.9	Y	1	24	8	N	11	38	N	0.5	0.5	0.2	N	0.5	0.6	0.3
M. sinensis Tea-101	Clumped	13	Y	0.4	N	140	120	N	9	36 / 42 / 35 / 39	1.7 / 1.9 / 1.7 / 1.9	Y	9	21	11	'N	11	30	N	0.5	0.4	0.2	N	0.4	0.8	0.5

Table D (*continued*)

ID	Habit (clumped,Spread)	Space between culms (cm)	Culm wax (Y/N)	Max culm width	Culm buds or Yanching (Y/N)	Plant height	Leaf number	Leaf variegation	Internode length (cm)	Max leaf length (cm)	Max leaf width (cm)	Leaf hairs	INFLORESCENCE	Inflorescence length	Inflorescence axis length	Axis hairs	Raceme length	Raceme number	Raceme axis hairs	Raceme internode	Upper pedicel length	Lower pedicel length	Pedicel hairs	Spikelet length	Spikelet hairs length	Awn length
M. sinensis Tea-102	Clumped	8	Y	0.5	N	140	160	N	8	35; 35; 36; 36	1.6; 1.4; 1.5; 1.7	Y	2	25	12	N	18	45	N	0.5	0.5	0.2	N	0.4	0.6	0.4
M. sinensis Tea-104	Clumped	25	N	0.6	N	150	65	N	9	43; 43; 44; 43	2.3; 2.1; 2.5; 2	N	2	25	14	N	19	65	N	0.5	0.5	0.2	N	0.5	0.8	0.6
M. sinensis Tea-105	Spread	23	Y	1.1	N	120	36	N	6	50; 53; 49; 46	2.4; 2.7; 2.6; 2	Y	3	30	16	Y	20	53	N	0.6	0.4	0.1	N	0.4	0.7	0.5
M. sinensis Tea-106	Clumped	4	N	0.5	N	130	42	N	7	47; 53; 44; 48	1.3; 1.6; 1.4; 1.9	Y	N	18	9	N	11	22	N	0.5	0.4	0.1	N	0.4	0.7	0.5
M. sinensis Tea-107	Clumped	7	N	0.6	N	150	15	N	7	40; 40; 48; 51	2; 1.9; 2; 1.9	Y	2	24	15	N	11	60	N	0.4	0.3	0.1	N	0.4	0.6	0.5
M. sinensis Tea-108	Clumped	6	N	0.5	N	100	42	N	7	29; 43; 39; 43	1.2; 1.5; 1.3; 1.4	Y	N	27	14	N	18	30	N	0.5	0.4	0.2	N	0.4	0.7	0.6
M. sinensis Tea-109	Clumped	8	N	0.6	N	150	35	N	10	47; 53; 54; 42	1.3; 1.4; 1.6; 1.4	Y	1	29	16	N	21	20	N	0.7	0.7	0.3	N	0.4	0.7	0.5
M. sinensis Tea-110	Clumped	4	N	0.5	N	160	160	N	9	39; 50; 36; 37	1.3; 1.1; 1.5; 1.1	Y	6	25	10	N	17	35	N	0.8	0.4	0.2	N	0.4	0.5	0.5
M. sinensis Tea-111	Clumped	6	N	0.9	N	150	50		7	*; *; *	*; *; *			29	16	Y	15	36	N	0.7	0.6	0.3	N	0.5	0.7	0.4
M. sinensis Tea-112	Clumped	11	N	0.5	N	160	70	N	10	38; 38; 34; 34	1.3; 1.3; 1.5; 1.6	Y	3	25	14	N	15	38	N	0.4	0.4	0.1	N	0.4	0.6	0.5
M. sinensis Tea-113	Clumped	5	N	0.7	N	170	100	N	14	31; 33; 46; 30	1.8; 1.8; 1.5; 1.7	Y	8	28	14	Y	18	40	N	0.7	0.5	0.3	N	0.4	0.7	0.4
M. sinensis Tea-114	Clumped	6	N	0.5	N	150	80	N	6	36; 39; 41; 39	1.2; 1.1; 1.1; 1	Y	1	22	11	N	15	54	N	0.5	0.4	0.1	N	0.4	0.6	0.4
M. sinensis Tea-115	Clumped	7	Y	1	N	150	70	N	13	51; 53; 40; 42	2; 1.9; 1.9; 1.8	Y	CLOSE	28	16	N	13	33	N	0.5	0.5	0.2	N	0.5	0.6	0.4
M. sacchariflorus x M. sinensis Tea-116	Clumped	3	N	0.5	N	150	90	N	10	38; 39; 39; 40	1.8; 1.8; 1.7; 1.9	N	Y	22	9	N	11	20	N	0.3	0.3	0.1	N	0.4	1	N
M. sacchariflorus x M. sinensis Tea-117	Clumped	4	N	0.5	N	150	300	N	10	39; 36; 39; 34	1.8; 1.8; 1.6; 2	N	8	23	11	N	11	25	N	0.5	0.5	0.2	N	0.5	0.9	0.2
M. sacchariflorus x M. sinensis Tea-118	Clumped	6	Y	0.7	N	90	80	N	6	28; 37; 34; 29	1.8; 1.5; 1.7; 1.4	N	4	29	14	N	14	26	N	0.5	0.4	0.2	N	0.4	0.7	N
M. sacchariflorus x M. sinensis Tea-119	Clumped	18	Y	0.8	N	100	210	N	5	51; 45; 51; 39	1.5; 1.8; 1.6; 1.9	N	4	35	20	N	20	35	N	0.7	0.5	0.2	N	0.5	0.6	0.2
M. sacchariflorus x M. sinensis Tea-120	Clumped	3	N	0.4	N	80	360	N	6	34; 38; 42; 34	1.3; 1.6; 1.2; 1.2	N	N	30	13	N	17	23	N	0.6	0.4	0.2	N	0.5	0.9	N
M. sacchariflorus x M. sinensis Tea-121	Clumped	5	N	0.5	N	60	210	N	4	43; 34; 41; 38	1.9; 1.6; 2; 2	N	N	21	9	N	13	14	N	0.7	0.5	0.2	N	0.5	0.9	N
M. sacchariflorus x M. sinensis Tea-122	Clumped	10	Y	0.9	N	130	240	N	7	44; 55; 48; 54	1.4; 1.8; 1.6; 1.8	N	N	30	16	N	13	25	N	0.6	0.6	0.2	N	0.5	0.9	0.3
M. sacchariflorus x M. sinensis Tea-123	Clumped	3	N	0.5	N	150	300	N	10	40; 40; 43; 34	1.6; 1.6; 1.4; 1.4	N	N	28	16	N	15	35	N	0.7	0.5	0.2	N	0.4	0.8	N
M. sacchariflorus x M. sinensis Tea-124	Clumped	19	Y	0.9	N	140	400	N	9	44; 44; 51; 43	2; 1.9; 1.7; 1.8	N	5	*	*	*	*	*	*	*	*	*	*	*	*	*
M. sacchariflorus x M. sinensis Tea-125	Clumped	12	Y	0.5	N	100	160	N	6	39; 41; 44; 42	1.2; 1.1; 1.2; 1.1	N	N	26	12	N	13	21	N	0.7	0.5	0.2	N	0.6	0.9	N
M. sacchariflorus x M. sinensis Tea-126	Spread	10	N	1.2	N	130	104	N	10	52; 50; 56; 51	1.7; 1.6; 1.8; 1.7	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. sacchariflorus x M. sinensis Tea-127	Clumped	7	N	0.8	N	150	210	N	9	47; 43; 44; 47	1.6; 1.7; 1.9; 1.8	N	N	37	19	N	20	39	N	1	0.6	0.3	N	0.5	1	0.3

Table D (*continued*)

ID	Habit (clumped, Spread)	Space between culms (cm)	Culm wax (Y/N)	Max culm width	Culm buds or Yanching (Y/N)	Plant height	Leaf number	Leaf variegation	Internode length (cm)	Max leaf length (cm)	Max leaf width (cm)	Leaf hairs	INFLORESCENCE	Inflorescence length	Inflorescence axis length	Axis hairs	Raceme length	Raceme number	Raceme axis hairs	Raceme internode	Upper pedicel length	Lower pedicel length	Pedicel hairs	Spikelet length	Spikelet hairs length	Awn length
Miscanthus sp. Tea-8	Spread	19	Y	1	Y	180	310	N	10	74 / 73 / 80 / 70	2.7 / 2.3 / 2.7 / 2.5	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. sacchariflorus Tea-128	Spread	32	Y	1.1	N	170	45	N	10	62 / 60 / 62 / 70	2.3 / 2.6 / 2.6 / 2.6	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. sacchariflorus x M. sinensis Tea-75	Clumped	7	N	0.8	N	170	520	N	10	46 / 45 / 43 / 40	1.7 / 1.6 / 1.6 / 1.8	N	6	26	11	N	15	33	N	0.5	0.5	0.2	N	0.4	0.8	0.3
M. sinensis goliath-like Tea-85	Clumped	5	Y	0.7	N	130	140	N	12	68 / 48 / 61 / 62	0.9 / 0.8 / 1 / 0.7	Y	N	35	20	N	16	37	N	0.6	0.7	0.3	N	0.6	0.9	0.5
M. sinensis Tea-86	Clumped	4	Y	0.9	N	140	360	N	9	65 / 58 / 60 / 65	1.8 / 1.5 / 1.5 / 1.9	Y	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. sacchariflorus x M. sinensis Tea-87	Clumped	5	N	0.8	N	160	400	N	10	40 / 43 / 41 / 36	1.6 / 1.9 / 1.7 / 1.7	N	N	24	13	N	11	27	N	0.4	0.6	0.3	N	0.5	0.7	N
M. sinensis Tea-88	Clumped	6	Y	1.2	N	130	270	N	11	52 / 57 / 55 / 63	2.2 / 2.2 / 2.3 / 2.3	Y	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-68	Clumped	6	Y	1.1	N	170	225	N	10	59 / 58 / 61 / 61	1.9 / 1.9 / 1.8 / 1.6	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-69	Clumped	12	Y	1	N	150	110	N	13	66 / 66 / 54 / 59	1.8 / 1.8 / 1.5 / 1.8	Y	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-70	Clumped	10	Y	0.9	N	140	90	N	6	44 / 52 / 60 / 43	1.2 / 1.3 / 1.4 / 1.2	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-71	Clumped	8	Y	1	N	130	48	N	9	61 / 62 / 54 / 58	1.8 / 1.4 / 1.5 / 1.5	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-72	Clumped	4	Y	0.7	N	150	150	N	10	64 / 60 / 61 / 65	2.2 / 2.4 / 2.2 / 2.5	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-73	Clumped	8	N	0.5	N	90	80	N	6	35 / 37 / 32 / 30	1.2 / 1 / 1.2 / 1.2	Y	2	15	5	N	11	16	N	0.5	0.4	0.1	N	0.5	0.6	0.6
Miscanthus sp. Tea-45	Spread	13	N	0.5	N	50	28	N	6	31 / 29 / 35 / 28	2 / 1.9 / 1.9 / 2	N	2	28	10	N	17	15	N	0.6	0.6	0.3	N	0.5	0.7	0.7
Miscanthus sp. Tea-47	*	*	N	0.4	N	50	7	N	2	39 / 45 / 40 / 32	1 / 1 / 0.9 / 1	Y	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-48	*	*	*	*	*	*	*	N		51 / 61 / 56 / 58	1.3 / 1.7 / 1.8 / 1.3	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-49	*	*	*	*	*	*	*	N		45 / 52 / 52 / 43	0.9 / 1.4 / 1.2 / 1.5	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-50	*	*	Y	0.8	N	70	8	N	4	49 / 58 / 56 / 58	1.2 / 1.7 / 1.6 / 2.3	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-51	*	*	*	*	*	*	*	N		49 / 58 / 66 / 44	1.1 / 1.6 / 1.6 / 1.5	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-52	*	*	*	*	*	*	*	N		61 / 60 / 69 / 64	1.6 / 1.8 / 1.7 / 1.5	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-53	Clumped	8	N	0.6	N	50	12	N	3	38 / 44 / 40 / 45	0.7 / 0.8 / 1.4 / 1.1	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp.Tea-54	*	*	*	*	*	*	*	N		55 / 44 / 50 / 46	1.2 / 1 / 1.1 / 1.4	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-6	Clumped	10	N	1.1	N	80	80	N	7	55 / 60 / 60 / 66	1.6 / 2.3 / 1.6 / 1.8	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-11	Clumped	6	Y	1.1	Y	190	310	N	17	67 / 67 / 61 / 71	2.6 / 2.3 / 2.3 / 2.3	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-21	Clumped	7	Y	0.9	Y	200	200	N	10	59 / 70 / 65 / 60	2.4 / 2.1 / 2.6 / 2.5	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*

Table D (*continued*)

IID	Habit (clumped, Spread)	Space between culms (cm)	Culm wax (Y/N)	Max culm width	Culm buds or branching (Y/N)	Plant height	Leaf number	Leaf variegation	Internode length (cm)	Max leaf length (cm)	Max leaf width (cm)	Leaf hairs	INFLORESCENCE	Inflorescence length	Inflorescence axis length	Axis hairs	Raceme length	Raceme number	Raceme axis hairs	Raceme internode	Upper pedicel length	Lower pedicel length	Pedicel hairs	Spikelet length	Spikelet hairs length	Awn length
Miscanthus sp. Tea-22	Clumped	4	Y	0.7	N	120	400	N	5	67; 59; 68; 71	2; 2.1; 2.2; 1.8	Y	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-27	Clumped	4	N	0.6	N	130	80	N	10	32; 41; 36; 34	1.5; 1.6; 1.6; 1.5	N	4	23	11	N	15	20	N	0.6	0.5	0.3	N	0.5	0.7	0.8
Miscanthus sp. Tea-28	Clumped	3	Y	0.9	Y	190	210	N	16	68; 71; 60; 61	2.1; 2.3; 2; 2.3	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-23	Clumped	4	Y	0.8	N	80	90	B	8	52; 54; 53; 49	0.9; 0.9; 0.9; 0.8	Y	N	37	19	N	21	41	N	0.7	0.6	0.3	N	0.6	0.8	0.4
Miscanthus sp. Tea-29	Spread	25	Y	0.7	N	50	25	N	3.5	38; 34; 33; 30	1.2; 1.3; 1; 1.2	Y	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp.Tea-37	Clumped	5	Y	0.5	N	100	270	N	9	53; 53; 51; 50	1.5; 1.6; 1.6; 1.7	Y	N	30	16	N	17	43	N	0.5	0.6	0.3	N	0.5	0.5	0.4
Miscanthus sp. Tea-38	Clumped	4	Y	0.5	N	60	80		4	*; *; *; *	*; *; *; *	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp.Tea-39	Clumped	5	N	0.5	N	120	50	*	15	*; *; *; *	*; *; *; *	*	*	26	11	N	16	24	N	0.6	0.5	0.2	N	0.4	0.7	0.5
Miscanthus sp. Tea-41	Clumped	5	N	1	N	150	70	*	5	*; *; *; *	*; *; *; *	*	*	26	14	N	11	35	N	0.8	0.7	0.3	N	0.5	0.6	0.6
Miscanthus sp. Tea-42	Clumped	5	N	0.6	N	100	160	*	11	*; *; *; *	*; *; *; *	*	*	28	11	Y	17	20	N	0.7	0.4	0.1	N	0.5	0.8	0.6
Miscanthus sp. Tea-43	Clumped	4	Y	0.8	N	120	120	*	8	*; *; *; *	*; *; *; *	*	*	38	20	N	17	36	N	0.6	0.7	0.3	N	0.5	0.8	0.5
M. sacchariflorus Tea-84	Clumped	10	Y	1	Y	170	110	*	10	*; *; *; *	*; *; *; *	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-1	Clumped	5	Y	1	Y	220	300	N	15	64; 57; 55; 65	2.2; 2; 2.1; 1.8	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-10	Clumped	4	N	1.1	N	140	180	N	7	70; 70; 74; 60	2.3; 2.5; 2.3; 1.8	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-18	Clumped	6	Y	1.2	N	110	210	N	10	45; 53; 53; 47	1.7; 1.8; 1.5; 1.5	Y	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp.Tea-55	Clumped	6	Y	0.9	Y	190	480	N	10	62; 58; 60; 61	2.1; 1.9; 2.1; 2.1	Y	1	*	*	*	*	*	*	*	*	*	*	*	*	*
M. sinensis 'zebrinus' Tea-2	*	*	*	*	*	*	*	N		48; 48; 41; 52	1.3; 1.3; 1.4; 1.4	Y	6	*	*	*	*	*	*	*	*	*	*	*	*	*
M. sinensis 'zebrinus' Tea-3	Clumped	3	N	0.8	N	120	210	N	5	41; 47; 36; 45	1.9; 1.7; 1.7; 1.7	Y	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. sinensis 'zebrinus' Tea-33	Clumped	4	Y	0.8	N	110	320	N	6	39; 40; 40; 37	1.7; 1.8; 1.7; 1.8	Y	1	35	15	N	18	48	N	0.5	0.6	0.2	N	0.6	0.8	0.4
Miscanthus sp. Tea-7	Clumped	7	N	1.1	N	160	140	N	10	35; 31; 38; 42	1.5; 1.5; 1.4; 1.5	Y	4	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-9	Clumped	4	Y	1.4	Y	190	325	N	15	31; 31; 35; 34	1.8; 2; 1.8; 1.7	Y	8	*	*	*	*	*	*	*	*	*	*	*	*	*
M. sinensis Tea-13	*	*	*	*	*	*	*	N		38; 39; 36; 34	1.4; 1.4; 1.4; 1.6	Y	2	*	*	*	*	*	*	*	*	*	*	*	*	*
M. sinensis Tea-14	*	*	*	*	*	*	*	N		42; 49; 45; 54	2.1; 2.4; 2.2; 2	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
Miscanthus sp. Tea-15	Clumped	4	Y	1	Y	200	180	N	11	45; 52; 45; 42	1.8; 1.9; 1.9; 1.4	Y	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. sinensis goliath-like Tea-19	Clumped	2	Y	0.8	N	150	240	N	6	49; 50; 52; 47	1.6; 1.7; 1.6; 1.6	Y	1	37	17	N	19	44	N	0.4	0.4	0.2	N	0.5	0.8	0.5

ID	Habit (clumped, Spread)	Space between culms (cm)	Culm wax (Y/N)	Max culm width	Culm buds or Yanching (Y/N)	Plant height	Leaf number	Leaf variegation	Internode length (cm)	Max leaf length (cm)	Max leaf width (cm)	Leaf hairs	INFLORESCENCE	Inflorescence length	Inflorescence axis length	Axis hairs	Raceme length	Raceme number	Raceme axis hairs	Raceme internode	Upper pedicel length	Lower pedicel length	Pedicel hairs	Spikelet length	Spikelet hairs length	Awn length
Miscanthus sp. Tea-24										47	2.3															
										45	2.1															
										52	2.1															
	Spread	13	Y	0.6	N	110	150	N	12	48	2	Y	2.5	28	11	N	13	30	N	0.4	0.4	0.1	N	0.4	0.8	0.6
Miscanthus sp. Tea-25										59	1.6															
										51	1.4															
										51	1.7															
	Clumped	6	N	0.4	N	100	90	N	6	43	1.1	Y	N	21	4	N	14	15	N	0.5	0.6	0.3	N	0.6	0.6	0.5
Miscanthus sp. Tea-26										39	1															
										42	1															
										41	1.2															
	Clumped	3	Y	0.5	N	130	200	N	15	46	1.2	Y	N	23	13	Y	12	15	N	0.6	0.5	0.2	N	0.5	0.7	0.6
M. sinensis Tea-30										36	1.1															
										38	1.2															
										36	1.1															
	Clumped	5	Y	0.8	N	120	200	N	9	36	1.1	Y	1.5	33	17	N	19	37	N	0.5	0.6	0.3	N	0.5	0.7	0.4
M. sinensis 'gross fontane' Tea-35										39	1.6															
										33	1.5															
										44	1.5															
	Clumped	10	Y	0.6	N	90	90	N	5	38	1.4	Y	3	28	15	N	13	33	N	0.5	0.6	0.2	N	0.5	0.7	0.6
M. sinensis 'gross fontane' Tea-36										40	1.5															
										38	1.5															
										30	1.4															
	Spread	7	Y	0.9	N	90	90	N	7	32	1.6	Y	3	*	*	*	*	*	*	*	*	*	*	*	*	*
M. sinensis Tea-40										42	2															
										47	2.3															
										55	2.2															
	Spread	10	Y	0.7	N	90	250	N	9	50	2	Y	5	31	17	N	17	50	N	0.6	0.5	0.2	N	0.5	0.5	0.5
M. sinensis 'goliath' Tea-56										34	1.1															
										34	1.2															
										32	1.2															
	Clumped	6	Y	0.8	N	110	90	N	12	27	1	Y	1	35	15	N	16	51	N	0.5	0.6	0.2	N	0.6	0.8	0.4
M. sinensis 'goliath' Tea-57										46	1.7															
										54	2.1															
										40	2															
	Clumped	6	Y	0.9	N	110	49	N	13	48	2	Y	1	32	18	N	15	36	N	0.5	0.5	0.2	N	0.6	0.7	0.5
M. sinensis 'sirene' Tea-58										41	2															
										36	1.8															
										40	2															
	Clumped	5	Y	0.9	N	80	140	N	5	41	1.7	N		30	13	N	16	33	N	0.6	0.5	0.2	N	0.5	0.7	0.3
M. sinensis 'strictus' Tea-59										44	1.8															
										44	2.2															
										46	2.1															
	Clumped	1	Y	0.3	N	50	100	N	5.5	43	1.8	N	2.5	*	*	*	*	*	*	*	*	*	*	*	*	*
M. sinensis 'strictus' Tea-60										29	1.5															
										29	1.6															
										25	1.5															
	Clumped	5	N	0.8	N	120	275	N	3	37	1.2	N	1	*	*	*	*	*	*	*	*	*	*	*	*	*
M. sinensis 'malaparteus' Tea-61										40	1.6															
										38	1.8															
										40	1.8															
	*	*	*	*	*	*	*	N		37	1.4	N	8	*	*	*	*	*	*	*	*	*	*	*	*	*
M. sinensis Tea-62										33	1.4															
										32	1.6															
										38	1.5															
	*	*	*	*	*	*	*	N		29	1.7	N	5	*	*	*	*	*	*	*	*	*	*	*	*	*
M. sinensis Tea-76										28	1.5															
										38	1.5															
										35	1.5															
	*	*	*	*	*	*	*	N		34	1.6	N	3	*	*	*	*	*	*	*	*	*	*	*	*	*
M. sinensis Tea-77										50	1.9															
										48	2															
										50	2															
	Clumped	3	N	0.5	N	120	30	N	6	46	1.9	N	N	18	5	Y	13	13	N	0.5	0.5	0.2	N	0.5	0.9	0.6
M. sinensis Tea-78										41	1.6															
										41	1.7															
										34	1.6															
	*	*	*	*	*	*	*	N		40	1.4	N	1	*	*	*	*	*	*	*	*	*	*	*	*	*
M. sinensis Tea-79										58	2															
										52	2.1															
										60	2.1															
	Clumped	2	N	0.6	N	120	150	N	14	54	1.8	N	3	23	8	Y	15	27	N	0.5	0.4	0.1	N	0.4	0.6	0.5
M. sinensis Tea-80										47	1.4															
										44	1.3															
										44	1.4															
	Clumped	5	N	0.4	N	110	60	N	11	44	1.5	N	5	19	4	Y	12	10	N	0.4	0.4	0.1	N	0.5	0.6	0.7
Miscanthus sp. Tea-16										52	1.4															
										50	1.6															
										59	1.5															
	*	*	*	*	*	*	*	N	*	59	1.6	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. condensatus Tea-44										43	2.1															
										44	1.5															
										48	2															
	*	*	*	*	*	*	*	N		42	1.5	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*
M. sacchariflorus Tea-129										59	2.3															
										56	2.5															
										54	2.3															
	Spread	20	Y	0.6	Y	90	40	N	5	63	2.2	N	N	*	*	*	*	*	*	*	*	*	*	*	*	*

Table E Haplotype information obtained with cpSSRs. Count= number of accession sharing the haplotype.

Haplotype Code	Count	Sac-2	Sac-3	Sac-10	Sac-13	Sac-17	Sac-26
1	36	255	269	275	288	230	176
2	34	252	266	277	290	227	175
3	16	256	269	275	288	230	176
4	5	252	265	277	289	227	175
5	3	252	266	277	289	227	175
6	3	255	268	275	288	230	175
7	3	255	-1	275	288	230	176
8	2	252	269	284	288	229	177
9	2	253	266	273	287	229	176
10	2	255	270	275	288	230	176
11	1	-1	265	276	288	227	175
12	1	-1	266	277	288	-1	175
13	1	236	265	294	290	229	178
14	1	247	-1	283	283	234	174
15	1	247	-1	286	-1	229	176
16	1	248	-1	278	-1	-1	169
17	1	248	-1	279	285	230	169
18	1	248	266	286	285	228	176
19	1	248	266	286	285	229	176
20	1	249	265	-1	-1	228	-1
21	1	249	264	273	-1	229	176
22	1	249	262	278	-1	234	174
23	1	250	265	275	288	229	176
24	1	250	264	283	-1	229	176
25	1	250	-1	283	287	228	176
26	1	250	262	284	288	229	175
27	1	250	272	285	290	227	175
28	1	251	-1	-1	-1	-1	-1
29	1	251	-1	-1	-1	228	175
30	1	251	263	269	284	238	174
31	1	251	-1	273	-1	228	175
32	1	251	262	273	283	217	175
33	1	251	263	273	286	227	175
34	1	251	265	273	286	229	176
35	1	251	267	275	-1	227	176
36	1	251	267	275	287	230	176
37	1	251	267	275	287	234	176
38	1	251	267	275	287	235	176
39	1	251	264	275	288	227	175
40	1	251	269	275	288	230	176
41	1	251	267	275	288	234	176
42	1	251	267	275	290	-1	175
43	1	251	267	276	287	-1	176
44	1	251	267	276	287	234	175
45	1	251	267	276	287	234	176
46	1	251	265	276	288	229	176
47	1	251	265	287	291	228	175
48	1	252	266	-1	290	227	175
49	1	252	268	275	-1	228	175
50	1	252	-1	275	287	230	176
51	1	252	-1	275	288	227	176
52	1	252	269	275	288	230	176
53	1	252	267	276	287	234	176
54	1	252	264	276	288	227	176
55	1	252	265	276	288	227	176
56	1	252	266	276	290	227	175
57	1	252	266	277	-1	227	175
58	1	252	266	277	290	-1	175
59	1	252	-1	277	290	227	175
60	1	252	266	277	290	229	175
61	1	252	267	284	291	228	176
62	1	252	263	289	291	229	176
63	1	253	269	275	283	229	176
64	1	253	266	275	287	229	176
65	1	253	269	275	288	230	176
66	1	253	267	275	291	228	176
67	1	255	-1	-1	288	227	175
68	1	255	-1	-1	288	230	176
69	1	255	269	271	288	230	176
70	1	255	268	274	288	230	175
71	1	255	269	275	-1	230	-1
72	1	255	266	275	-1	230	176
73	1	255	269	275	288	229	176
74	1	255	269	275	288	230	172
75	1	255	269	275	288	230	175
76	1	255	266	275	288	230	176
77	1	255	269	275	288	230	177
78	1	256	269	-1	288	230	175
79	1	256	266	-1	290	227	175
80	1	256	269	274	288	230	176
81	1	256	269	275	-1	230	176
82	1	256	269	275	288	229	175
83	1	256	-1	275	288	230	176
84	1	256	269	275	288	231	176
85	1	257	-1	279	-1	230	175

157

Table F List of 80 nSSRs developed from a microsatellite enriched library for *Miscanthus*

Primer name	SSR motif	SSR size	Forward sequence	Reverse sequence	PCR length (bp)
Mis-01	(TCTA)20	80	cagtccttggagcaggctat	aagatctcaaacctatagtc	202
Mis-02	(TATC)17	68	actttacaaaacaaacacac	cgaggcgcaagagtcaccat	307
Mis-03	(AGAT)17	68	acgactgactatacgccatcaa	ttgtatcactgtgcaagtgt	257
Mis-04	(AGAT)20	80	ttactactgagatcaaagca	aattattgttcgttggctga	281
Mis-05	(ATAG)16	64	taaggaggcctaatatccctt	tcgttggttacatcggcatg	253
Mis-06	(GATA)15	60	tatggttatgacttgtcaga	caagcaagttactgaaccta	200
Mis-07	(AGAT)26	104	gacaacccatttactactga	tacggcttataagccaagcg	305
Mis-08	(CATA)60	240	catattccatacatgcatgc	cagcttttctaagagtagtg	283
Mis-09	(TATC)17	68	tttacaaaacaaacacacag	cgccgacaattaaccatatgc	219
Mis-10	(TATC)19	76	gctgcgtagctattgcatca	cttagatcaatatctcaaat	304
Mis-11	(TACA)26	104	tgagctgctgaacttgtcag	cgtttgcaatagtgtcgatcaa	344
Mis-12	(AGAT)18	72	ttactactgagatcaaagca	tacggcttataagccaagcg	263
Mis-13	(TAGA)19	76	cggactaacttgtgaatctt	gtccttggagcaggctatga	230
Mis-14	(GATA)15	60	gtagctgcaactgctagtgt	actcgcattggttggtatga	141
Mis-15	(ATCT)16	64	actactgcatgcatcatgatg	tgcttcgccggcgaagtttca	195
Mis-16	(TATC)13/(TCTA)16	52+64	atcttgcctaggatgcattag	tggtctattacaacaaggct	264
Mis-17	(ATAC)17	68	acgctagctgatggacca	tggcgacctctgagcacagc	226
Mis-18	(ATAG)15	60	tcagccgatcgatggattag	ctaccgagcatgcaagta	166
Mis-19	(ATAG)16	64	atcaatatctcaaatcacat	agcggcgagcagctcgttg	243
Mis-20	(TCTA)17	68	tagctgagctgtctatggta	tagccattgaggctaaggat	249
Mis-21	(AGAT)16	64	caggccacatgtccatgcac	ctactgcatgcatcatgatg	169
Mis-22	(TAGA)17	68	cgagcgagcctgcatgtgtg	ttgacgtcagcaagatattg	173
Mis-23	(ATCT)15	60	cacgaactgaatcagcatgc	gtagctgcaactgctagtgt	240
Mis-24	(AGAT)15	60	atacacgatccaaacatgtc	atgtgctcacccaagagatg	324
Mis-25	(ATAG)15	60	atatctcaaatcacatctaag	gttggttacatcggcatgtc	226
Mis-26	(TG)16(AG)22	76	atgtttcctgtttccatcag	acattaggcatttcgccatt	277
Mis-27	(AG)38	76	tcaaccattgctcctggatg	tgtattggtgcaactgcaag	249
Mis-28	(AC)10(GA)25	70	cggcgctccgcgatgctcag	tgacggtaacatctaagttc	266
Mis-29	(TC)24	48	taacaacttagccaaggata	atcttaattaggagtcactg	245
Mis-30	(AG)29	58	gtaatttctcgtatgctgcg	cgtggccggtagcttggtgc	282
Mis-31	(AG)20	40	atccaacagtgataggacgt	cctagacccacttggacgat	198
Mis-32	(CT)22	44	aggtgattcagtctcaggtta	ttgcttgcggattagatgg	239
Mis-33	(CT)20	40	tgacataggctacacatat	cgagtgaggcagctagttca	242
Mis-34	(TC)25	50	cgcactgccatggccaggac	tgagctggccacagagttgacc	223
Mis-35	(GA)30	60	gtcgaactctaatctaggca	ctgcattagcaagctttagg	175
Mis-36	(TC)11c(CT)21	65	agcactgcatgccttcagat	ggttttatatgattcagcat	231
Mis-37	(TC)34	68	gaatgcagtcatcagcagct	tggacatctctaggttgatc	218
Mis-38	(CT)40	80	cttgatcagaaggtatgctt	agacctctggtaatttgtag	212
Mis-39	(GA)22	44	taaggtagctactcacagac	cgcccacgcccagcaaggca	230
Mis-40	(AG)24	48	agtgttaacacgagtcactg	gaaatactatgctgtgagtg	222
Mis-41	(GA)24	48	ataatgcaggtcagttcaac	cgcagctagctgcttgtcag	226
Mis-42	(AG)31	62	gccgccaggctcccaagcct	atccgagccatgtatgcacg	206
Mis-43	(CT)27	54	agcatgcatggctgctgagc	tgagctagtctgcatgcatc	271
Mis-44	(CT)28	56	ggtgcccaacatacacaatc	cgttcgatcaagcgatgaac	148
Mis-45	(AG)26	52	acaaaacggaatccttgaca	gccagcatcgtgctcatcgc	221
Mis-46	(GA)53	106	tagcaccgcctgttccctga	ctaatgaatccaagacatttac	241
Mis-47	(GA)28	56	atggaagcagcatggcttcg	ataggagttctcgactcacc	179
Mis-48	(CT)30	60	catgtatgcacggcagcacg	cgccaggctcccaagcctaa	194
Mis-49	(GA)26	52	catggtttatcagccaagcg	ggaatatgcctggctccctg	229
Mis-50	(GA)21	42	tacggacgattaaccaagcc	cgcaaggtgcaggaccatca	230
Mis-51	(TC)20	40	gatccatcacggattcatca	atcataggcaaaacggatcg	164
Mis-52	(GA)19	38	ttattggtgcccaaaggtgt	aacaagccctcaagcttcct	370
Mis-53	(GA)19	38	aggcagcacctcacaaaact	ggtggagatgctcttcttgc	173

Table F (*continued*)

Primer name	SSR motif	SSR size	Forward sequence	Reverse sequence	PCR length (bp)
Mis-54	(CT)18	36	taagaaacgcagcagcagaa	agtctccggcttctcacaa	226
Mis-55	(GA)18	36	cggcttcgagtgatacctt	taccggatttaaggggcttt	250
Mis-56	(CT)18	36	gctagtcttgcctgctgctt	gccatggaagtcatggttct	208
Mis-57	(GA)17	34	tcgcgtaatgcgtcttgtta	gcacacagtcaccactcacc	171
Mis-58	(GA)17	34	tgacagtcattgctccttgc	ctctcccatccttccctctc	243
Mis-59	(GA)16	32	gagctgatcgcgtagcaag	ttcgataaacaggggattgg	152
Mis-60	(GA)16	32	agatggcagcttgctcttgt	ccatttgttgagcacgatgt	190
Mis-61	(TATC)15	60	cccaagagatggatggaaaa	gcttgataaaatgccgggta	226
Mis-62	(TCTA)14	56	catgaattgaggacagggaag	gagccccaaagtgaaacgat	207
Mis-63	(TCTA)14	56	aggctagcacttcctccaaa	ctgcctggtgacccctataa	234
Mis-64	(AGAT)14	56	tcccttagtgtccgtgaag	gaggcaggtgtagtcggaga	236
Mis-65	(AGAT)13	52	acgacgccttagcatgtctt	gtgcagtttgcatctgtgct	245
Mis-66	(CTAT)13	52	catggctacaggcacctaaaa	ataacgagaaaatggccgatg	165
Mis-67	(TCTA)13	52	cctctgcggatatgaggtgt	gaagtgacaacatgcgatgg	175
Mis-68	(AGAT)13	52	acgacgccttagcatgtctt	gtgcagtttgcatctgtgct	245
Mis-69	(TCTA)13	52	cctctgcggatatgaggtgt	gaagtgacaacatgcgatgg	175
Mis-70	(TATC)12	48	tcgcacctttaattttgcat	ttatgaacccgacagggaga	249
Mis-71	(TAGA)12	48	caaccatgagcacttctcca	aacataggaggccaagcaaa	179
Mis-72	(TATC)12	48	aagaggccacaatcaaatgc	cgtcaaccaataacgagtagca	203
Mis-73	(TAGA)11	44	cggtctcttggacgatttgt	cgccaaatctcgtatgtatagaa	246
Mis-74	(AG)16	32	agccagtggttagacggatg	tgttttcctgcaaactttctca	175
Mis-75	(TC)15	30	atcttagcccttccgactgg	tcgtacccctcactcctcac	485
Mis-76	(CT)15	30	cccggctacaataatggtgt	ggctccatttcgtttgttga	155
Mis-77	(AG)15	30	ctgcagtacattgcaggatca	tacggggcatagagttacgg	187
Mis-78	(CT)15	30	tctgcaggtgacaaggaaga	gtcaaccggcatagttcgat	167
Mis-79	(CT)15	30	gccaactcgtggatttgagt	cgtagcaagaggggaacaaa	248
Mis-80	(GA)14	28	ggcttgatccttcacttggt	cttgctcttccacccttgtcc	240