Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

SCIENZE E TECNOLOGIE AGRARIE, AMBIENTALI E ALIMENTARI

Ciclo 28

Settore Concorsuale di afferenza**: 07/E1**

Settore Scientifico disciplinare**: AGR/07**

TITOLO TESI

Physical mutagenesis in giant reed (*Arundo donax* L.) and phenotypic and genomic characterization of mutagenized clones

Presentata da**: Fabio Valli**

Coordinatore Dottorato — **Prof. Giovanni Dinelli**

Relatore — **Prof. Silvio Salvi**

Esame finale anno 2017

CONTENTS

ABSTRACT

Giant reed (*Arundo donax*L.) is a C_3 perennial, warm-season, rhizomatous grass of emerging interest for bioenergy and biomass-derivatives production, and for phytoremediation. It only propagates vegetatively and very little genetic variation is found among ecotypes, basically precluding breeding efforts. With the objective to increase the genetic variation in this species, we developed and applied a mutagenesis protocol based on γ-irradiation of *in-vitro* cell cultures from which regenerants were obtained. Based on a radio-sensitivity test, the irradiation dose reducing to 50% the number of regenerants per callus (RD_{50}) was estimated at 35 Gy. A large mutagenic experiment was carried out by irradiating a total of 3,120 calli with approx. 1x, 1.5x and 2x RD_{50}. A total of 1,004 regenerants from irradiated calli were hardened in pots and transplanted to the field. Initial phenotypic characterization of the collection showed correlated responses of biomass-related quantitative traits to irradiation doses. Approx. 10% of field grown clones showed remarkable morphological aberrations including dwarfism, altered tillering, abnormal inflorescence, leaf variegation and others, which were tested for stability over generations. Clone lethality reached 0.4%. Our results show for the first time that physical mutagenesis can efficiently induce new genetic and phenotypic variation of agronomic and prospective industrial value in giant reed. The methodology and the plant materials described here may contribute to the domestication and the genetic improvement of this important biomass species.

100 clones of *A. donax* were chemically analyzed for several key chemical components of plant biomass including lignin, cellulose, hemicelluloses, non-structural carbohydrates, ash, and others.

The complete chloroplast genome of *A. donax* is 139353 bp (bais pairs) in length subdivide in two inverted repeat region (IRa and IRb) of 22,227 bp each separated by a small-single-copy-region of of 12,275 bp (SSC) and a large-single-copy-region of 82,124 bp (LSC). The genome includes 112

individual genes including 72 protein coding genes, 30 tRNA, 6 rRNA, 3 open reading frames and one pseudogene.

FISH and GISH analysis were performed on three species *A. donax*, *A. plinii* and *P. australis* in order to define the genetic structure of these species and the phylogenetic relationship existing among them.

CHAPTER 1.

In-vitro physical mutagenesis of giant reed (*Arundo donax* L.)

1.1. INTRODUCTION

Giant reed *(Arundo donax* L.) is a wild perennial rhizomatous C_3 Poaceae species of emerging interest for bioenergy production. It is characterized by very low requirements in terms of cultivation inputs, it adapts to a large range of soil and/or climatic conditions (eg. from semi-arid to water logging-prone environments), and has unusually high photosynthetic and carbon accumulation capacity, making it one of the most promising biomass crops for the Mediterranean area (Corno *et al*., 2014; Ge *et al.,* 2016; Monti & Zegada-Lizarazu, 2016; Webster *et al*. 2016). Biomass produced by giant reed has been used in industrial cellulose production and for thermo chemical conversion processes and more recently has attracted interest for production of bioethanol and other liquid biofuels, biomethane (by anaerobic digestion), and added-value bioproducts (Pilu *et al.,* 2014; Ge *et al.,* 2016). Giant reed was also used for riverbank consolidation and shows potential for soils and water phytoremediation (Nsanganwimana *et al.,* 2013).

Giant reed is traditionally considered as native of sub-tropical Eurasia (including the Mediterranean, Middle East and Northern India regions. Hardion *et al*., 2014). However, a relatively recent colonization of the Mediterranean region by one or very few invasive clones of Asian origin was also suggested (Mariani *et al.,* 2010, Hardion *et al.,* 2014). Different studies (reviewed in Bucci *et al.,* 2013) reported a range of chromosome numbers (2n from 40 to 110), although two cytotypes seemed to prevail: 2n = 108 in both Europe and Asia, and 2n = 72 in Asia only (Hardion *et al.,* 2014). The most likely base numbers appear to be x = 6 or 12, making giant reed highly polyploid (Saltonstall *et al.,* 2010; Hardion *et al.,* 2015). With the exception of few reports of seed reproduction (Perdue, 1958; Bor, 1970; Brach & Song, 2006), the species

appears sterile and only propagates vegetatively by rhizome and shoot fragmentation during flooding events, and by shoot layering (Boland, 2006; Ceotto & Di Candilo, 2010; Saltonstall *et al.,* 2010). It is unclear whether its sterility is caused by polyploidy (Bucci *et al.,* 2013) or self-incompatibility (Hardion *et al.,* 2015). Besides two recent transcriptome analyses (Sablok *et al.,* 2014; Barrero *et al.,* 2015), no sequencing-based genome investigation has been tempted in giant reed.

Likely as a consequence of both its propagation mechanism and history of geographical diffusion giant reed is a species with reduced diversity as amply shown by molecular markers-based investigations (Khudamrongsawat *et al.,* 2004; Ahmad *et al.,* 2008; Mariani *et al.,* 2010; Hardion *et al*., 2012; Pilu *et al.,* 2014) and by field experiments comparing ecotypes for traits of agronomic importance (Cosentino *et al.,* 2006). However, stable phenotypic differences among ecotypes were more recently detected for traits such as phenology (Cantaluppi *et al.,* 2016), cold, drought or salinity tolerance (Pompeiano *et al.,* 2013; Sanchez *et al*. 2015; Haworth *et al.,* 2016) and biomass and bioenergy production (Pilu *et al.,* 2014; Amaducci and Perego, 2015).

Mutagenesis is one of the most effective tools available for the genetic improvement of annual and perennial crop species and has contributed to the development of thousands of cultivars worldwide (Bado *et al.,* 2015; Bradshaw, 2016). Among the different types of mutagenic treatments, irradiation has been the most frequently used, with a predominance of γ-rays (Mba, 2013). Seeds are the preferred target for mutagenic treatments, however, buds, shoot apices and *in-vitro* cultured tissues can be treated in vegetatively propagated species (Predieri *et al.,* 2001; Bado *et al.,* 2015). Mutagenesis has already been applied to biomass species such as miscanthus (Lee *et al.,* 2012; Perera *et al.,* 2015), poplar (Douglas, 1986) and switchgrass (Wang *et al.,* 2013) and *in-vitro* physical mutagenesis has been applied to sugarcane (Nikam *et al.,* 2014).

1.2. OBJECTIVE

In-vitro physical mutagenesis is a well established strategy for increasing genetic variability for breeding purposes. Although giant reed (*A. donax* L.) is a species devoted of genetic diversity, to the best of our knowledge mutagenesis (including physical mutagenesis) has never been attempted.

The objective of this work was to establish an efficient mutagenesis protocol of giant reed based on γ-ray irradiation of *in-vitro* tissues (calli), with the final aim to produce novel and useful genetic variation for the genetic improvement of the species. A first description of field-grown giant reed morphological mutants are presented.

1.3. MATERIALS AND METHODS

1.3.1. Plant material and micropropagation protocol

A single wild plant of giant reed characterized by vigorous growth was collected in the Po valley region (Italy) and used as source of *in vitro*-propagated meristematic tissues. A giant reed-specific micro propagation protocol was developed to induce and establish stable multiplication of undifferentiated calli before preparing the tissue for the irradiation treatment. Calli were induced from sterile 2-5 cm long immature inflorescences after a 2-month period incubation on an induction medium (Table S1). Subsequently, induced calli were grown and multiplied on multiplication medium and sub-cultured every four weeks for a period of at least four months before the mutagenic treatment. Organs differentiation of mutagenized calli took place in a differentiation medium (Table S1).

1.3.2. Radiosensitivity test experiment and identification of the 50% reduction dose (RD_{50})

A test experiment was carried out in order to assess the radio-sensitivity of the plant materials and identify the ideal irradiation dose for large-scale mutagenesis. The optimal dose should be characterized by an acceptable compromise between high frequency of induced mutations and high calli regeneration rate, expected to increase and decrease, respectively, by increasing irradiation dose. Viable 5-mm diameter calli were transferred in multiplication media in Petri dishes (20 calli/dish) and immediately irradiated at different doses using a ^{60}Co source Gamma-cell 220 apparatus (MDS Nordion International Inc., ON, Canada). Six dishes were irradiated per dose, corresponding to a total of 120 treated calli/dose. Irradiated calli were cultured onto the multiplication medium for four weeks before transferring them onto the differentiation medium for the regeneration of rooted plantlets via organogenesis. Petri dishes were kept at 25°C with a 16/8 hours light/dark cycle. Calli were sub-cultured onto a fresh differentiation medium every four weeks for a period of 10 months.

An early visual estimate of radio-sensitivity was obtained by periodically scoring calli for their tissue-browning index (TBI), which was defined as the degree of callus browning as compared to a colored reference ladder (Fig. 1a). TBI scores ranged from 0 (perfectly green) to 100 (fully browned). Each callus was scored and the average TBI score per dish and per treatment calculated. TBI scores were estimated at 10, 30, 50, 180 and 300 days after irradiation treatment (DAT).

The rate of viable regenerants (VR) per callus was utilized as the final index of radio-sensitivity. The concept of 50% reduction dose (RD_{50}), as described in Kodym et al. (2012), was utilized in place of the more traditional 50% lethal dose (LD_{50}). RD_{50} indicates the irradiation dose causing a 50% reduction in a biological index (in our case, the number of regenerant per callus) as compared with control (untreated) dose. RD_{50} was estimated after collecting and graphing VR values from the radio-sensitivity test experiment across the seven

doses, and fitting a logistic regression curve (Fig. 1b). Observations on TBI and VR stopped at 300 DAT.

After roots and shoots differentiation, plantlets (1-2 cm height) were transferred to 48-well trays filled with sandy soil for a 4-week period, at 25°C, 16/8 hours light/dark and 95% relative humidity. This time period enabled plantlets to adapt to reduced air humidity and a non aseptical environment. When plants reached 5-cm height they were transferred into 1-liter pots for the final hardening phase at greenhouse conditions.

Based on the results of the radio-sensitivity test experiment, a larger γ-irradiation experiment was carried out. *In-vitro* culture, irradiation procedure, plant regeneration and hardening were performed as described for the test experiment, with the exception that hardened plants were transferred and kept in 12-liter pots for 8 to14 months at greenhouse conditions.

1.3.3. Field evaluation and definition of mutant classes

Mutagenized, independently regenerated hardened plants were randomly (irrespectively of the irradiation dose) transplanted in the field (at Ca' Bosco, near Ravenna, Italy) during spring (April 2015). Eleven control plants derived from untreated calli and subjected to the same regeneration and hardening processes were transplanted along with the irradiated plants. Plants were spaced at 2.5 m x 2.5 m between and within row and cultivated following standard agronomic practice. At the end of the first growing season (November 2015), all shoots were mechanically shredded.

For each of the 1,011 field grown giant reed clones, three biomass-related quantitative traits were collected in October 2015: maximum plant height, number of shoots (or tillering), stem diameter. Maximum plant height was expressed in m and collected by manually measuring the highest shoot for each plot. Number of shoots was obtained by directly (visually) counting all shoots taller than 0.1 m per plot. Stem diameter (expressed in mm) was the average value

from 5-10 representative shoot stems per plot (measurements were taken with a handheld caliper by considering the external stem diameter, between two stem nodes, at 0.5 m height from the soil level).

Putative morphological mutants were identified for the three morphological metric traits described above and were classified in four mutant classes: *short stature, thin stem, high-tillering* and *low-tillering*. For these traits, the putative mutants corresponded with the extreme outlier plants as identified by comparison with the box-plot distributions of control (0 Gy) plants (Fig. 2). Two additional mutant classes were identified, namely *variegate* and *abnormal shoots*. Mutants were classified as *variegate* when showing at least one shoot with clearly variegated leaves or leaf sectors (white or yellow colored) on multiple leaves of the shoot. Mutants were classified as *abnormal shoot* when showing obvious aberrant shoot development in terms of stem, leaves or inflorescences shape and/or architecture. For all these putative mutants, their extreme phenotypic values were confirmed in 2016.

Phenotype stability of five putative morphological mutants (along with an untreated control, for comparison) was further tested by clonal propagation (by rhizome subdivision, five sub-clones per clone) in pots. Propagation was carried out in October 2015, and subclones were cultivated in pots in greenhouse during the winter and transplanted in the field in spring 2016. Observations were collected in pots (winter) and in the field during the second growing season (2016),

1.3.4. Statistical analysis

Analysis of variance was performed to test the effect of different irradiation doses (0, 40, 60 and 80 Gy doses; dose as fixed factor) on biomass-related quantitative traits collected during field evaluation. Trait mean comparisons between doses were conducted using Tukey's test.

All statistical analyses were conducted using R (R Development Core Team, 2015), with Tukey's HSD function and Tuckey's test.

1.4. RESULTS

1.4.1. Development of the mutagenic protocol

The optimal irradiation dose for *in-vitro* physical mutagenesis of giant reed was searched in a test-experiment. Calli treated at ≤ 20 Gy and control (untreated) calli suffered little or no tissue browning and differentiated shoots throughout the whole period of observation (Fig. 1a and Table 1). However, higher irradiation doses (from 40 to 100 Gy) showed a dramatic TBI increase and significantly reduced VR (from 23% to 6%, respectively. Table 1) as compared to untreated control ($P < 0.01$, Tukey's test). Overall, increasing irradiation doses correlated positively with TBI ($r^2 = 0.94$, $P < 0.01$, at 50 days after treatment. Table 1) and negatively with VR ($r^2 = -0.93$, $P < 0.01$), as expected. Based on the above test-experiment, the RD_{50}, namely the irradiation dose corresponding to a 50% reduction of the number of viable regenerated plantlets per callus as compared with untreated control, was estimated to be 35 Gy (Fig. 1b).

Table 1. Effect of different *γ*-irradiation doses on *in-vitro* plant regeneration in giant reed. TBI = Tissue browning index; DAT = Days after irradiation treatment; VR = viable regenerants.

Dose (Gy)	Treated calli	TBI (%) at different DAT					VR [1] (No.)	VR [1] % [2]
		10	30	50	180	300		
0	120	1	1	1	1	1	101	84 (a)

10	120	1	3	6	3	1	103	86 (a)
20	120	3	3	9	7	3	85	71(a)
40	120	3	5	18	47	100	28	23 (b)
60	120	4	8	33	58	100	21	18 (b)
80	120	4	9	57	100	100	9	8 (c)
100	120	4	7	42	100	100	7	6 (c)

[1] Viable regenerants at 300 days after irradiation treatment.

[2] Percentage of viable regenerants per callus. Different letter indicate significantly different values ($P < 0.05$, Tukey's test).

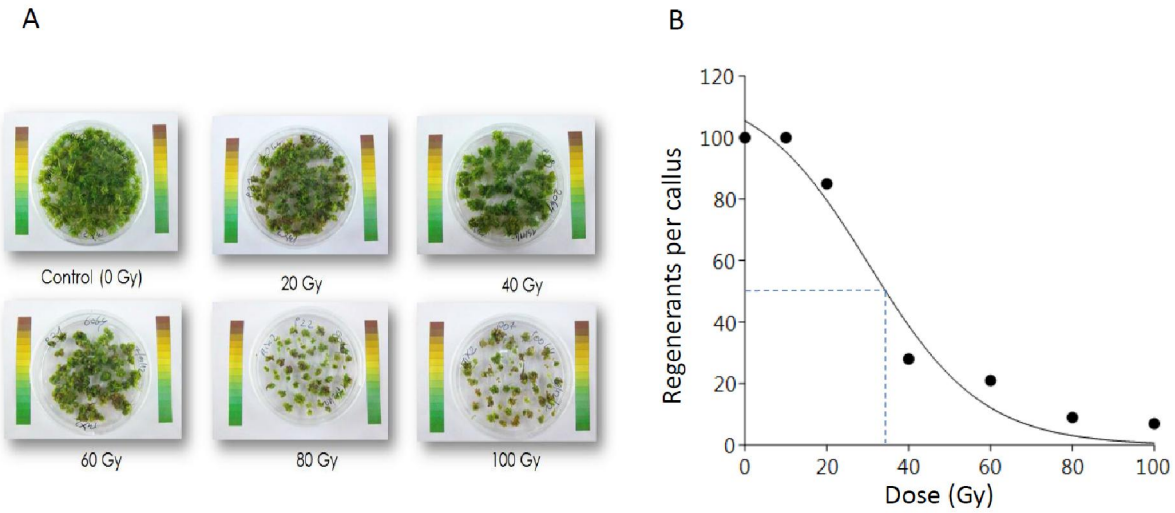


Figure 1. a) Representation of γ-irradiation sensitivity of giant reed calli at 120 DAT (days after treatment) as utilized to derive the tissue-browning index (TBI). Effect of 10 Gy dose is not shown. **b)** Reduction dose curve indicating the radio-sensitivity response of giant reed calli. Curve indicates the reduction in the number of regenerated plantlets per callus (expressed as percentage of control untreated samples), based on seven different γ-irradiation doses (0 to 100 Gy). Dots represent mean values of six replicates. Interpolate curve was fitted using logistic regression. Dashed lines indicate the reduction dose 50% (RD_{50}), namely, the γ-irradiation dose that resulted in 50% reduction in number of regenerated plantlets per callus.

1.4.2. Large-scale mutagenic experiment

Based on the estimated RD_{50}, 40, 60 and 80 Gy doses (corresponding to approx. 1x, 1.5x and 2x RD_{50}) were utilized to treat 1,200, 1,200 and 720 calli, respectively, which produced a total of 1,004 regenerants (Table 2). VR was negatively related to irradiation dose, with values of 34.8, 33.1 and 26.3% for 40, 60 and 80 Gy doses, respectively. As in the test experiment, 40 and 60 Gy doses provided very similar VR, while VR was significantly lower for 80 Gy ($P < 0.05$; Tukey). This notwithstanding, in this experiment VR was generally higher than that observed in the test experiment, so that even the 80 Gy dose provided a sizeable number of plantlets (Table 2).

Table 2. Number of giant reed viable regenerant plants (VR) obtained from the large-scale mutagenesis experiment per dose level.

Dose (Gy)	Irradiated calli (No.)	VR (No.)	VR % [1]
40	1,200	418	34.8 (a)
60	1,200	397	33.1 (a)
80	720	189	26.3 (b)
Total	3,120	1,004	32.2

[1] Percentage of viable regenerants per callus. Different letters indicate significantly different values ($P< 0.05$, Tukey's test).

1.4.3. Effects of irradiation dose on quantitative traits and identification of putative mutants

A collection of 1,015 pot-hardened clones were transplanted in the field (1,004 clones regenerated from irradiated calli and 11 regenerated from untreated calli). Four clones (0.4%) died by the end of the first growing season (October 2015). Thus, phenotypic scores were collected on a total of 1,011 clones.

In order to test whether different irradiation doses had any effect on giant reed phenotypic traits, three shoot architecture traits (maximum shoot height, number of shoots per plant and stem diameter) were analyzed. The three traits were significantly influenced by irradiation dose (Fig. 2; Table S2). Specifically, the number of shoots per clone increased with irradiation dose, from 15.3 shoots for control (untreated) to 21.8, 21.9 and 24.1 shoots in the 40, 60, and 80 Gy treated plants, respectively (Fig. 2a). Conversely, shoot height was negatively related to the irradiation dose, since the maximum shoot height (2.8 m) was observed in controls and decreased progressively to 2.7, 2.5, and 2.2 m, in the 40, 60, and 80 Gy-treated plants, respectively (Fig. 2b). Plants obtained from untreated calli showed significantly larger stem diameter (18.2 mm, $P < 0.05$. Fig. 2c) compared to the irradiated ones, however, no significant difference was observed for stem diameter among 40, 60 and 80 Gy irradiation doses (Fig. 2c).

Based on shoot trait measurement and/or scoring, 93 clones (corresponding to approx. 10% of all treated clones) were identified as putative morphological mutant clones and were preliminarily classified in six classes (Table 3). The class *abnormal shoot* included a curly leaves/inflorescences mutant (Fig. 3b), a mutant characterized by patent leaves and a mutant characterized by shortened, deviated internodes. The class *variegate* included four similar mutants showing different levels and colors of leaf variegation (Fig. 3a); all four variegate mutants were chimeric. *High-tillering* and *low-tillering* classes included six and four mutants showing extremely high or low propensity (Fig. 4a) to develop additional shoots per plot,

respectively. The class *short stature* included 64 clones. This was a relatively heterogeneous class including both weak, stunted plants (Fig. 3c, 4c) and plants with shortened internodes. The *thin stem* class includes 12 putative mutants characterized by stems remarkably thinner than the untreated control (Fig. 4b). No mutant was found with shoot taller or stem larger than the untreated control. As expected, the frequency of putative mutants increased with irradiation dose from 4.8% (40 Gy) to 22.8% (80 Gy; Tab. 3).

In order to further test the stability of the observed phenotypes, five putative mutant clones which were already clearly identified at the end of the first field growing season were vegetatively propagated by rhizome subdivision and phenotypic traits collected after approx. 10 months (Table 4 and Table S2; Fig. 4). All five mutants confirmed to be stable, supporting a genetic basis of the observed phenotypes. Also in this case, the original five mother-plants continued to show the same phenotypes during the second year (checked on September 2016, not shown).

Many additional clones showed strongly modified expression of typical quantitative traits such as flowering time, inflorescence size and architecture, shoot habitus, leaf size and others. However, since variation for these traits is highly quantitative (ie. affected by multiple genes, environmental factors and uncontrolled managing effects related with the regeneration, hardening, and transplantation phases), it will only be possible to confirm these clones as mutants after they will be tested in replicated experiments.

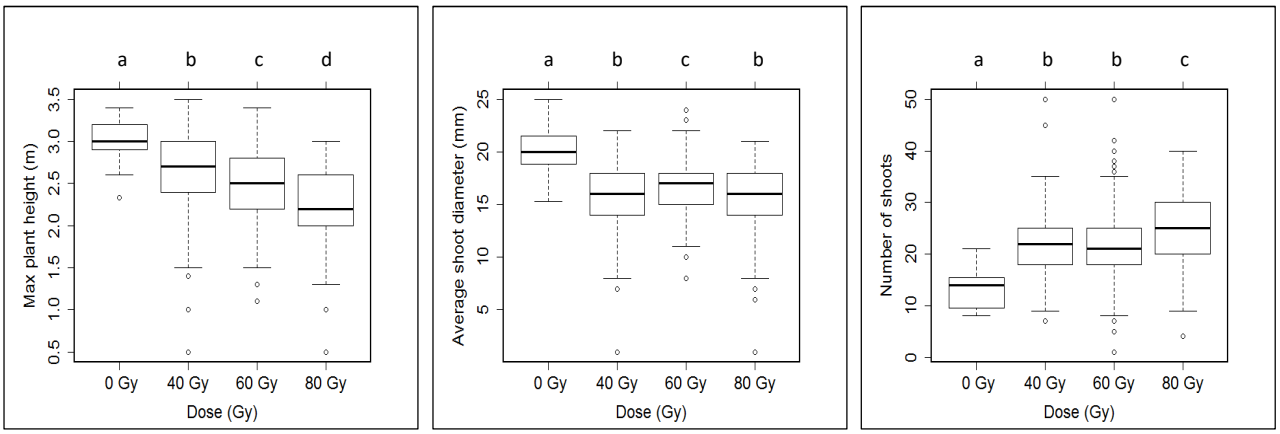


Figure 2. Effects of different *γ*-irradiation doses on giant reed biomass-related traits measured on clones obtained from the large-scale mutagenesis experiment, six months after field-transplantation. Box plots report median, first and third quartile and 95% confidence interval of the median. Different letters indicate different mean values ($P < 0.05$, Tukey's test).

Table 3. Summary of giant reed putative mutant clones for qualitative traits identified after field observations

Mutant class	Clone per irradiation dose (No.)			Total (No.)
	40 Gy	60 Gy	80 Gy	
Abnormal shoot	-	2	1	3
Variegate	1	2	1	4
Short stature	12	22	30	64
High-tillering	1	3	2	6
Low-tillering	1	2	1	4
Thin stem	2	5	5	12
Lethal [1]	*1*	*2*	*1*	*4*
All putative mutants	*18*	*38*	*41*	*97*
All clones	*397*	*418*	*189*	*1004*
% mutant clones	*4.5*	*9.1*	*21.7*	*9.7*

[1] Number of dead clones at September 2016 (15 months after field transplantation).

Table 4. Giant reed mutant clones tested for phenotypic stability. Additional data are provided in Table S2.

Mutant code	Mutant class and observations [1]	Observation level
5-60-10	*Variegate*	Pots
157-40-221	*Thin stem* (and patent leaves)	Pots
22-80-2	*Low tillering* (and tall shoots)	Pots
182-60-17	*Short stature* (and bushy, leafy)	Pots
182-60-60	*Short stature* (and erect leaves, semi dwarf)	Pots

[1] Mutant classes are fully described in Materials and Methods.

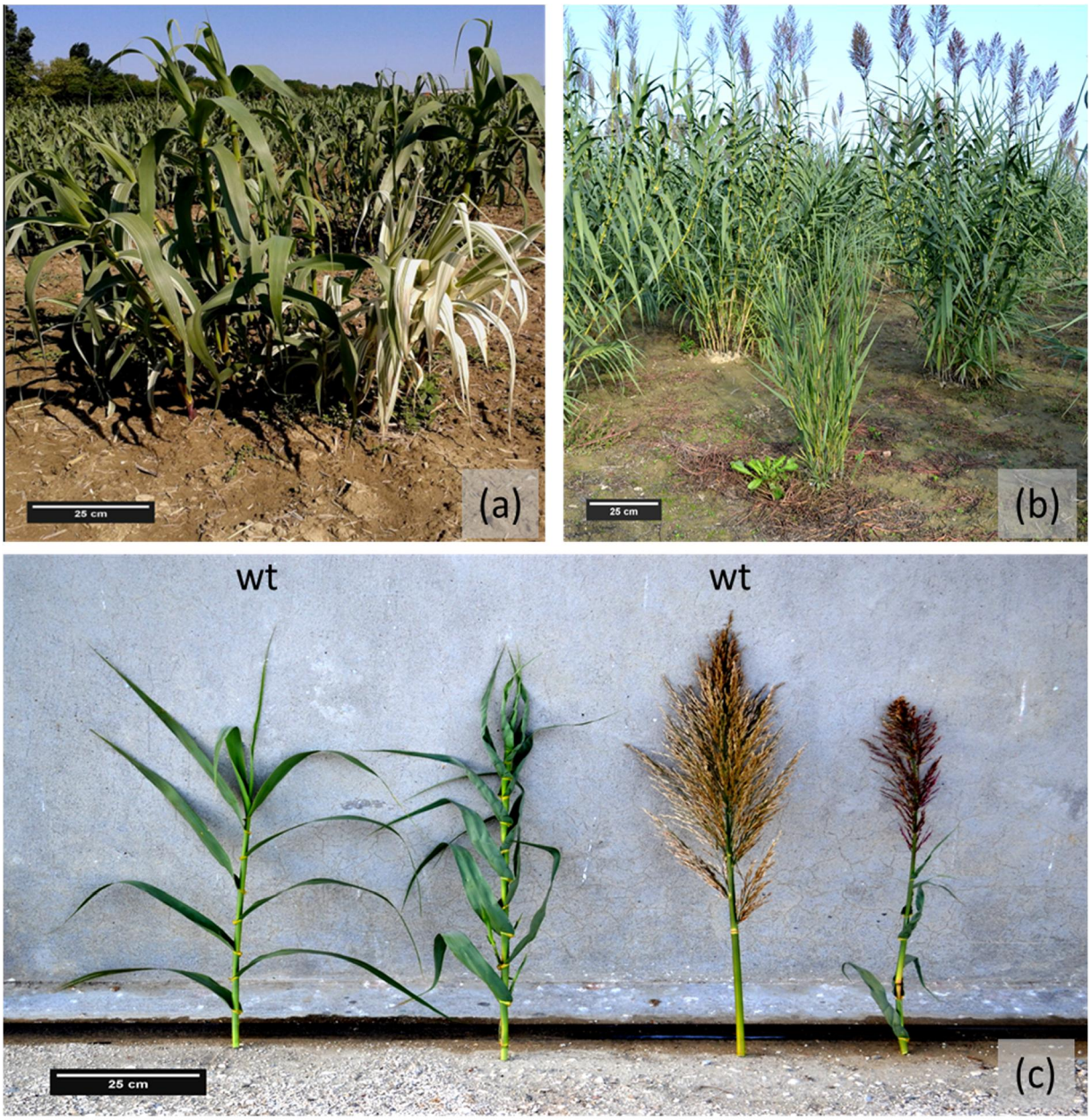


Figure 3. Examples of giant reed putative mutant clones. (A) *Variegate*, showing example of chimerism; (B) *Short stature* (in the foreground); (C) *Abnormal shoot* (wild type -wt- vegetative shoot and inflorescence are shown as comparison).

Figure 4. Representative images of phenotypic stability of giant reed mutant clones. Plants shown here were obtained by vegetative propagation (by rhizome subdivision) from putative mutant clones identified in the field based on visual observations. Photos were taken at approx. six months of cultivation in greenhouse after rhizome subdivision. Clones of untreated control plants are shown on the right side of each image. Additional details in text,Table 4 and Table S2.

1.5. DISCUSSION

A mutagenesis program requires to handle a high number of individuals. Giant reed is a relatively large plant with neither sexual reproduction nor seed production. Thus, while mutagenesis on pollen or seed was precluded, *in-vitro* cell cultures mutagenesis followed by regeneration appeared a suitable choice in order to produce a large number of mutagenized clones. This choice was also supported by the availability of reliable protocols for giant reed *in-vitro* culture and regeneration (Takahashi *et al.,* 2010) and by the former successful application of similar approaches in other species (reviewed in Jain *et al.,* 2010 and Suprasanna *et al.,* 2012) including physical (γ) irradiation in the botanically related species sugarcane (Nikam *et al.,* 2014). Based on regeneration rate per callus, irradiation doses in the 40-60 Gy range seemed appropriate in order to maximize the recovery of mutants in giant reed. Our RD_{50} estimate appears in the range of RD_{50}/LD_{50} estimates previously obtained in other polyploid species such as sugarcane (Nikam *et al.,* 2014), banana (Sales *et al.,* 2013), cassava (Magaia *et al.,* 2015) and rose (Bala &Singh, 2013). Higher irradiation doses (1.5x and 2x RD_{50}) were also applied in our experiment in order to maximize the recovery of visible mutant plants. Indeed, in polyploid species, the expression of phenotypes upon artificial mutagenesis is expected to be masked by high gene functional redundancy (Comai *et al.,* 2005). Therefore, higher doses should increase the probability of deleting two or more copies of the same genes and thus to increase the visible mutant recovery rate.

The frequency of clones (9.7%) showing aberrant phenotypes was in our experiment relatively high (for a highly polyploid species). Different mechanisms likely participated to this result. First, gene functional redundancy is probably less than that predicted based on genome duplications, given the high grade of genome re-arrangements, rapid sequence loss (of homeolog genes) and genome downsizing occurring after most polyploidization events (Parry

et al., 2009; Tayalè & Parisod, 2013; Soltis *et al.,* 2015). Indeed, genome downsizing mechanisms have likely acted in giant reed given its relatively low genome size (n ≈ 2.4 pg ≈ 2.3 Gb) as compared to its presumed high ploidy (9 or 18*x*). This is also exemplified by the relatively reduced number (six) of 5S, 45S loci and nucleolus, which are usually maintained as single locus in a genome (Hardion *et al.,* 2015). Second, the multiple genomes which originated extant giant reed were likely heterogeneous (Hardion *et al.,* 2015), therefore mutational events (e.g. deletions) where one dominant allele would turn to a recessive one at a heterozygous locus (and thus inducing a phenotype) are expected to be relatively frequent (Suprasanna *et al.,* 2012). Third, dominant mutations are also expected to occur, albeit with drastically reduced frequency (< 1% of recessive mutations), as previously empirically shown in other systems (Gottschalk and Wolff, 1983).

The molecular nature of the induced variants remains to be investigated. Some of the variants might have resulted from somaclonal variation since *in-vitro* cultures were involved in our protocol. Somaclonal variation is known to include several types of molecular events including change in chromosome number and structure, transposon activation and movement, point mutations and/or methylation changes (Kaeppler *et al.,* 2000; Bairu *et al.,* 2011; Ong-Abdullah *et al.,* 2015). Since γ-irradiation is thought to mostly induce deletions, a molecular analysis disclosing deletions across the mutant collection should indirectly enable us to quantify the impact of somaclonal variation in our materials. Applications of next-generation sequencing to characterize genome features of this collection are currently in progress.

Because multicellular tissues were irradiated, it is likely that some regenerated plants were chimeric (i.e. carrying sectorial genetic differences of cells and tissues in the same individual). However, protocols suitable to gradually separate chimeras and produce solid homohistons by several generations of vegetative propagation have been described and can be applied when needed (Predieri, 2001; Mba, 2013). Notably, the four “variegate” mutants that

were identified within our collection all turned out to be not solid (i.e. the variegation was clearly of different grade among different shoots of the same plant). This notwithstanding, "variegate" clones require further investigations as they might not simply be due to chimeras generated by genetic mutations (Marcotrigiano, 1997).

Interestingly, the average number of shoot per plant and shoot height correlated positively and negatively, respectively, with irradiation dose. More specifically, the higher the irradiation dose, the higher the number of buds which were initiated and/or developed in the rhizome of field-grown plants, which resulted in more shoots of smaller size. Previous investigations had already shown that increased shoot and root branching is one of the stress-induced morphogenic response in plants, which may involve the perception of reactive oxigen species and altered phytormone physiology (Potters *et al.,* 2007). Additionally, moderate stresses are well known to induce bud formation in perennial plant species (Grossnickle, 2012). However, whether the response observed in our case was directly triggered by the primary cellular injury caused by the irradiation treatment and perceived early by the calli/early regenerants, or was a consequence of a secondary type of stress due to the mutation load is currently unclear.

Although all clones showing aberrant phenotypes should in principle be considered putative mutants until further investigations will be carried out, three observations suggest that these clones may indeed be real mutants. First, all highlighted clones (Table 3) showed phenotypic stability based on field-plot observations during two successive years of cultivation, where shoots (ie. any above-ground plant structures) were completely mechanically shredded at the end of the first year. Second, five randomly chosen clones confirmed their abnormal phenotypes in vegetatively propagated sub-clones. Third, several of our giant reed putative mutants bear obvious resemblance with similar mutants described in other grass species. For instance, some giant reed short stature mutants showed similarities (including shorter internodes and erect leaves) to typical dominant dwarf maize mutants (Sheridan, 1988).

Similarly, high or low-tillering mutants are well known in maize, barley and rice (Neuffer *et al.,* 1997; Chuck *et al.,* 2007; Hussien *et al.,* 2014). Many additional clones showed more subtle, quantitative, differences in comparison with untreated control plots, for traits such as flowering time, shoot habitus (erect *vs*. prostrate), leaf size and others. Replicate experiments will be required to test the stability of these phenotypes.

Mutants for several agronomically and industrially important phenotypes could be searched in our collection. Ample variation for quantitative traits linked with yield/biomass, such as the number of shoots, shoot height and others was clearly observed and could be verified in replicated trials. Additionally, mutants could be searched for improved response to multiple cuts per year and the extension of leaf juvenility, which are growth-related traits recently shown to improve biomethane yield (Di Girolamo *et al.,* 2013; Ragaglini *et al.,* 2014), and for enhanced cold and drought tolerance since little variation exists between giant reed ecotypes for these traits (Pompeiano *et al.,* 2013; Haworth *et al.,* 2016). Extreme variants could be searched for (lower) ash content, which is an important quality parameter in several energy-generating processes and is known to vary between giant reed ecotypes (Amaducci and Perego, 2015), or for cell wall properties, including cellulose, hemicellulose and lignin, or their chemical modifications, which may strongly impact energy transformation or be important as independent bioproducts (Chen &Dixon, 2007; Marriott *et al.,* 2014; Wang *et al.,* 2016). Based on similar hypotheses, mutants characterized by improved saccharification properties were identified in a mutagenized population of *Brachypodium* (Marriott *et al.,* 2014). Finally, clones characterized by reduced stem fragmentation and rooting propensity at shoot nodes, and/or reduced rhizome diffusion could help to mitigate the supposed giant reed invasiveness, a perceived threat in some environments (Saltonstall *et al.,* 2010).

1.6. CONCLUSIONS

The paucity in giant reed genetic diversity, the lack of organized breeding efforts and the increasing interest in multiple industrial applications make the identification of novel giant reed genotypes and the release of genetically improved giant reed cultivars two important priorities in the field of plant feedstock research. In this work we showed that it is possible to generate remarkable genetic and phenotypic variation for agronomically relevant traits in this highly polyploidy species by means of *in-vitro* physical mutagenesis. A collection of one thousand mutants is now under further molecular and agronomical characterization. Thus, the protocol and the materials described here could represent the beginning of giant reed genetic improvement and could be of interest in breeding programs of other vegetatively propagated species suffering of low genetic variability.

1.7 ACKNOWLEDGEMENTS

The research reported here was conducted in WATBIO (Development of improved perennial non-food biomass and bioproduct crops for water-stressed environments) which is a collaborative research project funded from the European Union's Seventh Programme for research, technological development and demonstration under grant agreement No. 311929. We thank Stefano Vecchi, Pierluigi Meriggi, Giovanni Miccoli, Simona Corneti, Riccardo Bovina, Riccardo Di Pasquale, Danara Ormanbekova, Martina Costacurta, Lucia Pernigotto, Stefano Predieri, Mila D'Angelantonio and Giuseppe Di Girolamo for technical assistance.

1.8 REFERENCES

Amaducci S, Perego A (2015) Field evaluation of *Arundo donax* clones for bioenergy production. *Industrial Crops and Products*, **75**, 122-128.

Ahmad R, Liow P, Spencer D, Jasieniuk M, (2008) Molecular evidence for a single genetic clone of invasive Arundo donax in the United States. *Aquatic Botany,* **88**, 113-120

Bado S, Forster BP, NielenS, Ali AM, Lagoda PJ, Till BJ, Laimer M (2015) Plant mutation breeding: current progress and future assessment. *Plant Breeding Reviews*, **39**, 23-87.

Bairu MW, Aremu AO, Van Staden J (2011) Somaclonal variation in plants: causes and detection methods. *Plant Growth Regulation*, **63**, 147-173.

Bala M, Pal Singh K (2013) In vitro mutagenesis of rose (*Rosa hybrida* L.) explants using gamma-radiation to induce novel flower colour mutations. *The Journal of Horticultural Science and Biotechnology*, **88**, 462-468.

Barrero RA, Guerrero FD, Moolhuijzen P, Goolsby JA, Tidwell J, Bellgard SE, Bellgard MI (2015) Shoot transcriptome of the giant reed, *Arundo donax*. *Data in Brief*, **3**, 1-6.

Boland JM (2006) The importance of layering in the rapid spread of *Arundo donax* (giant reed). *Madrono*, **53**, 303–312.

Bor NL (1970) Gramineae. In: Rechinger KH, ed. *Flora Iranica*, **70**, 147–244.

Brach AR, Song H (2006) eFloras: new directions for on line floras exemplified by the Flora of China Project. *Taxon*, **55**, 188–192.

Bradshaw JE (2016) Mutation Breeding. In: *Plant Breeding: Past, Present and Future*, PP. 529-560.Springer, Berlin.

Bucci A, Cassani E, Landoni M, Cantaluppi E, Pilu R (2013) Analysis of chromosome number and speculations on the origin of *Arundo donax* L. (Giant Reed). *Cytology and Genetics*, **47**, 237-241.

Cantaluppi E, Cassani E, Puglisi D, Corno L, Munaro M, Landoni M, Pilu R (2015) Study on the inflorescences of *Arundo donax* L. clones sampled in Italy. *Brazilian Journal of Botany*, **39**, 275-285.

Ceotto E, Di Candilo M (2010) Shoot cuttings propagation of giant reed (*Arundo donax* L.) in water and moist soil: The path forward?. *Biomass and Bioenergy*, **34**, 1614-1623.

Chen F, Dixon RA (2007) Lignin modification improves fermentable sugar yields for biofuel production. *Nature Biotechnology*, **25**, 759-761.

Chuck G, Cigan AM, Saeteurn K, Hake S (2007) The heterochronic maize mutant *Corngrass1* results from overexpression of a tandem microRNA. *Nature Genetics*, **39**, 544-549.

Comai L (2005) The advantages and disadvantages of being polyploidy. *Nature Reviews Genetics*, **6**, 836-846.

Corno L, Pilu R, Adani F (2014) *Arundo donax* L.: a non-food crop for bioenergy and bio-compound production. *Biotechnology advances*, **32**, 1535-1549.

Cosentino SL, Copani V, D'agosta GM, Sanzone E, Mantineo M (2006) First results on evaluation of *Arundo donax* L. clones collected in Southern Italy. *Industrial Crops and Products*, **23**, 212–222.

Di Girolamo G, Grigatti M, Barbanti L, Angelidaki I (2013) Effects of hydrothermal pre-treatments on Giant reed (*Arundodonax*) methane yield. *Bioresource Technology*, **147**:152–159.

Douglas GC (1986) Effects of gamma radiation on morphogenesis and mutagenesis in cultured stem explants of poplar. Nuclear techniques and in vitro culture for plant improvement. *Procedure International Symposium*,IAEA, Vienna, PP. 121-128.

Ge X, Xu F, Vasco-Correa J, Li Y (2016) Giant reed: A competitive energy crop in comparison with miscanthus. *Renewable and Sustainable Energy Reviews*, **54**, 350-362.

Gottschalk W, Wolff G (1983) Induced Mutations in Plant Breeding. *Springer Science & Business Media*, LLC, USA.

Grossnickle SC (2012) Why seedlings survive: influence of plant attributes. *New Forests*, **43**, 711 - 738

Hardion L, Verlaque R, Baumel A, Juin M, Vila B (2012) Revised systematics of Mediterranean *Arundo* (Poaceae) based on AFLP fingerprints and morphology *Taxon*, **61**, 1217–1226.

Hardion L, Verlaque R, Saltonstall K, Leriche A, Vila B (2014) Origin of the invasive *Arundo donax* (Poaceae): a trans-Asian expedition in herbaria. *Annals of Botany*, **114**, 455–462.

Hardion L, Verlaque R, Rosato M, Rosselló JA, Vila B (2015)Impact of polyploidy on fertility variation of Mediterranean *Arundo* L. (Poaceae). *Comptesrendus biologies*, **338**, 298-306.

Haworth M, Centritto M, Giovannelli A, Marino G, Proietti N, Capitani D, Loreto F (2016) Xylem morphology determines the drought response of two *Arundo donax* ecotypes from contrasting habitats. *Global Change Biology*.DOI: 10.1111/gcbb.12322

Hussien A, Tavakol E, Horner DS, Muñoz-Amatriaín M, Muehlbauer GJ, Rossini L (2014) Genetics of tillering in rice and barley. *The plant genome*, **7**, 1.

Jain SM, Ochatt SJ, Kulkarni VM, Predieri S (2010) In vitro culture for mutant development. *ActaHorticultural*, **865**, 59–68.

Kaeppler SM, Kaeppler HF, Rhee Y (2000) Epigenetic aspects of somaclonal variation in plants. *Plant molecular biology*, **43**, 179-188.

Khudamrongsawat J, Tayyar R, Holt JS (2004) Genetic Diversity of Giant Reed (*Arundo donax*) inthe Santa Ana River, California. *Weed Science*, **52**, 395-405.

Kodym A, Afza R, Forster B, Ukai Y, Nakagawa H, Mba C (2012) Methodology for physical and chemical mutagenic treatments. In: Mutation Techniques in Plants - Principles and Applications. Q.Shu (FAO/IAEA), H Nakagawa (Japan), B Forster (UK) (Eds). CABI Publishing. pp 169-181.

Lee GJ, Zhang L, Choi YI, Chung SJ, Yoo YK, Kim DS, Kim SH (2012) Biomass and molecular characteristics of multi-tillering Miscanthus mutants. *Korean Journal of Plant Resources*, **25**, 745-752.

Magaia HE Joseph J, Elsy CR, Francis RM, Bastian D (2015) Creation of variability by in vitro mutagenesis in cassava (*Manihot esculenta*, Crantz). *Journal of Tropical Agriculture*, **53**, 123-130.

Marcotrigiano M (1997) Chimeras and variegation: patterns of deceit. *Horticultural Science*, **32**, 773-784.

Marriott PE, Sibout R, Catherine Lapierre C *et al* (2014) Range of cell-wall alterations enhance saccharification in *Brachypodium distachyon* mutants. *PNAS*, **111**, 14601-14606

Mariani C, Cabrini R, Danin A, Piffanelli P, Fricano A, Gomarasca S, Soave C (2010) Origin, diffusion and reproduction of the giant reed (*Arundo donax* L.): a promising weedy energy crop. *Annals of Applied Biology*, **157**, 191–202.

Mba C (2013) Induced mutations unleash the potentials of plant genetic resources for food and agriculture. *Agronomy*, **3**, 200-231.

Monti A, Zegada-Lizarazu W (2016) Sixteen-year biomass yield and soil carbon storage of giant reed (*Arundo donax* L.) grown under variable nitrogen fertilization rates. *BioEnergy Research*, **9**, 248-256.

Neuffer MG, Coe EH, Wessler SR (1997) Mutants of maize. *Cold Spring Harbor Laboratory Press*.

Nikam, AA, Devarumath RM, Shitole MG, Ghole VS, Tawar PN, Suprasanna P (2014) Gamma radiation, in vitro selection for salt (NaCl) tolerance, and characterization of mutants in sugarcane (*Saccharum officinarum* L.). *In Vitro Cellular & Developmental Biology-Plant*, **50**, 766-776.

Nsanganwimana F, Marchand L, Douay F, Mench M (2013) *Arundo donax* L., a candidate for phyto managing water and soils contaminated by trace elements and producing plant-based feedstock. A review. *International Journal of Phytoremediation*, **16**, 982-1017.

Ong-Abdullah M, Ordway JM, Jiang N *et al.* (2015)Loss of *Karma* transposon methylation underlies the mantled somaclonal variant of oil palm. *Nature*, **525**, 533-537.

Parry MAJ, Madgwick PJ, Bayon C *et al.* (2009) Mutation discovery for crop improvement. *Journal of Experimental Botany*, **60**, 2817-2825.

Perdue RE (1958) *Arundo donax*: source of musical reeds and industrial cellulose. *Economic Botany*, **12**, 368–404.

Perera D, Barnes DJ, Baldwin BS, Reichert NA (2015) Mutagenesis of in vitro cultures of *Miscanthus* × *giganteus* cultivar Freedom and detecting polymorphisms of regenerated plants using ISSR markers. *Industrial Crops and Products*, **65**, 110-116.

Pilu R, Cassani E, Landoni M, Badone FC, Passera A, Cantaluppi E, Corno L, Adani F (2014)Genetic characterization of an Italian Giant Reed (*Arundo donax* L.) clones collection: exploiting clonal selection. *Euphytica*, **196**, 169-181.

Pompeiano A, Vita F, Miele S, Guglielmetti L (2013) Freeze tolerance and physiological changes during cold acclimation of giant reed [*Arundo donax* (L.)]. *Grass Forage Science*, **70**, 168–175.

Potters G, Pasternak TP, Guisez Y, Palme KJ, Jansen MAK (2007) Stress-induced morphogenic responses: growing out of trouble? *Trends in Plant Science*, **12**, 98–105.

Predieri S (2001) Mutation induction and tissue culture in improving fruits. *Plant Cell, Tissue and Organ Culture*, **64**, 185–210.

Ragaglini G, Dragoni F, Simone M, Bonari E (2014) Suitability of giant reed (*Arundo donax* L.) for anaerobic digestion: effect of harvest time and frequency on the biomethane yield potential. *Bioresource Technology*, **152**, 107–115.

Sablok G, Fu Y, Bobbio V, Laura M, Rotino GL, Bagnaresi P, Li M (2014) Fuelling genetic and metabolic exploration of C3 bioenergy crops through the first reference transcriptome of *Arundo donax* L. *Plant biotechnology journal*, **12**, 554-567.

Sales EK, Lopez J, Espino RRC, Butardo NG, Gonzalez Diaz L (2013) Improvement of Bananas through Gamma Ray Irradiation. *Philippine Journal of Crop Science* (PJCS), **38**: 47-53.

Saltonstall K, Lambert A, Meyerson LA (2010) Genetics and reproduction of common (*Phragmites australis*) and giant reed (*Arundo donax*). *Invasive Plant Science and Management*, **3**: 495-505.

Sanchez E, Scordia D, Lino G, Arias C, Cosentino S, Nogues S (2015) Salinity and water stress effects on biomass production in different *Arundo donax* L. clones. *Bioenergy Research*, **8**, 1461–1479.

Sheridan WF (1988) Maize developmental genetics: genes of morphogenesis. *Annual Review of Genetics*, **22**, 353-385.

Soltis PS, Marchant DB, Van de Peer Y, Soltis DE (2015) Polyploidy and genome evolution in plants. *Current opinion in genetics & development*, **35**, 119-125.

Suprasanna P, Jain SM, Ochatt SJ, Kulkarni VM, Predieri S (2012) Applications of in vitro techniques in mutation breeding of vegetatively propagated crops. In: Shu QY, Forster BP, Nakagawa H (eds) *Plant mutation breeding and biotechnology*. CAB International, Wallingford, PP. 369–383.

Takahashi W, Takamizo T, Kobayashi M, Ebina M (2010) Plant regeneration from calli in giant reed (*Arundo donax* L.). *Grassland science*, **56**, 224-229.

Tayalé A, Parisod C (2013) Natural pathways to polyploidy in plants and consequences for genome reorganization. *Cytogenetic and Genome Research*, **140**, 79-96.

Wang Y, Fan C, Hu H, Li Y, Sun D, Wang Y, Peng L (2016) Genetic modification of plant cell walls to enhance biomass yield and biofuel production in bioenergy crops. *Biotechnology advances*,**34**, 997-1017.

Webster RJ, Driever SM, Kromdijk J *et al* (2016) High C3 photosynthetic capacity and highintrinsic water use efficiency underlies the high productivity of the bioenergy grass *Arundo donax*. *Scientific Reports*, **6**:20694.

Zegada-Lizarazu W, Elbersen W, Cosentino SL, Zatta A, Alexopoulou E, Monti A (2010) Agronomic aspects of future energy crops in Europe. *Biofuel Bioproducts and Biorefining*, **4**, 674–691.

CHAPTER 2

Chemical analysis of *Arundo donax* mutants

2.1. INTRODUCTION

The steady recent increase of energy demand and the UE2009/28/CE directive (each member state must produce 20% of its energy demand from renewable sources within 2020) are driving the European agriculture system towards increasing use of bioenergy crops. This will unavoidably lead to competition between food and bioenergy crops for existing land and other agricultural resources and production factors. Besides the use of wastes such as sludge from organic industrial wastes and organic domestic wastes, one way to overcome this drawback is to use ligno-cellulosic biomasses obtained from agricultural (such as stovers of cereals), or industrial residues, or from dedicated plant biomass crops grown in marginal areas and thus not directly competing with food crops (Di Girolamo, 2014). In general, the characteristics of such ideal bioenergy crop are (McKendry, 2002):

- high yield (in terms for biomass per hectare per year);
- low energy input to produce;
- low production costs;
- low presence of contamitants and low level of production of pollutants;
- it can be grown under low fertility conditions (including poor or contaminated soils) and it requests little or no input in terms of water, nutrients, etc..

The cultivation of dedicated bioenergy crops could bring several advantages, including to sustain local agrosystems, to provide alternative source of income in rural zones, to promote the use of marginal lands and to reduce the level of atmospheric CO2. (Zegada- Lizarazu *et al*.,

2010). Conversion of biomass to energy can be obtained in several ways. Following the classification of McKendry 2002, these processes are:

1. thermo-chemical conversion, which includes:
    a. combustion,
    b. pyrolysis,
    c. gasification,
    d. liquefaction;
2. biochemical/biological, which includes:
    a. anaerobic digestion (production of biogas, a mixture of mainly methane and carbon dioxide)
    b. fermentation (production of ethanol).
3. mechanical extractions.

Among the methods above, anaerobic digestion is one of the most promising thanks to its several advantages both from economical and environmental standpoints. Specifically, the use of anaerobic digestion for the production of biogas and biomethane over natural (= fossil) gas offers crucial benefits: i) biomethane is a renewable resource; ii) the net input of greenhouse in the atmosphere is null; iii) it does not require to depend from outsource (other country) as it can be produced locally; iv) it reduces the production of pollutants and wastes (McKendry, 2002b). However, anaerobic digestion of lingo-cellulosic substrates can be seriously constrained by chemical and structural characteristics of the substrates themselves, including level of cellulose crystallinity, polymeration grade, surface for enzymatic attack, and lignin content. Lignin is the most recalcitrant component and can seriously hinder anaerobic digestion (Taherzadeh e Karimi, 2008). A main negative effect comes from the action of protection of lignin to cellulose and hemicellulose, which reduces the efficiency of enzymatic degradation (Frigon e Guiot, 2010).

First generation energy crops include maize, sorghum triticale and others, but their use for energy production is not supported by current national and European policies. Main second generation lignocellulosic energy crops species (non food crop species not directly competing with food and feed crops) can be broadly divided in four categories (Zegada-Lizarazu et al. 2010): herbaceous annual, perennial herbaceous, short rotation woody crops and oilseed crops. Herbaceous annual crop species include hemp (Cannabis sativa L.) and sweet sorghum (Sorghum bicolor). Main perennial herbaceous crops include switchgrass (Panicum virgatum), miscanthus (Miscanthus spp.) and giant reed (Arundo donax). This crops are grass crops that are being developed

for biomass production in Europe and North America. They are characterized by very low production costs, can be grown on marginal lands, they usually require relative low water amounts and low nutrient and agrochemical inputs, and provide positive environmental benefits such as carbon storage in a well-developed root system (Zegada-Lizarazu et al. 2010).

2.2. OBJECTIVE

The Plant Genetics group of the Department of Agricoltural Sciences, University of Bologna, has produced a mutagenized population of one thousand A. donax clones using g-irradiation. Initial characterization of this population showed a large number of clones with solid and stable phenotypic abnormalities in the shoot part of the plant, which were attributed to genetic mutations. Abnormalities included strongly reduced plant vigor, altered stem dimension, leaf color and shape, number of tillers/shoots per plot, change in early vigor and flowering time, etc. From the energy production standpoint, it would be of great interest to identify mutants with altered chemical compositions, providing improved propensity to biomass transformation to biofuel and biomethane. For instance, it is well recognized that lower lignin vs cellulose rate,

higher concentration of sugar and other non-structural carbohydrates, lower ash content are all positive attributes which enhance chemical or biochemical transformation of Arundo biomass in biomethane or ethanol.

This work aims to obtain a preliminary screening of the induced genetic variation present in the one thousand A. donax mutagenized clone collection in order to identify clones with improved biofuel and/or biomethane yield per unit of biomass. To reach this objectives, we have chemically analysed 100 clones of A. donax for several key chemical components of plant biomass including lignin, cellulose, hemicelluloses, non-structural carbohydrates, ash, and others.

2.3. MATERIALS AND METHODS

2.3.1. Source of plant materials and chemical analysis

Details on the origin of *A. donax* mutagenized clones are provided in Chapter 1 of this thesis. Briefly, approximately 1,000 *A. donax* mutagenized clones were transplanted in the field during spring (April) 2015, at locality Ca' Bosco, near Ravenna, in the eastern Po Valley, north Italy. Order of transplanting was random (and independent from irradiation dosage) and included 11 wild-type (control, untreated) plants. Distance between plants was 2.5 m between rows and 2.5 m between two plants on the same row. After transplantion plants were treated following standard agronomic practice. At the end of the first and second growing seasons (November 2015 and November 2016) plants (more precisely, the above-ground part = shoots of each plant) were completely mechanically shredded.

Samples for chemical analysis described in this chapter were collected in October 2015 from 98 mutagenized clones and two wild type (control untreated) plants. Three samples per plant (clone) were collected as follows: basal (sample including stem and leaves from the 2nd, 3rd and

4th node, counting from the soil level); central (sample including stem and leaves from mid stem node, and the two adjacent nodes); apical (sample including stem and leaves from the 2nd, 3rd and 4th node counting from the top of the plant). Chemical analysis which were carried out were:

- dry residue (dr) (% fresh weight);
- protein (% dr);
- total fiber (% dr) by Wendee method;
- lipid (% dr);
- ash (% dr);
- NDF (Neutral Detergent Fiber, % dr, includes lignin, cellulose and hemicellulose), by Van Soest method;
- ADF (Acid Detergent Fiber, % dr, includes lignin and cellulose), by Van Soest method;
- ADL = lignin (Acid Detergent Lignin, % dr), by Van Soest;
- hemicellulose (% dr; obtained as NDF − ADF);
- cellulose (% dr; ADF − ADL);
- pH
- NSC (Non-structural carbohydrate, % dr), computed as NSC = 100 − (% NDF + % protein + % lipid + % ash).

Neutral Detergent Fiber (NDF): it is the most common measure of fiber used for animal feed analysis, it measures most of the structural components in plant cells (i.e. lignin, hemicellulose and cellulose), but not pectin. Further analysis can be done to the sample to determine individual components such as ADF analysis (source: Wikipedia).

Acid Detergent Fiber (ADF): The fibrous component represents the least digestible fiber portion of forages and includes lignin, cellulose, silica and insoluble forms of nitrogen but not hemicellulose. Forages with higher ADF are lower in digestible energy than forages with lower

ADF. During laboratory analysis, ADF is the residue remaining after boiling a forage sample in acid detergent solution (Source: http://georgiaforages.caes.uga.edu/glossary/A.html).

All chemical analysis were carried out by Lab Analisi Zootecniche Sas, Gonzaga MN (Italy) using a NIR based approach on a FOSS 5000 (FOSS Analytical, Denmark, www.foss.dk), where not differently indicated.

2.3.2. Statistical analysis

Correlation analysis was carried out by applying the Pearson's *r* coefficient of correlation. Correlation between traits was also assessed by means of Correspondence Analysis. All statistical analyses were carried out using Past 3.0 (Hammer et al. 2001).

2.4. RESULTS AND DISCUSSION

2.4.1. Variation for chemical traits within the collection of mutagenized *A. donax* clones.

Chemical traits collected on shoot samples of 98 mutagenized clones and two wild type (wt) untreated control plants showed a sizeable range of variation within reliable and expected values, with only one exception (clone #154-40.218. For this clone, strongly divergent values were obtained for several chemical parameters, therefore we preferred to exclude it from most of statistical synthesis - eg. computation of mean values, while we are carrying out all the due checks). The values from the three tissue samples (top, mid and base of stems) per plant were first analysed as mean values (Table 1 and Table 2). Coefficient of variation (CV) ranged from 2.77% for pH to 21.02% for NCS, which therefore resulted the least and the most variable traits among those collected. For NCS, the lower value was 8.00 (% of dry residue, dr) and the highest was 18.18 % dr (more than twice higher). Samples from wt untreated clones showed intermediate values (11.57 - 13.71). For ADL (lignin), variation ranged from 6.60-11.01% dr,

with a mean value of 9.36% dr. Again, samples from wt clones showed intermediate values (9.30 - 9.90 % dr).

Variation of chemical traits when the position of the samples within the plants (top, mid or basal part of the stem) is considered is illustrated in Table 3. Generally, and as expected, samples taken from the basal portion of the stems showed higher fiber content, in terms of total Fiber, NDF, ADF, ADL and cellulose, than the samples taken from the mid or apical portions. This is likely the consequence of the older and much more structured presence of mature cell walls around fully extended (ie. larger) cells. On the contrary, apical samples showed the higher values for Hemicellulose, Protein, Lipids and Ash, likely as a consequence of the higher number of cells per unit of tissue samples, and of the thinner and younger cell-wall structure of the same cells. Samples with higher proportion of NSC (12.76% dr) were from the mid part of the stems.

Table 1. Range and coefficient of variation for chemical traits collected on 100 *A. donax* clones. wt1 and wt2: values for the two wild-type untreated controls. cv: coefficient of variation. sd: standard deviation. Values were computed by excluding clone #154-40.218 (see text).

Clone id	DR	Protein	Fiber	Lipid	Ash	NDF	ADF	ADL	Hemic.	Cellul.	pH	NSC
min	29.78	7.60	34.58	0.70	4.35	62.35	38.62	6.60	21.17	31.29	4.67	8.00
mean	35.86	9.88	39.74	1.40	6.34	70.61	46.26	9.36	24.35	36.90	4.95	11.77
max	43.14	13.00	44.32	2.04	8.79	77.06	51.12	11.01	26.89	40.51	5.28	18.18
cv	8.53	10.49	5.51	15.77	12.71	4.96	6.22	8.79	4.93	5.85	2.77	21.02
sd	3.06	1.04	2.19	0.22	0.81	3.50	2.88	0.82	1.20	2.16	0.14	2.47
wt1	*40.46*	*10.18*	*38.50*	*1.48*	*6.44*	*68.19*	*44.46*	*9.30*	*23.73*	*35.16*	*5.13*	*13.71*
wt2	*35.06*	*9.77*	*40.89*	*1.06*	*6.61*	*70.98*	*46.25*	*9.99*	*24.74*	*36.25*	*5.00*	*11.57*

Table 2. Mean values of chemical traits collected on 100 mutant clones of *A. donax.*

Clone id		DR	Protein	Fiber	Lipid	Ash	NDF	ADF	ADL	Hemic.	Cellulose	pH	NSC
157-40.24	1	30.62	9.11	40.58	1.26	6.43	72.56	48.57	9.81	24.00	38.76	4.73	10.63
27-40.1	2	36.50	9.73	40.13	1.38	5.90	72.21	47.12	9.75	25.09	37.37	5.02	10.79
15-25.1	3	36.52	9.74	41.28	1.20	6.34	72.93	47.54	10.47	25.39	37.07	5.04	9.80
178-60.14	4	31.30	9.55	40.67	1.29	6.74	71.64	46.83	9.40	24.81	37.43	4.88	10.78
10-60.1	5	33.44	9.76	40.10	1.43	5.74	72.59	46.47	9.41	26.12	37.06	4.90	10.47
182-60.94	6	34.90	9.77	38.04	1.87	6.12	67.64	44.49	8.49	23.15	36.01	4.91	14.60
130-40.4	7	32.23	9.89	40.90	1.32	6.99	73.33	46.66	9.32	26.67	37.35	4.82	8.47
215-60.1	8	32.46	8.61	40.59	1.33	5.88	73.73	49.19	9.87	24.54	39.32	4.74	10.46
157-40.26	9	36.15	9.70	41.59	1.33	6.58	73.57	48.56	10.39	25.01	38.17	4.90	8.82
182-60.85	10	32.86	9.97	39.10	1.65	7.41	71.64	46.65	9.41	24.99	37.23	4.88	9.33
182-60.60	11	30.73	9.21	40.29	1.42	8.79	71.46	47.93	9.68	23.53	38.25	4.81	9.12
5-60.7	12	32.75	8.48	41.38	1.39	6.96	74.88	49.66	10.14	25.22	39.52	4.82	8.29
182-60.99	13	30.97	9.11	41.29	1.30	7.16	73.91	49.02	9.60	24.88	39.42	4.75	8.52
35-94.13	14	39.03	7.60	43.80	1.11	5.12	76.47	51.12	10.61	25.35	40.51	4.81	9.70
154-40.14	15	38.84	8.49	42.69	1.32	6.26	75.39	49.29	9.71	26.10	39.58	5.00	8.53
154-40.54	16	40.27	9.27	41.51	1.32	5.45	74.70	48.31	10.07	26.38	38.24	5.02	9.26
154-40.29	17	32.35	8.63	42.73	1.10	6.90	74.95	49.99	10.43	24.96	39.56	4.78	8.42
182-60.95	18	37.44	9.76	41.99	1.28	7.34	73.56	48.50	10.10	25.06	38.39	4.99	8.06
142-40.3	19	36.84	8.63	42.36	1.35	6.24	75.79	50.82	10.39	24.96	40.43	4.82	8.00
12-80.4	20	34.34	9.52	41.71	1.30	5.91	74.36	47.62	9.54	26.75	38.07	4.98	8.90
182-60.37	21	33.12	10.03	40.66	1.33	7.03	72.00	47.51	9.82	24.49	37.68	4.87	9.62
154-40.38	22	39.41	8.62	43.23	1.24	4.46	76.89	50.00	10.39	26.89	39.61	4.89	8.80
154-40.56	23	36.88	9.54	41.69	1.34	6.24	72.83	47.78	9.80	25.04	37.98	4.91	10.05
182-60.17	24	34.05	9.98	39.09	1.52	6.21	72.38	45.74	9.49	26.63	36.26	4.98	9.92
5-60.11	25	38.72	9.35	39.88	1.42	5.88	69.98	45.73	8.77	24.25	36.97	5.00	13.37
154-40.37	26	36.55	9.37	40.31	1.48	5.97	73.35	48.63	10.01	24.71	38.62	4.88	9.82
301-100.2	27	41.08	9.42	41.91	1.41	4.94	74.61	48.87	10.18	25.75	38.69	5.01	9.62
31-60.3	28	36.30	10.45	40.26	1.42	6.97	71.83	46.75	9.76	25.08	36.98	5.02	9.34
182-60.53	29	30.96	10.41	39.51	1.47	7.27	71.84	46.01	9.51	25.83	36.49	4.94	9.01
182-60.72	30	32.20	11.21	38.17	1.72	6.83	70.22	44.82	8.59	25.39	36.23	4.90	10.02
5-60.9	31	31.40	9.80	39.81	1.35	7.09	70.75	47.63	9.74	23.12	37.89	4.77	11.01
182-60.50	32	37.56	8.64	41.99	1.35	6.14	74.88	50.08	10.57	24.80	39.51	4.85	8.99
154-40.1	33	36.87	9.23	41.56	1.36	7.21	72.82	48.61	10.32	24.21	38.29	4.96	9.38
40-20.1	34	37.74	9.54	40.50	1.40	6.43	71.97	47.89	9.97	24.08	37.93	4.98	10.66
178-60.8	35	37.71	9.60	40.82	1.45	5.84	73.92	48.64	9.57	25.29	39.07	4.92	9.19
182-60.10	36	39.15	9.13	41.33	1.35	5.83	73.44	48.38	10.27	25.06	38.11	4.98	10.25
8-100.1	37	36.30	9.33	42.28	1.21	6.86	73.12	48.82	9.83	24.30	38.99	4.92	9.48
39-40.9	38	37.53	10.27	40.86	1.25	7.42	69.88	46.65	9.53	23.22	37.12	5.06	11.18
154-40.15	39	34.19	8.58	41.16	1.30	6.35	74.33	49.40	9.91	24.94	39.49	4.87	9.44
146-40.1	40	32.70	10.08	39.99	1.64	8.30	69.81	47.23	9.89	22.58	37.34	4.84	10.17
154-40.20	41	37.83	8.37	43.24	0.99	4.39	74.57	49.65	10.38	24.92	39.27	4.83	11.68
154-40.68	42	38.66	9.32	41.39	1.16	5.37	73.31	48.03	10.58	25.28	37.45	4.97	10.84
182-60.38	43	33.57	9.79	39.75	1.38	7.07	70.58	47.05	9.53	23.53	37.52	4.86	11.18
154-40.11	44	38.98	9.53	41.18	1.24	5.93	71.41	47.70	9.78	23.71	37.92	4.97	11.89
182-60.90	45	33.15	9.62	39.12	1.47	6.17	71.19	47.17	9.23	24.02	37.93	4.86	11.56
157-40.13	46	33.38	9.57	38.95	1.53	5.91	69.69	45.84	8.97	23.85	36.87	4.91	13.30
215-60.2	47	32.22	8.87	41.36	1.23	5.75	73.47	49.08	9.52	24.40	39.55	4.74	10.68
182-60.82	48	29.78	7.90	40.46	1.43	5.72	72.78	49.32	9.31	23.46	40.01	4.67	12.19
182-60.31	49	31.65	8.65	41.63	1.20	5.90	73.15	48.99	10.04	24.16	38.96	4.73	11.10
235-80.161	50	40.04	8.30	44.21	1.06	5.15	76.84	51.09	11.01	25.74	40.08	4.84	8.66
235-80.99	51	31.55	8.94	40.13	1.38	6.57	71.27	48.44	9.88	22.82	38.57	4.70	11.85
228-60.29	52	33.29	9.64	38.82	1.62	6.54	70.05	47.35	9.14	22.70	38.21	4.78	12.16
235-80.59	53	31.90	10.13	38.84	1.31	5.99	69.73	46.15	9.04	23.58	37.11	4.78	12.85
235-80.31	54	34.82	8.63	41.30	1.24	4.90	74.46	49.08	10.09	25.39	38.98	4.79	10.78
235-80.27	55	31.16	9.05	41.07	1.28	6.94	72.34	48.17	9.58	24.17	38.59	4.79	10.39
157-40.85	56	39.98	10.29	40.48	1.29	6.00	71.43	46.48	9.51	24.95	36.97	5.10	10.99
154-40.209	57	37.46	9.38	40.84	1.35	5.94	72.63	47.86	9.69	24.77	38.17	4.94	10.70
228-60.95	58	33.06	10.59	38.48	1.68	7.05	70.13	45.64	8.99	24.48	36.66	4.93	10.56
Clone id		DR	Protein	Fiber	Lipid	Ash	NDF	ADF	ADL	Hemic.	Cellulose	pH	NSC
228-60.48	59	33.21	10.19	37.86	1.76	6.97	70.07	45.93	8.74	24.14	37.19	4.88	11.01
228-60.44	60	35.27	10.18	39.55	1.43	6.35	71.30	46.03	8.93	25.27	37.10	5.02	10.75
157-40.92	61	37.48	8.78	44.32	0.95	4.97	77.06	50.39	10.38	26.67	40.01	4.91	8.24
154-40.97	62	40.04	9.61	42.68	1.20	4.35	73.60	48.47	10.13	25.13	38.34	4.91	11.24
182-60.41	63	35.27	9.12	41.57	0.70	5.91	71.47	47.68	10.03	23.80	37.65	4.85	12.80

157-40.117	64	35.46	10.69	38.71	1.34	6.16	67.66	44.00	8.75	23.66	35.25	4.92	14.16
182-60.79	65	35.83	11.14	38.60	1.30	6.58	65.46	42.23	8.07	23.23	34.16	5.12	15.51
128-60.33	66	40.17	10.48	37.30	1.42	6.74	64.90	42.46	8.21	22.45	34.24	5.25	16.46
WT (untreated)	67	35.06	9.77	40.89	1.06	6.61	70.98	46.25	9.99	24.74	36.25	5.00	11.57
5-60.4	68	31.75	11.20	38.59	1.37	8.10	65.96	43.28	8.20	22.68	35.08	4.94	13.38
178-60.18	69	32.74	12.16	38.33	1.27	6.80	65.37	41.82	7.76	23.55	34.06	4.92	14.40
22-80.2	70	36.47	10.85	39.18	1.14	5.96	64.46	43.16	8.71	21.31	34.44	4.97	17.59
235-80.10	71	35.74	10.18	39.96	1.20	5.94	68.77	45.03	9.11	23.75	35.91	4.95	13.92
127-40.3	72	36.76	10.31	38.95	1.26	5.24	67.61	44.58	8.91	23.03	35.67	4.92	15.59
228-60.28	73	33.99	11.43	37.15	1.42	6.96	65.62	41.57	7.28	24.04	34.30	5.00	14.58
235-80.29	74	34.87	10.84	36.17	2.04	6.60	66.59	42.74	7.75	23.85	34.98	5.04	13.93
235-80.95	75	33.75	10.87	38.52	1.43	7.13	66.98	44.81	9.08	22.16	35.73	4.91	13.59
235-80.54	76	37.84	10.39	38.50	1.62	6.17	69.20	45.07	9.04	24.13	36.03	4.95	12.62
235-80.153	77	39.17	12.31	35.32	1.62	6.91	64.77	39.62	8.33	25.15	31.29	5.24	14.39
157-40.56	78	37.06	9.26	41.68	1.13	6.49	72.63	47.52	10.02	25.11	37.50	5.04	10.48
228-60.43	79	34.88	12.83	34.96	1.67	7.47	62.35	38.62	6.60	23.73	32.02	5.09	15.68
235-80.150	80	37.01	12.24	35.38	1.66	6.51	64.62	40.12	8.29	24.50	31.83	5.14	14.97
228-60.92	81	35.69	10.20	38.28	1.74	6.73	68.82	45.06	8.85	23.76	36.21	4.97	12.52
178-60.47	82	38.23	11.42	35.01	1.82	6.05	65.23	40.98	7.96	24.25	33.02	5.18	15.49
235-80.84	83	38.71	10.58	34.96	2.02	5.76	67.99	42.12	8.45	25.87	33.67	5.19	13.65
154-40.220	84	37.07	9.97	37.67	1.45	6.02	67.14	43.17	8.60	23.97	34.57	5.09	15.43
154-40.206	85	40.08	9.23	39.30	1.25	4.86	69.35	45.63	9.43	23.72	36.20	5.04	15.31
178-60.86	86	37.03	8.82	39.81	1.42	6.31	69.72	47.24	9.58	22.48	37.66	4.77	13.72
154-40.177	87	40.34	9.22	38.90	1.59	6.38	70.64	46.10	9.53	24.54	36.57	5.09	12.16
178-60.49	88	33.64	13.00	34.71	1.79	7.37	63.31	39.13	7.58	24.18	31.55	5.17	14.54
157-40.79	89	36.44	10.98	36.00	1.63	5.58	67.20	42.51	8.82	24.70	33.69	5.18	14.61
154-40.61	90	41.33	9.45	38.08	1.42	6.33	66.48	44.49	9.50	21.99	34.99	5.15	16.33
154-40.159	91	43.14	11.14	36.46	1.67	5.67	63.34	42.00	8.42	21.34	33.58	5.20	18.18
WT (untreated)	92	40.46	10.18	38.50	1.48	6.44	68.19	44.46	9.30	23.73	35.16	5.13	13.71
235-80.128	93	37.81	11.74	34.58	1.89	5.77	66.32	40.44	8.42	25.88	32.02	5.28	14.27
178-60.92	94	36.98	10.81	37.28	1.55	5.56	66.74	42.88	7.99	23.85	34.89	5.05	15.34
15-25.2	95	38.86	10.68	39.65	1.31	6.82	68.86	44.86	9.31	24.00	35.55	5.12	12.33
154-40.218 *	96	39.54	10.43	47.99	1.07	10.36	74.01	37.78	9.78	36.24	28.00	5.42	4.13
157-40.59	97	32.25	10.96	36.88	1.56	7.78	63.66	42.49	8.15	21.17	34.35	4.94	16.04
154-40.191	98	38.25	10.90	37.74	1.45	7.26	66.68	42.77	8.81	23.91	33.96	5.19	13.70
154-40.198	99	40.64	10.58	37.19	1.60	6.86	67.39	43.13	9.04	24.26	34.09	5.24	13.57
31-60.2	100	39.91	9.73	40.32	1.34	6.01	68.70	46.21	9.71	22.48	36.51	4.85	14.21
	Mean	***35.90***	***9.89***	***39.82***	***1.40***	***6.38***	***70.64***	***46.17***	***9.36***	***24.47***	***36.81***	***4.96***	***11.70***

*) Several chemical values (Ash, cellulose, NSC) for clone #154-40.218 appear exceedingly extreme and need to be double-checked, therefore they have not been considered in the analysis of results.

Table 3. Mean values of chemical parameters per sampling position on the *A. donax* stems (1 = top portion of stems; 2 = mid portion of stems; 3 = basal portion of stems) over the 100 clones sampled for chemical analysis.

		DR	Protein	Fiber	Lipid	Ash	NDF	ADF	ADL	Hemic.	Cellulose	pH	NSC
1	**Top**	33.56	13.52	36.19	2.04	7.42	67.09	41.71	7.57	25.38	34.14	5.09	9.93
2	**Mid**	34.38	10.10	37.83	1.59	7.13	68.42	43.70	8.76	24.72	34.94	5.09	12.76
3	**Base**	39.75	6.04	45.46	0.57	4.58	76.42	53.12	11.76	23.30	41.35	4.68	12.40
	Mean	***35.90***	***9.89***	***39.82***	***1.40***	***6.38***	***70.64***	***46.17***	***9.36***	***24.47***	***36.81***	***4.96***	***11.70***

2.4.2. Analysis of correlation among chemical traits.

Also as a consequence of the chemical assays and computation of values utilized for their estimations, several traits related with cellulose and lignin amount resulted positively correlated. For instance, the content of Cellulose and ADL (lignin) resulted highly positively correlated (r = 0.745, p < 0.01. Table 4). Quite interestingly, the accumulation of cellular components such as proteins, lipids and NSC (non-structural carbohydrate) resulted mostly positively correlated, and correlated negatively with fibers, cellulose and lignin (Table 4). For instance, NSC resulted highly negatively correlated with Fiber (r = -0.78, p < 0.01), with ADL (lignin, r = -0.73. p < 0.01) and with cellulose (r = -0.59. p < 0.01). Ash content did not correlate strongly with any of the other traits, suggesting an independent genetic control and independent physiological determination. Therefore, the production of low ash cultivar of *A. donax* seems possible without detrimental effect on other agronomical or chemical traits of interest.

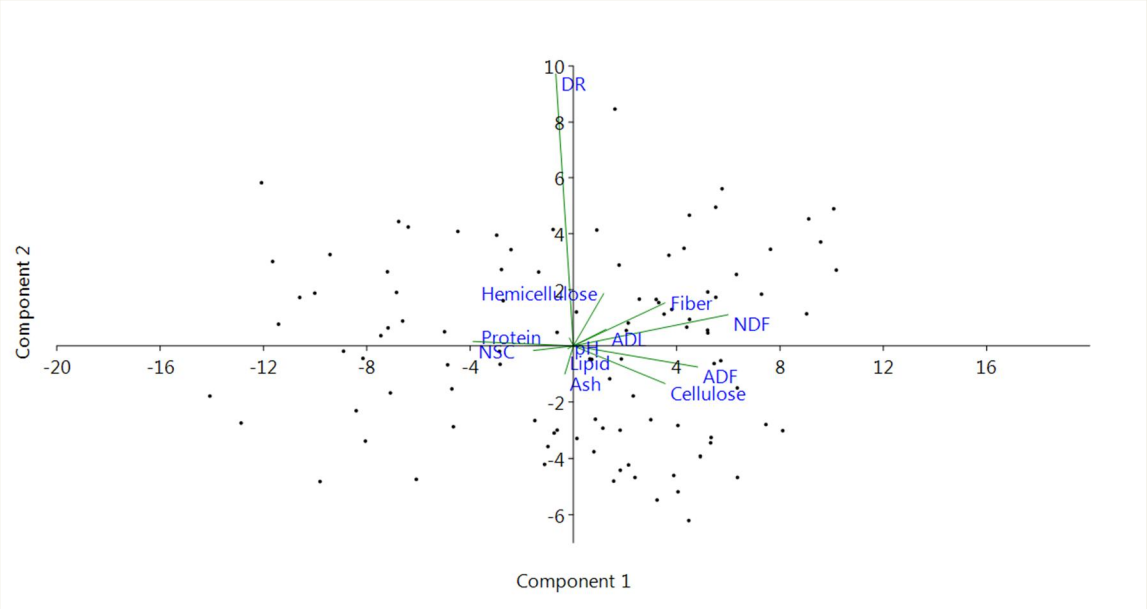
Correlation among traits was also analysed with Principal Component Analysis (PCA), in order to better describe the association between the variation of different traits (Fig. 1). Based on this analysis, general fiber content (ADF, NDF, etc) appear negatively correlate with Protein and NSC contents, as they lay on opposite directions on the first principal component axes, (PC1, accounting for 65.4% of variability). The amount of dry residue (DR) explains most of the

variation along the component 2 (18.7% of variation explained). Variation for ash and hemicelluloses appear also orientated on the PC2 and thus independent (orthogonal) to variation of fibers and cellular components (ig. protein and NSC).

Table 4. Correlation analysis (Pearson's *r*) among chemical traits analysed on 100 *A. donax* clones. *r* values are reported below the diagonal, with $r > 0.5$ and $r < -0.5$ highlighted in red and green, respectively. *P* values are reported above the diagonal.

	DR	Prot.	Fiber	Lipid	Ash	NDF	ADF	ADL	Hemi.	Cell.	pH	NSC
DR		0.792	0.643	0.699	0.000	0.646	0.183	0.247	0.161	0.033	0.000	0.066
Protein	0.027		0.000	0.000	0.000	0.000	0.000	0.000	0.135	0.000	0.000	0.000
Fiber	0.047	-0.757		0.000	0.361	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Lipid	-0.039	0.534	-0.777		0.061	0.000	0.000	0.000	0.050	0.000	0.001	0.000
Ash	-0.371	0.363	-0.092	0.188		0.010	0.000	0.014	0.288	0.000	0.090	0.105
NDF	-0.046	-0.838	0.869	-0.542	-0.256		0.000	0.000	0.000	0.000	0.000	0.000
ADF	-0.134	-0.898	0.734	-0.524	-0.360	0.877		0.000	0.633	0.000	0.000	0.000
ADL	0.117	-0.825	0.836	-0.621	-0.245	0.881	0.854		0.001	0.000	0.000	0.000
Hemic.	0.141	-0.150	0.505	-0.196	0.107	0.523	0.048	0.315		0.628	0.023	0.000
Cellulose	-0.214	-0.862	0.647	-0.454	-0.376	0.815	0.983	0.745	-0.049		0.000	0.000
pH	0.624	0.622	-0.418	0.333	0.171	-0.519	-0.736	-0.439	0.227	-0.790		0.000
NSC	0.185	0.564	-0.777	0.369	-0.163	-0.885	-0.659	-0.726	-0.670	-0.591	0.367	

Figure 1. Principal component analysis (PCA) of chemical traits variation. The graph illustrates the correlation analysis between traits (green vectors and blue tags) using PCA.



2.4.3. Identification of *A. donax* mutant clones characterized by chemical trait expression values of potential interest.

The evaluation of the chemical parameters on 98 mutagenized clones (and two wt untreated clones) enabled us to preliminarily identify clones with extreme expression values (= outliers, defined as showing sample values near or beyond 2 standard deviations from the mean) for one or few chemical traits. These clones are reported in Table 5 and are candidates to further phenotypic evaluation including agronomic test. Three clones (154-40.38, .97 and .20) showed very low Ash content (4.35 - 4.46 % dr, compared with approx. 6.5% dr in wt, and approx. 6.3% dr as population mean). Elevated ash content is a typical negative features of *A. donax* tissues, especially in leaves, when compared with other biomass or bioenergy crops (Monti et al. 2008), therefore mutagenized clones with a reduction of c. 30% in ash content could represent a significant improvement over standard cultivated ecotypes. These three clones were identified from the same plate and level of irradiation, so they may represent the same mutational event. This supports the genetic basis of the observed interesting phenotype.

Four clones (235-80.29, 178-60.49, 228-60.28 e 228-60.43) showed particularly low level of lignin (ASL), ranging from 6.6 - 7.8 %dr, as compared to 9.3 - 9.9 % dr in wt, and 9.4 % dr as population mean. Low lignin clones, associate with standard levels of cellulose and other carbohydrate would be particularly important for improving anaerobic digestion process (and other transformation processes) and therefore biogas and biomethan yield, given the general inhibitory/delaying role played by lignin (Ragaglini et al. 2014).

Table 5. Selection of interesting clones for further chemical and phenotypic evaluation. Values in blue represent lower outlier values (sample value < [mean value −2sd], or near), values in orange represent higher outlier values (sample > [mean value + 2sd], or near). sd = standard deviation.

Clone id	DR	Prot.	Fiber	Lipid	Ash	NDF	ADF	ADL	Hemic.	Cellul.	pH	NSC
154-40.38	39.41	8.62	43.23	1.24	**4.46**	76.89	50.00	10.39	26.89	39.61	4.89	8.80
154-40.97	40.04	9.61	42.68	1.20	**4.35**	73.60	48.47	10.13	25.13	38.34	4.91	11.24
154-40.20	37.83	8.37	43.24	0.99	**4.39**	74.57	49.65	10.38	24.92	39.27	4.83	11.68
235-80.29	34.87	10.84	36.17	**2.04**	6.60	66.59	42.74	**7.75**	23.85	34.98	5.04	13.93
178-60.49	33.64	**13.00**	**34.71**	**1.79**	7.37	63.31	39.13	**7.58**	24.18	**31.55**	5.17	14.54
228-60.28	33.99	11.43	37.15	1.42	6.96	65.62	41.57	**7.28**	24.04	34.30	5.00	14.58
228-60.43	34.88	12.83	**34.96**	1.67	7.47	62.35	38.62	**6.60**	23.73	32.02	5.09	15.68
154-40.61	41.33	9.45	38.08	1.42	6.33	66.48	44.49	9.50	21.99	34.99	5.15	**16.33**
128-60.33	40.17	10.48	37.30	1.42	6.74	64.90	42.46	8.21	22.45	34.24	5.25	**16.46**
22-80.2	36.47	10.85	39.18	1.14	5.96	64.46	43.16	8.71	21.31	34.44	4.97	**17.59**
154-40.159	**43.14**	11.14	36.46	1.67	5.67	63.34	42.00	8.42	21.34	33.58	5.20	**18.18**
WT1	*40.46*	*10.18*	*38.50*	*1.48*	*6.44*	*68.19*	*44.46*	*9.30*	*23.73*	*35.16*	*5.13*	*13.71*
WT2	*35.06*	*9.77*	*40.89*	*1.06*	*6.61*	*70.98*	*46.25*	*9.99*	*24.74*	*36.25*	*5.00*	*11.57*
min	29.78	7.60	34.58	0.70	4.35	62.35	37.78	6.60	21.17	31.29	4.67	8.00
mean	35.90	9.89	39.82	1.40	6.34	70.64	46.17	9.36	24.47	36.90	4.96	11.77
max	43.14	13.00	47.99	2.04	8.79	77.06	51.12	11.01	36.24	40.51	5.42	18.18
sd	3.06	1.03	2.33	0.22	0.81	3.50	2.98	0.82	1.68	2.16	0.14	2.47
mean -2sd	29.77	7.82	35.17	0.95	4.72	63.65	40.20	7.72	21.10	32.58	4.67	6.82
mean + 2sd	42.03	11.95	44.48	1.84	7.95	77.64	52.14	11.00	27.84	41.22	5.24	16.72

2.5 REFERENCES

Frigon, J.C., Guiot, S.R., 2010. Biomethane production from starch and lignocellulosic crops: a comparative review. Biofuel Bioprod. Bior. 4, 447-458.

Hammer, Ø., Harper, D.A.T., and P. D. Ryan, 2001. PAST: Paleontological Statistics Software Package for Education and Data Analysis. Palaeontologia Electronica 4(1): 9pp.

McKendry, P. (2002). Energy production from biomass (part 1): overview of biomass. Bioresource technology, 83(1), 37-46.

Monti, A., Di Virgilio, N., & Venturi, G. (2008). Mineral composition and ash content of six major energy crops. Biomass and Bioenergy, 32, 216-223.

Ragaglini, G., Dragoni, F., Simone, M., & Bonari, E. (2014). Suitability of giant reed (Arundo donax L.) for anaerobic digestion: Effect of harvest time and frequency on the biomethane yield potential. Bioresource technology, 152, 107-115.

Taherzadeh, M.J, Karimi, K., 2008. Pre-treatment of lignocellulosic wastes to improve ethanol and biogas production: a review. Int. J. Mol. Sci. 9, 1621-1651.

CHAPTER 3.

The chloroplast genome sequence of *Arundo donax* L.

3.1. INTRODUCTION

Chloroplasts are cytoplasmic organelles responsible for photosynthesis: their main functions concern the production of ATP and the conversion of CO_2 to carbohydrates through the photosynthetic process. Moreover, chloroplasts are capable of synthesizing amino acids, fatty acids, and the lipid components of their own membranes. Their structure is characterized by a double membrane called the chloroplast envelope and by a third innermost membrane called the thylakoid membrane, which is involved in the production of ATP. The thylakoid membrane forms a network of flattened discs called thylakoids, which are frequently arranged in stacks called grana. The space between the internal layer of the chloroplast envelope and the thylakoid membrane called stroma contains the chloroplast chromosome and a variety of metabolic enzymes, including those responsible for the critical conversion of CO_2 to carbohydrates during photosynthesis (Cooper, 2000). Chloroplasts evolutionarily originated about 1 billion years ago from cyanobacterial-like prokaryote, which were engulfed in an eukaryotic cell establishing a mutualistic relationship and leading to cells capable of photosynthesize. This scenario is known as the Endosymbiotic Theory (Raven and Allen, 2003; McFadden, 2001; Kowallik, 1997). Enunciated for the first time in 1905 and 1910 by the Russian botanist Konstantin Mereschkowsk were then corroborated with microbiological evidence in 1967 by Lynn Margulis. The theory was finally accepted by the scientific community in 1975 following the paper from Linda Bonen and Ford Doolittle (Bonen and Doolittle, 1975), which for the very first time validated the theory with a quantitative characterization of T1 RNAs catalogues of cyanobacteria and chloroplasts. Further evidence came from Hans Kössel who applied the DNA

sequencing to define the primary and secondary structure of maize chloroplast rRNA confirming that the chloroplast genome was more closely related to the bacteria *Escherichia coli* than to eukaryote. Soon after the first chloroplast genome was sequenced: the liverwort *Marchantia polymorpha's plastome* is 121,024 bp in length and became the model genome for the subsequent study concerning the chloroplast. The chloroplast genome consists of circular DNA molecules that includes two copies of an IR region (inverted region), IRa and IRb, that separate large and small single-copy (LSC and SSC) regions, and is present with several copies per organelle. Generally a chroroplast genome ranges between 120 to 160 kb in length and contains approximately 120-130 genes organized as operons and primarily involved in photosynthesis, translation, and transcription. Non-photosynthetic plastids are generally characterized by a smaller number of genes (approx.70), but it can also be bigger (eg. 120 genes in red algae). The chloroplast genes encode both RNAs and proteins involved in gene expression, as well as a variety of proteins involved in the functioning of the photosystems I and photosystem II. The current state of algae and plant chloroplast results from a series of genomic events which lead to a loss of genes during the evolution of the mutualistic relation between the cyanobacteria and eukaryotic cells. As a matter of fact, the analysis of the genome of the cyanobacterium *Synechocystis* allowed to detect the most likely original number of genes which characterized the ancestor of organisms capable of photosynthesize. This study permitted to understood that although the chloroplast genome is independent from the nuclear genome about 90% of the proteins involved in operating in the chloroplast are encoded by nuclear genes, synthetized by cytosolic ribosomes, sorted and transported to the right location within the chloroplast (Cooper *et al*., 2003; Howe *et al*., 2003). This is the consequence of a massive transfer of genes from the cyanobacterium to the nucleus in a phenomena called horizontal gene transfer. About 2,000 nuclear genes contribute to the functioning of chloroplast. Therefore, the mutualism between the cell and the plastid required a whole resetting of the communication

system that operate inside the cell itself. One of the first evidence regarding the dependency of chloroplast from the nucleus came from the analysis of the gene Rubisco small subunit (RbcS), which identified a precursor form produced in the cytosol and then targeted to be transferred to the chloroplast in order to be assembled there (Kanevski and Maliga, 1994). Since the first chloroplast genome, almost 900 accessions were reported from seed plants (National Center for Biotechnology Information (NCBI) Organelle Genome Resources, http://www.ncbi.nlm.nih.gov/genomes/).

The comparison among choroplast genomes showed that both gene content and gene order are well conserved among seed plant species even if there are several exceptions including rearrangements and loss of genes and entire gene families, and inversions and losses of inverted region (IRa and IRb). For example several studies demonstrated that the chloroplast genome of grasses (Poaceae) is characterized by the presence of three inversions when compared to non-grasses species; the largest inversion being 28 kb in length, the second inversion being 6 kb in length and the third inversion characterizing the *tRNAT* gene. The smallest inversion is considered as a synapomorphy since it is unique to grasses while the other two inversions were found in some other non-grass families such as Restionaceae, Ecdeiocoleaceae, and Joinvilleaceae, The sister group of Poaceae is still matter of debate since both Ecdeiocoleaceae and Joinvilleaceae families are closely related to them. The family Joinvillaceae shares morphological traits with Poaceae, however, based on mitochondrial and chloroplast genomic sequence comparisons, the sister group would be Ecdeiocoleaceae.

The sequencing and characterization of chloroplast genome is a valuable source for interpreting the phylogenetic relationship existing among plants. At higher taxonomic levels (e.g. family and order level) usually protein-coding genes are used in order to deciphering the evolution of taxa, while at lower taxonomic levels (e.g. genus and species levels) complete chloroplast genome sequence is more suitable to untangling the phylogenetic relationship

existing among taxa. From a plant breeding standpoint, understanding the phylogenetic relationship between cultivated crops and their wild relatives facilitate the introgression of specific advantageous traits in the varieties of interest. Additionally, it can provide information on ways to create new synthetic species. Specifically, synthetic hexaploid wheat plants have been re-produced by controlled hybridization of diploids wild wheat (Ogbonnaya *et al*., 2013); the cultivated triploid energy crop *Miscanthus* × *giganteus* (3n = 57, x = 19) can at least be theoretically re-synthesized from controlled hybridizations of diploid ancestors *M. sinensis* and *M. sacchariflorus* (Clifton-Brown *et al*., 2008).

3.2. OBJECTIVE

The objective of this work was to obtain the first fully sequenced chloroplast genome of *Arundo donax*, in order to provide a suitable resource for the development of marker useful for the classification of this species and in the long term to contribute to the genetic improvement of the species.

3.3. MATERIALS AND METHODS

3.3.1. Plant material and DNA extraction

A. donax fresh leaves were collected from the same natural stand (cluster) utilized as source of plant material for the mutagenic experiment, and freeze-dried. Total genomic DNA was extracted from a pool of these leaves using The Wizard® Genomic DNA Purification Kit (Promega). DNA was quantified and quality checked using Infinite 200 NanoQuant

spectrophotometers (**Tecan Group Ltd,** Männedorf, Switzerland), while DNA integrity was verified by a 0.8% agarose gel.

3.3.2. Primer design and chloroplast genome sequencing

The complete chloroplast genome sequence of *A. donax* was amplified from the total genomic DNA by a long-PCR-based approach. Amplicons ranging from about 6,000 to 15,000 bp and supposedly covering the entire sequence were obtained using primers designed on chloroplast genome sequences of three related species, *Setaria italica* (Genbank code KJ001642.1), *Phragmites australis* (Genbank code EU732697.1), and *Brachypodium distachyon* (Genbank code U170609.1),. Long range PCR were performed using GoTaq Long PCR Master Mix with default conditions, with the exception of the annealing temperatures, which were changed accordingly with the pairs of primers used in each reaction. The list of primers is reported in Table 1. An aliquot of each amplicons were checked for quality, integrity and size on 1% agarose gel. Subsequently the total amount of each amplicon were loaded on an agarose gel and purified using QIAquick PCR Purification Kit (Qiagen, Hilden, Germany). Each purified amplicon was quantified using Infinite 200 NanoQuant spectrophotometers and the concentrations were normalized at 100 ng/ul. Fragments were pooled together to obtain a final amount of 1 ug of product. Sequencing was carried out by Illumina MiSeq (Illumina, San Diego, CA, USA). A total number of 111,434 reads were analyzed and de novo assembled using Bowtie2 software (ver. 2.2.3, http://bowtie-bio.sourceforge.net/bowtie2/index.shtml; Langmead and Salzberg 2012) to produce 235 superconting. The whole set of superconting were then aligned with the chloroplast genome sequence of *S. italica*: three supercontigs covered the entire sequence of *S. italica*. Most of the other supercontigs belong to *Burkholderia*, a microorganism, and therefore considered as contaminants.

Table 1. Primers name and sequence

Prime name	Sequence 5'-3'
CP_1F_donax_ trnfM-CAU	CGTCTCTAATTCAAAACCGAACATG
CP_1R_donax_ trnC-GCA	AAAGAGCAGTTTTCTAGTGTTAGCA
CP_2F_donax_ petN	AAAGACTACCATTAAAGCAGCCCAAG
CP_2R_donax_ rps2	CAACCTCGTCATATATTTGATCCCGC
CP_3F_donax_ rps2	TTGAAAGAGATGATAGAAGCGGGAGT
CP_3R_donax_ rps14	ATGGTTTATGCATGTTTGTTACCGGG
CP_4F_donax_ trnR-UCU	ATTAGACAATGGACGCTTTTCTTTCG
CP_4R_donax_ trnF-GAA	ACCCTTTCTTGTGCATCATCCTAGTA
CP_5F_donax_trnT-GGU	GAGTATTGCTTTCATACGGCGGGAG
CP_5R_donax_trnW-CCA	AGGTTCAAATCCTACAGAGCGTGAT
CP_6F_donax_petL	CTGCTTTAACTATAACCCCAGCTCT
CP_6R_donax_petB	ACACACTCATATTCCAGAGATACCGA
CP_7F_donax_psbN	AGGATCGAATCTATGGAAGCATTGGT
CP7R_donax_rpl22	TCACTAAAGCCGAAGTCAGTAGGGG
CP_8F_donax_rps3	ATCAACGAGCAGAATGGGAAGACTG
CP_8R_donax_rps7	AGGAAGCGACTCATAGAATGGCAGA
CP_9F_donax_trnL-CAA	GGACTCGAACCTCCACGCTCTTTAG
CP_9R_donax_trnA-UGC	TCCTTTTGCCACATTTCGCTCAAAG
CP_10F_donax_ycf68	CATGGCGTACTCCTCCTGTTTGAAT
CP_10R_donax_trnL-UAG	TCCTAAGAGCAGCGTGTCTACCAAT
CP_11F_donax_rpl32	ACAGAAAGAAGGATTGATTTGCGGC
CP_11R_donax_trnN-GUU	AAACAAGGATAGCGAACGGGTTACC
CP_12F_donax_rps15	ACTTCTCCCGAATATCCAACTGACTGA
CP_12R_donax_trnV-GAC	GGCTTTCTTTCCGCACTCTTATGGA
CP_13F_donax_ycf68	AAACCTGCTCCCATTTCGAGTCAAG
CP_13R_donax_trnI-CAU	TCGACTTTCCCTCCTATGCTCTGAG
CP_14F_donax trnI-CAU	GGAGGGAAAGTCGATTTATGGATGGA
CP_14R_donax MatK	TGATTATGGATTAAATGGTGCCGAGC
CP_15F_donax_MatK	TCCCAAAATTCTGCTGATACATTCGA
CP_15R_donax_trnfM-CAU	AAATCAACCAACGTCGACTATAACCC

3.3.3. Chloroplast gene annotation

Chloroplast genome annotation was performed using DOGMA (Wyman *et al*., 2004), cpGAVAS (Liu *et al*., 2012), Open Reading Frame (ORF) finder program from the National Center for Biotechnology Information and Basic Local Alignment Search Tool (BLAST; BLASTN, PHI-BLAST and BLASTX). A circular map of the chloroplast genome was drawn using Genome VX (http://wolfe.ucd.ie/GenomeVx/). A-T contents were evaluated by MEGA 7 (http://www.megasoftware.net; Tamura et al 2013). Repeating sequences were analyzed by Tandem Repeats Finder, ver. 4.07b (https://tandem.bu.edu/trf/trf.html; Benson 1999) and by Phobos ver. 3.3.12.

3.3.4. Phylogenetic analysis

Bayesian inference analysis was conducted using MrBayes software (MrBayes 3.2.6, http://mrbayes.sourceforge.net; Ronquist *et al*., 2012) with the time reversible model. Two independent runs of $2x10^5$ of Markov Chain Monte Carlo were performed with a burn-in-fraction of 0.01, each starting with a different random tree. The protein codin gene *trnK/matK* of A. donax and 39 grasses species were align using the algorithm CLUSTAL, with default parameters, implemented in MEGA. DNA sequences were gathered from GenBank (table 2). The whole set of sequences were aligned using MEGA 7 with the default parameter of the CLUSTAL algorithm. The results were analysed using FigTree 1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/

Table 2. Species selected for the phylogenetic, Genbank code and taxonomic classification.

Species	Genbank code	Subfamily/Family
Arundo donax		Arundinoideae
Arundo formosana	FR821319.1	Arundinoideae
Phragmites mauritianus	HF558517.2	Arundinoideae
Molinia caerulea	HE586092.1	Arundinoideae
Phragmites australis	EU732697.1	Arundinoideae
Monachather paradoxus	HE574410.1	Arundinoideae
Elytrophorus spicatus	J920230.1	Arundinoideae
Hakonechloa macra	KJ920232.1	Arundinoideae
Chasmanthium latifolium	HE573997.1	Panicoideae
Setaria italica	KJ001642.1	Panicoideae
Dichanthelium acuminatum	HF558501.1	Panicoideae
Panicum bisulcatum	FR821330.1	Panicoideae
Setaria viridis	KT289405.1	Panicoideae
Miscanthus sinensis	KR822688.1	Panicoideae
Panicum virgatum	HQ822121.1	Panicoideae
Zea mays	NC_001666.2	Panicoideae
Aristida purpurea	KJ920224.1	Aristidoideae
Bambusa arnhemica	KJ870989.1	Bambusoideae
Arundinaria gigantea	JX235347.1	Bambusoideae
Bambusa oldhamii	FJ970915.1	Bambusoideae
Phyllostachys edulis	KP684145.1	Bambusoideae
Bambusa emeiensis	HQ337797.1	Bambusoideae
Hordeum vulgare subsp. vulgare	NC_008590.1	Pooideae
Triticum aestivum	NC_002762.1	Pooideae
Brachypodium distachyon	U170609.1	Pooideae
Triticum urartu	KJ174105.1	Pooideae
Lolium perenne	AM777385.2	Pooideae
Triticum monococcum	KC912690.1	Pooideae
Diarrhena obovata	KM974739.1	Pooideae
Avena sativa	KM974733.1	Pooideae
Puelia olyriformis	NC_023449.1	Puelioideae
Merxmuellera decora	EU400740.1	Danthonioideae
Danthonia californica	KJ920229.1	Danthonioideae
Eriachne stipacea	KJ920231.1	Micraioideae
Isachne distichophylla	KJ920233.1	Micraioideae
Chionochloa macra	KJ920227.1	Chloridoideae
Leptaspis cochleata	HF558509.1	Pharoideae
Ecdeiocolea monostachya	DQ257528.2	Ecdeiocoleaceae
Georgeantha hexandra	DQ257531.2	Ecdeiocoleaceae
Joinvillea plicata	DQ257535.2	Joinvilleaceae

3.4. RESULTS

3.4.1. General features of the *A. donax* chloroplast genome

The complete chloroplast genome of *A. donax* resulted 139,353 bp in length (figure 1). The chloroplast genome sequence is characterized by a pair of inverted regions (IRa and IRb) of 22,227 bp each separated by a small-single-copy-region of of 12,275 bp (SSC) and a large-single-copy-region of 82,124 bp (LSC). The genome includes 112 individual genes including 72 protein coding genes, 30 tRNA, 6 rRNA, 3 open reading frames and one pseudogene. IR regions contain 25 duplicated genes. Based on their predicted functions, these genes can be divided into three categories, genes related to transcription and translation, genes related to photosynthesis, genes related to the biosynthesis of amino acids, fatty acids, etc., and some functionally unknown genes Ten genes contain one intron, ycf3 genes has two introns (table 3).

Table 3. Genes containing introns and their respective length.

Gene	Strand	Start	End	Exon I	Intron I	Exon II	Intron II	Exon III
atpF	+	35042	36420	160	812	407		
ycf3	-	44229	46197	134	730	229	720	156
rpl2	-	83152	84633	394	660	428		
ycf15	+	86986	87497	135	289	88		
ndhB	-	89077	91318	870	619	753		
ndhA	-	114198	116306	553	1020	536		
ndhB	+	131075	133311	775	704	758		
ycf15	-	134889	135397	138	286	85		
rpl23	+	137452	137797	138	64	144		
rpl2	+	137804	139300	409	660	428		

Almost half part (49.0%) of the *A. donax* chloroplast genome consists of gene-coding regions (42.6% protein and 6.4% RNA coding regions), whereas the intergenic spacers (including 11 introns) account for the 51.0%. The overall A-T content of *A. donax* chloroplast genome is

61.7%, which is similar to the A/T content found in grasses, while GC content is 38.3% (Cotton *et al*., 2015; Wang and Gao, 2015).

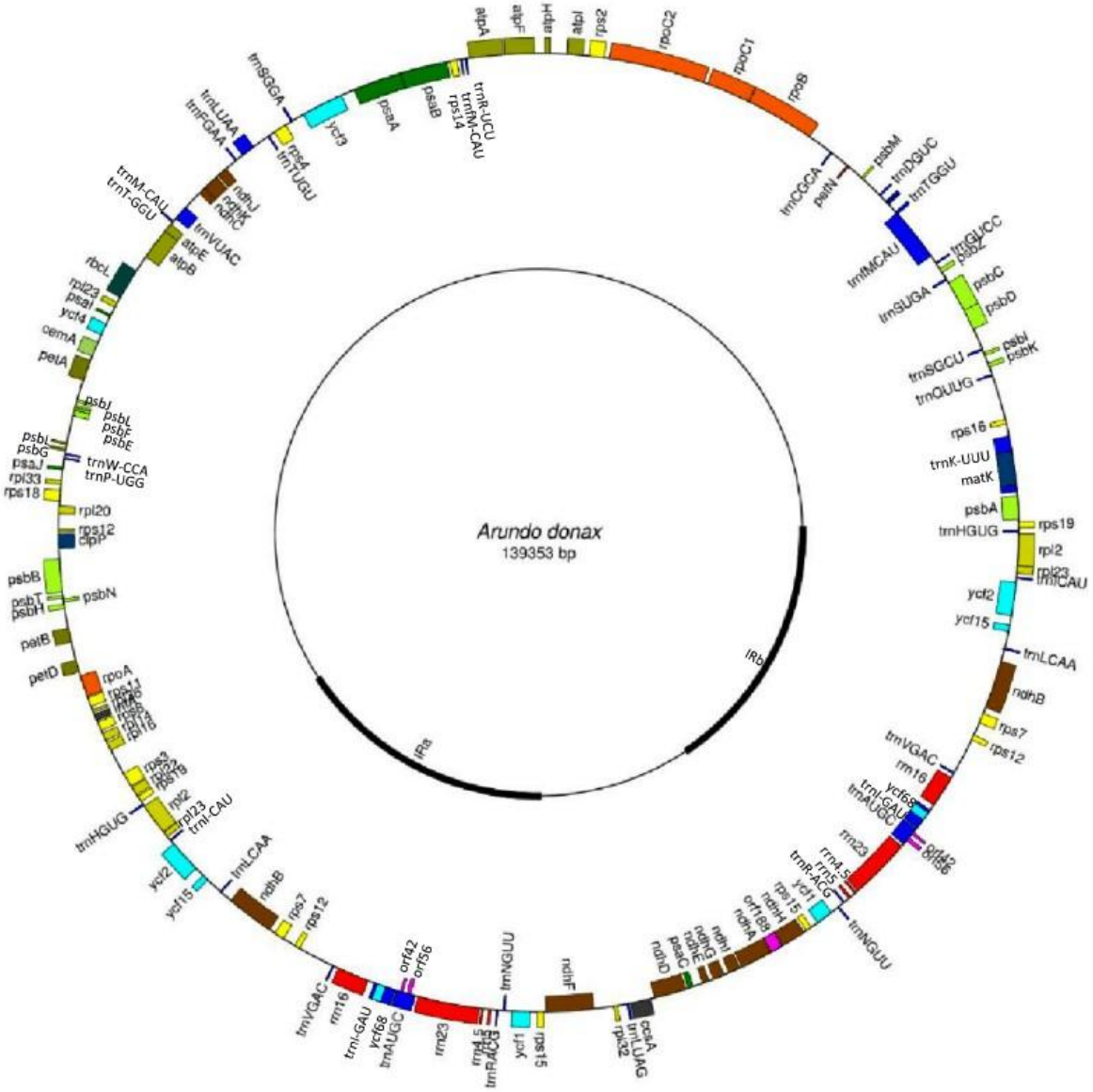


Figure 1. The complete chloroplast genome map of *Arundo donax*, total length 139,353 bp. The inner circle reported the inverted repeats (IRa and IRb; 22,227 bp each), which separate the large single copy region (LSC; 82,124 bp) from the small single copy region (SSC; 12,775 bp). Genes drawn outside the circle are transcribed clockwise, while genes drawn inside the circle are transcribed counterclockwise.

3.4.2. Analysis of repeat unit in *A. donax chloroplast* genome

A total number of 31 repeat units were found. The repeated size range from 14 bp, the smallest one, to 117 bp the biggest one. The majority of repeat units size is 22 bp. All the repeats

analyzed fall in the category of direct sequences with a variable number of copy ranging from 2 to 5 copies per repeat units. No palindromic or dispersed sequences were found during the analysis. Repeat units were also characterized as mononuleotide, dinucleotide, trinucleotide, tetranucleotide, pentanucleotide, hexanucleotide, 7-nucleotide, 8-nucleotide, 9-nucleotide and 10-nucleotide. The main repeat units are reported in Table 4.

Table 4. Classification of repeat units found in the chloroplast genome of *Arundo donax*. The slash symbol divide the different nucleotide(s) which characterize the repeats.

Repet unit length	**Unit**
Mononucleotide	A
Dinucleotide	AG/T
Trinucleotide	AAC/G/T/CT/GG
Tetranucleotide	AAAG/T/TG/TT/CC/CG/CT/GC/GAT/GGAT
Pentanucleotide	AAAAC/G/T/TC/CAT/CTT/GAC/GAT/GCG/GCT/TAG/TAT/TGC/CACT/CACC/TCC/TCG
Hexanucleotide	AAAAAG/T/TC/CT/CC/GG/GT/GAT/CGG/CTC/GAG/TAC/TAG/TAT/TCC/TTG/CGGG/CTAC/CTTC/GAAT/GATC/GCAT/CTATC/GATT/GCAG/GCAT/GCCC/GCGG/GAGAT
7-nucleotide	AAAAAAG/T/C/TA/GAT/TAT/TAG/TCT/GAAA/GATA/TAAT/TACC/TACT/TTCC/CCCCA/CTACT/TGATC/TCTGG/TGATG/GAGGAT
8-nucleotide	AAAATACC/TACCT/GAAAT/GATG/TCTC
9-nucleotide	AAAATATAC
10-nucleotide	AAAGATACTG/CAATATTC

3.4.3. Phylogenetic analysis

The Bayesian inference analysis was conducted on 40 species belonging to ten subfamily of Poaceae and two family, Joinvinlleaceae and Ecdeiocoleaceae which are considered as the outgroup o grasses. The analysis was performed using the sequence of the combination between a transcription and a protein coding gene *trnK/matK*, which is reported in literature as one of the gene more useful to define the phylogenetic relationship among species. The phylogenetic tree clearly show that the clades belonging to the species of Puelioideae, Bambusoideae, Pooideae and Aristoideae were put in the same cluster but forming a robust picture of the relationship among the species within the same subfamily with a posterior probability close to 100%. Moreover, the subfamily Pharoideae with one species, can be

considered as strongly supported, and separated from the rest of the subfamily analysed (posterior probability 97%). The other subfamilies analyzed were not clearly divided into cluster. Infact the clades characterizing these taxa were not roboustely supported, as showed by the prosterior probability which is 51% and 65% (figure 2).

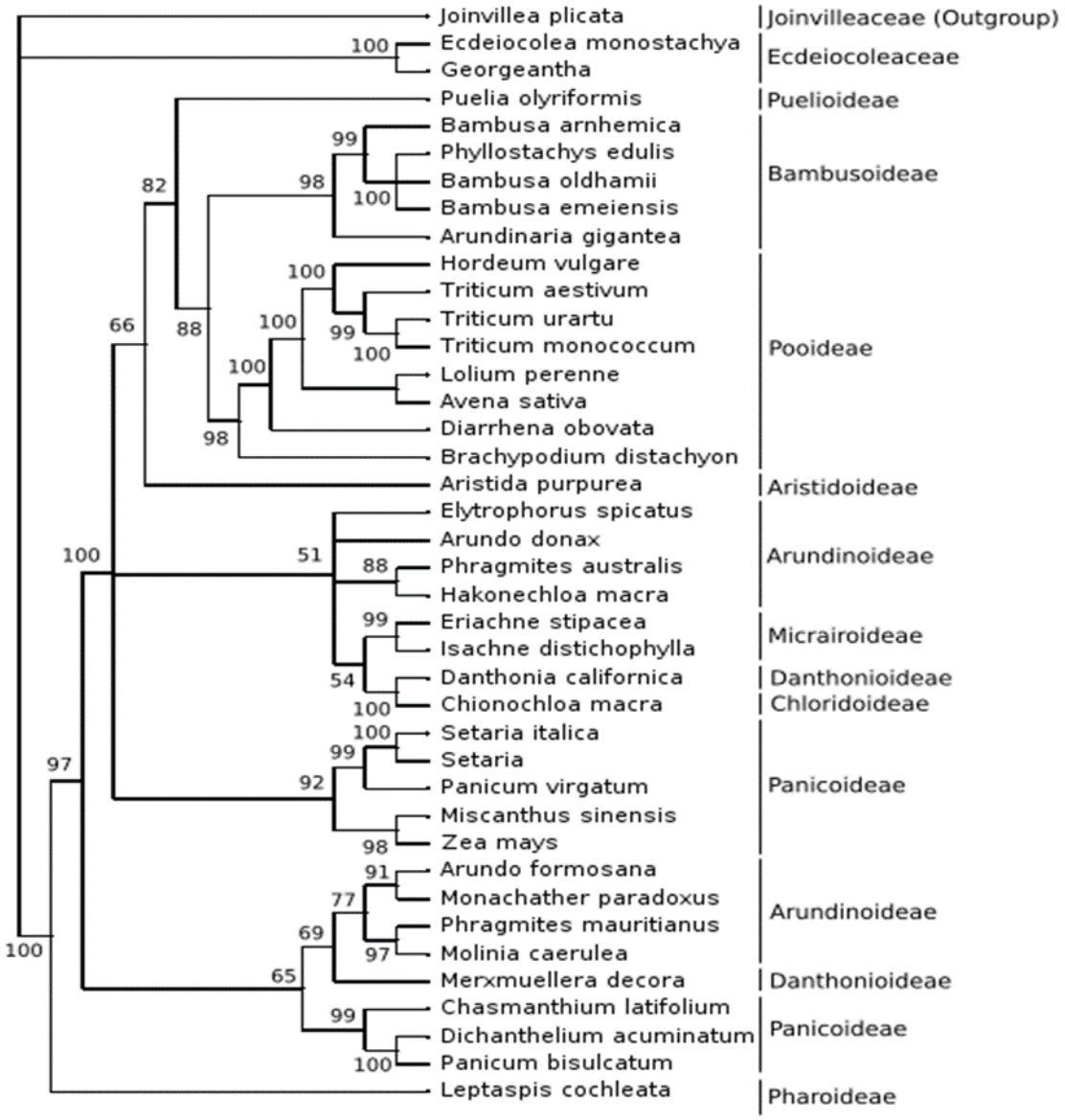


Figure 2. Phylogenetic tree of 40 species belonging to Poaceae family and *trnk/matK* plastid protein-coding genes using Bayesian analysis. Numbers close to the nodes indicate Bayesian posterior probability confidence values.

3.5. DISCUSSION AND CONCLUSION

The complete chloroplast genome of *A. donax* were obtained after sequencing PCR amplicons which were extracted from the genomic DNA using primers deigned on closely related species. The total length of is 139353 bp, with IR of 22,227 bp and the LSC and SSC region respectively of 82,214 bp and 12,775 bp. The complete sequence contain 112 unique genes with 25 of them duplicated in the IR regions for a complessive number of 137 genes. The number of genes is in the range of that that were reported from other chloroplast genome sequenced in Poaceae (Zhang *et al*., 2012; Ye *et al*., 2014; Middelton *et al*., 2014, Wang and Gao, 2015, Tanaka *et al*., 2016). The analyses of repeate units a large number of repats located both, in intergenic and intron regions of LSC and SSC regions. The phylogenetic reconstruction of Poaceae, considered the main representive of this taxa, was elaborated using the *trnK/matK* sequences of 40 species, including three species, one from Joinvinlleaceae and two from Ecdeiocoleaceae, which several authors had identified as the sister group of Poaceae and used in their phylogenesis (Doyle *et al*., 1992; Melvin *et al*., 1996; Khidir et al., 1999; Zhang *et al*., 2000; Melvin *et al*., 2007; Soreng *et al*., 2015).

The relationship here presented put in evidence that within Poaceae the classification of some taxa can be very tricky. As showed by the tree the species belonging to the three subfamily Arundinoideae, Panicoideae and Danthoniodeae are mixed and not clustered together as expected. The interpretation made with the analyses show that the these taxa are closely related and therefore in order to separate them into different cluster on the basis of the subfamily they belong to, a higher number of markers is needed, considering not only protein coding region but also regions with a higher rates of mutations (i.e intergenic and intronic regions). In fact in several studies (Melvin *et al*., 2007; Aliscioni *et* al., 2012; Cotton *et al*., 2015) regarding the

phylogeny of grasses, were shown that the relationship between Micrairoideae/Arundinoideae and Danthonioideae/ Chloridoideae were weakly supported.

The emerging interest in *A. donax* as energy crop require the increase in knowledge about its taxonomic positioning. Therefore, the complete sequence of chloroplast obtained will be a useful source of DNA marker for the identification and classification of this species. In fact the choloroplast genome is an incredible tool that provide information about the evolutionary radiation from the area of origin and to reconstruct the history that characterize its diffusion in the world.

3.6 REFERENCES

Aliscioni S, Bell HL, Besnard G, Christin PA, Columbus JT, Duvall MR, Edwards EJ, Giussani L et al (2012) New grass phylogeny resolves deep evolutionary relationships and discovers C4 origins. Grass Phylogeny Working Group II. The New Phytologist, 193, 304-312.

Barker NP, Linder HP, Morton CM, Lyle M (2003) The paraphyly of cortaderia (Danthonioideae; Poaceae): evidence from morphology and chloroplast and nuclear DNA sequence data. Annals of the Missouri Botanical Garden, 90, 1-24.

Benson G, (1999) "Tandem repeats finder: a program to analyze DNA sequences"
Nucleic Acid Research, 27, 573-580.

Bonen L, Doolittle WF (1975) On the Prokaryotic Nature of Red Algal Chloroplasts (ribosomalRNA/oligonucleotidecataloging/evolution/endosymbiosis). Proceedings of the National Academy of Sciences, 72, 2310-2314.

Conant GC, Wolfe KH (2008) GenomeVx: simple web-based creation of editable circular chromosome maps. Bioinformatics, 24, 861–862.

Cotton JL, Wysocki WP, Clark LG, Kelchner SA, Pires JC, Edger PP, Mayfield-Jones D and Duvall MD (2015) Resolving deep relationships of PACMAD grasses: a phylogenomic approach. BMC Plant Biology, 78, 11-151.

Doyle JJ, Davis JI, Soreng RJ, Garvin D, Anderson MJ () Chloroplast DNA inversion and the origin of the grass family (Poaceae). Proceedings of the National Academy of Sciences, 89, 7722-7726.

Duvall MR, Davis JI, Clark LG, Noll JD, Goldman DH (2007) Phylogeny of the Grasses (Poaceae) Revisited. Aliso: A journal of systemayics and evolunionary botany, 23, 237-247.

Howe CJ, Barbrook AC, Koumandou VL,. Nisbet ER, Symington HA, Tom F. Wightman TF (2003) Evolution of the chloroplast genome. Phylosophical transaction of the Royal society B: biological science, 29, 99–107.

Jordan GE, Piel WH (2008) PhyloWidget: web-based visualizations for the tree of life. Bioinformatics, 24, 1641-1642.

Kanevski I, Maliga P (1994) Relocation of the plastid rbcL gene to the nucleus yields functional ribulose-1,5-bisphosphate carboxylase in tobacco chloroplasts. Proceedings of the National Academy of Sciences, 91, 1969-1973.

Khidir WH, Lawrence AA, Hongping L (1999) Phylogeny of Poaceae Inferred from matK Sequences. Annals of the Missouri Botanical Garden, 86, 835-851.

Kowallik KV (1997) Origin and Evolution of Chloroplasts: Current Status and Future Perspectives. In: Schenk H.E.A., Herrmann R.G., Jeon K.W., Müller N.E., Schwemmler W. (eds) Eukaryotism and Symbiosis. Springer, Berlin, Heidelberg.

Geoffrey MC (2000) The Cell, 2nd edition: a molecular approach. NCBI Bookshelf.

Liu C, Shi L, Zhu Y, Chen H, Zhang J, Lin X, Guan X (2012) CpGAVAS, an integrated web server for the annotation, visualization, analysis, and GenBank submission of completely sequenced chloroplast genome sequences. BMC Genomics, 13, 715.

Mayer C (2006-2010) Phobos 3.3.11, <http://www.rub.de/spezzoo/cm/cm_phobos.htm>

McFadden GI (2001) Chloroplast Origin and Integration. Plant Physiology, 125, 50–53.

Middleton CP, Senerchia N, Stein N, Akhunov ED, Keller B, et al. (2014) Sequencing of chloroplast genomes from wheat, barley, rye and their relatives provides a detailed insight into the evolution of the Triticeae tribe. PLoS One, 9.

Raven JA, Allen JF (2003) Genomics and chloroplast evolution: what did cyanobacteria do for plants? Genome biology, 4, issue 3, 209.

Ronquist F, Huelsenbeck JP (2003) MRBAYES 3: Bayesian phylogeneticinference under mixed models. Bioinformatics, 19, 1572-1574.

Soreng RJ, Peterson PM, Romaschenko K, Davidse G, Zuloaga FO, Judziewicz EJ, Filgueiras TS, Davis JI, Morrone O (2015) A worldwide phylogenetic classification of the Poaceae (Gramineae). Journal of Systematics Evolution, 53, 117-137.

Sudhir Kumar, Glen Stecher, and Koichiro Tamura (2015) MEGA7: Molecular Evolutionary Genetics Analysis version 7.0. Molecular Biology and Evolution.

Tanaka H, Hirakawa H, Muguerza M, Hashiguchi M, Tabata S, Akashi R, Sato S (2016) The complete chloroplast genome sequence of Zoysia matrella (L.) Merr. Crop Science, 56, 1206–1212.

Ye CY, Lin Z, Li G, Wang YY, Jie Qiu J, Fu F, Zhang H, Chen L (2014) Echinochloa Chloroplast Genomes: Insights into the Evolution and Taxonomic Identification of Two Weedy Species. PLoS ONE 9.

Wang S, Gao LZ (2016) Complete chloroplast genome sequence of green foxtail (Setaria viridis), a promising model system for C4 photosynthesis. Mitochondrial DNA. Part A, DNA Mapping, Sequencing, and Analysis, 27, 3707-3708.

Wyman SK, Jansen RK, and Boore JL (2004) Automatic annotation of organellar genomes with DOGMA. Bioinformatics, 20, 3252–3255.

CHAPTER 4

Cytogenetics investigations of *Arundo donax* L. evolutionary origin

4.1. INTRODUCTION

The worldwide distribution of *Arundo donax* is the result of the human activities, which spread this species along with their migration. This species is considered as an archeophyte since its introduction in the Mediterranean basin was dated between the 1500 AD, during the diffusion of agriculture in the Neolithic, and the discovered of America (Hardion *et al*., 2014). Throughout the whole geographical distribution this species is sterile and therefore not able to produce viable seeds. Evidences of seed production were only reported for population recovered in the Middle east (Saltonstall *et al*., 2010). The several uses of *A. donax* such as agriculture, fodder, construction, weaponry, fishing, hunting, music, erosion control, medicine and fuel made this species cultivated all around the world from Asia, Europe, Africa to United States, where especially in Southern California was largely employed along ditches as erosion control (Saltonstall *et al*., 2010; Lambert *et al*., 2010). Although the origin of *A. donax* was long debated (Mariani *et al*., 2010; Hardion *et al*., 2012; Hardion *et al*., 2014; Hardion *et al*., 2015) there are several proofs supporting the hypothesis of its ancient introduction from eastern to western Eurasia (Hardion et al., 2014) making giant reed as one of the oldest invasive species. The eastern origin is supported by both by cytogenetic and molecular evidence (Hardion *et al*., 2012; Hardion *et al*., 2014). The ploidy level has been investigated in several population distributed worldwide defining *A. donax* as subdiveded in two cytotypes. The lower cytotypes (2n = 12x, 72 chromosomes) has been reported in Thailand, India and Uzbekistan (Hardion *et al*., 2014) while the higher cytotypes (2n = 18x, 108-110 chromosomes) characterize the Europe and America area of distribution. The higher ploidy level could explain the sterility of this

species since it could interfere with the production of meiocytes. Recently Bucci *et al* (2013) proposed two different scenarios aiming at clarifying the origin of *A. donax.* The authors suggest that *A. plinii* (2n = 12x, 72 chromosomes) could be the ancestor of *A. donax* and elaborated two different hypotheses that could explain the ploidy level 2n = 18x with 108-110-chromosome genome of giant reed. The first scenario proposed that *A. plinii* experienced a poliploidization event that lead to the formation of a fertile tetraploid characterized by 144 chromosome. The admixture of population with different ploidy level but still capable of crossing led to a fusion of reduced and unreduced gametes, respectively with 36 and 72 chromosomes, producing a sterile triploid with 108 chromosome. The presence of population with 110 chromosomes could be due to aneuploidy. The second scenario suggests instead an interspecific origin for *A. donax*. The fertile tetraploid A. plinii (or a related species) crossed with *Phragmites australis*, characterized by 96 chromosomes, and this cross produced a sterile hybrid with 120 chromosomes, which produced the sterile *A. donax* after losing 10 chromosomes.

4.2. OBJECTIVE

The objectives of this work was to test the hypothesises that A. donax genome is derived from polyploidization events involving *A. plinii*, using a cytogenetic approach. The results of this work would shed light on the evolutionary origin of *A. donax* and could indicate a possible strategy to produce a synthetic highly productive and fertile new species, to be exploited in plant breeding and as a crop.

4.3. MATERIALS AND METHODS

Plant materials consist of natural accessions of *Arundo donax*, *Arundo plinii*, *Phragmites australis*. *A. donax* accession was collected in the Po valley, while *A. plinii* and *P. australis* were collected near the river bank of Reno in Bologna, Italy.

A. donax genome was probed by fluorescent in-situ hybridization (FISH) and genomic in-situ hybridization (GISH) with A. plinii and P. australis probes. Four types of probes were used: pTa71 rDNA from *Triticum aestivum*, was labelled with digoxigenin-11-dUTP (Roche), pTa794 rDNA from *Triticum aestivum* was labelled with tetramethyl-rhodamine-5-dUTP (Roche) and total genomic DNA from *A. plinii* and *P. australis*. *A. plinii* was labeled by digoxigenin-11-dUTP (Roche), while *P. australis* was labelled by tetramethyl-rhodamine-5-dUTP (Roche). All probes were labelled by nick translation according to manufacturer's instructions (Roche). FISH and GISH procedure were adapted from Hasterok et al. (2001) with modifications. The Hybridization mix consisted of 100% formamide, 50% dextran sulphate, 20XSSC. The ribosomal DNA probes and the genomic DNA probes were mixed to a final concentrations of 100 ng per slide.

4.4. RESULTS

The chromosome count at metaphase plate confirmed the number of chromosomes previously defined by Bucci *et al.* (2013). The FISH experiment with the rDNA probes allowed us to physically map the position of the pTa794 and pTa71 genes on the genome of *A. donax* and *A. plinii*. The genome of *A. donax* (Fig. 1) was showed to have six pTa794 rDNA sites, all located proximally to the centromere, and six distal pTa71 rDNA sites. The genome of *A. plinii* (Fig.1B), instead, appeared as characterized by four proximal pTa794 rDNA sites and four

(three larger and one smaller) pTa71 rDNA sites, all located distally next to the telomere. The GISH experiment on *A. donax* chromosome preparations (Fig. 2) was performed using total genomic DNA from *A. plinii* and *P. australis* as probes. Six chromosomes of *A. donax* clearly hybridized with the genomic DNA of *P. australis*, one chromosome has whole arm hybridized with the probe. The hybridization with *A. plinii* instead was not so evident, but it showed a weakly dispersed signal along all chromosomes.

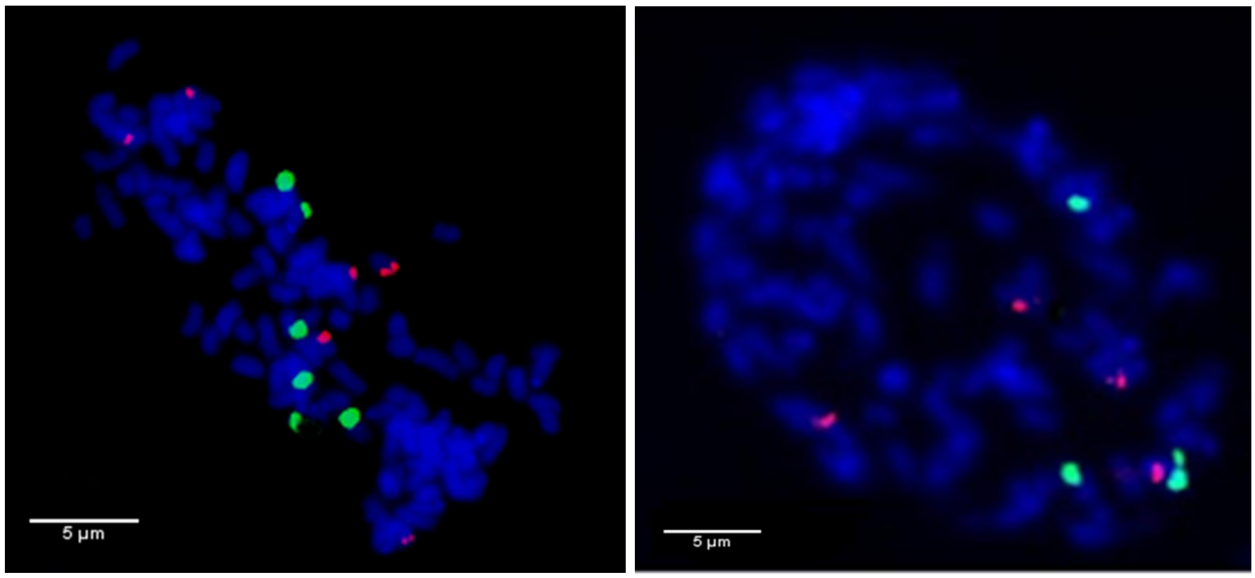


Figure 1. FISH images of somatic metaphase chromosome spread; A) A. donax shows six prossimal pTa794 rDNA sites (red), and six distal pTa71 rDNA sites (green) ; (B) A. plinii shows four prossimal pTa794 rDNA sites (red), and four distal pTa71 rDNA sites (green

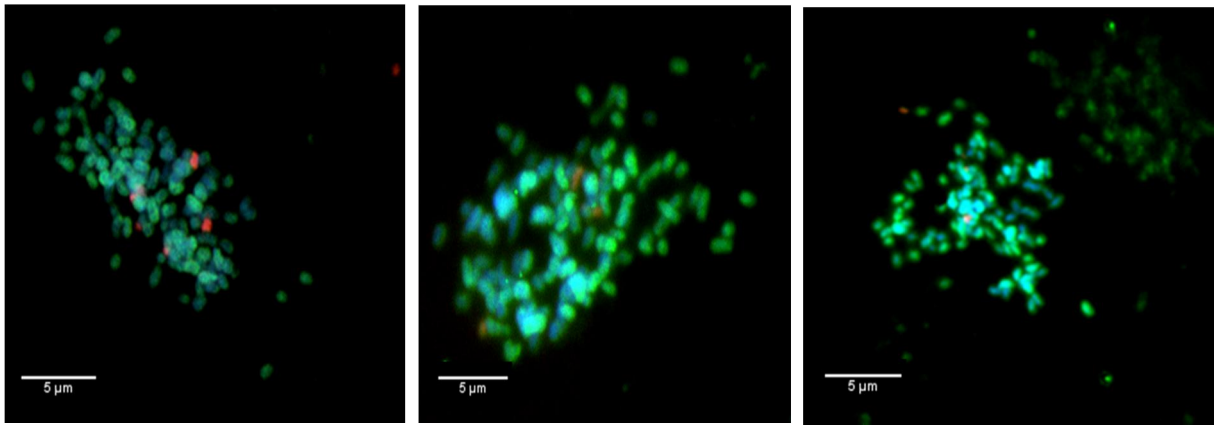


Figure 2. GISH images of *A. donax* somatic metaphase spread with total genomic DNA from *A. plinii* (green) and from *P. australis* (red) as probes.

4.5. DISCUSSION AND CONCLUSIONS

The FISH pattern may suggest an autopoliplody origin of *A. donax* as the results of a cross between a partial fertile tetraploid (144 chromosomes) and a diploid (2n = 72 chromosomes) individual of *A. plinii*, that produced a sterile triploid (108 chromosomes) follow by an acquisition of two chromosomes, as proposed by Bucci *et al* (2013). GISH results were less clear and additional investigations are required in order to support these conclusions. The hybridization with *P. australis* limited to six chromosomes could be explained as the presence of conserved rDNA sequence between the two species as the pattern is similar to that usually shown by ribosomal sequences. Therefore, for the time being, *P. australis* should probably be excluded as one of the direct ancestor of *A. donax*.

4.6 REFERENCES

Saltonstall K, Lambert A, Meyerson LA (2010) Genetics and reproduction of common (*Phragmites australis*) and giant reed (*Arundo donax*). *Invasive Plant Science and Management*, **3**: 495-505.

Hardion L, Verlaque R, Baumel A, Juin M, Vila B (2012) Revised systematics of Mediterranean *Arundo* (Poaceae) based on AFLP fingerprints and morphology *Taxon*, **61**, 1217–1226.

Hardion L, Verlaque R, Saltonstall K, Leriche A, Vila B (2014) Origin of the invasive *Arundo donax* (Poaceae): a trans-Asian expedition in herbaria. *Annals of Botany*, **114**, 455–462.

Hardion L, Verlaque R, Rosato M, Rosselló JA, Vila B (2015) Impact of polyploidy on fertility variation of Mediterranean *Arundo* L. (Poaceae). *Comptesrendus biologies*, **338**, 298-306.

Mariani C, Cabrini R, Danin A, Piffanelli P, Fricano A, Gomarasca S, Soave C (2010) Origin, diffusion and reproduction of the giant reed (Arundo donax L.): a promising weedy energy crop. Annals of Applied Biology, 157, 191–202.

Bucci A, Cassani E, Landoni M, Cantaluppi E, Pilu R (2013) Analysis of chromosome number and speculations on the origin of *Arundo donax* L. (Giant Reed). Cytology and Genetics, 47, 237-241.

Lambert, A. M., Dudley, T. L., & Saltonstall, K. (2010). Ecology and impacts of the large-statured invasive grasses Arundo donax and Phragmites australis in North America.

Hasterok, R., Jenkins, G., Langdon, T., Jones, R. N., & Maluszynska, J. (2001). Ribosomal DNA is an effective marker of *Brassica* chromosomes. Theoretical and Applied Genetics, 103(4), 486-490.

CHAPTER 5.
ACKNOLEDGEMENTS

Ringrazio il Prof.re Silvio Salvi, il mio tutor, grande scienziato e grande appassionato di scienza, che durante questi anni di dottorato mi ha permesso di crescere sia professionalmente sia umanamente.
Ringrazio il Prof.re Roberto Tuberosa, la guida del gruppo di genetica.
Ringrazio tutti i miei colleghi nonché amici che ho avuto la fortuna di incontrare in questi anni di dottorato, per il supporto e l'amicizia dimostratomi.
Ringrazio la mia famiglia: mia madre, Stefania, che mi ha supportato e sopportato durante tutta la mia carriera universitaria, credendo in me quando nemmeno io ero in grado di farlo; i miei fratelli Marco e Matteo, per tutto l'aiuto che hanno saputo darmi in ogni situazione. In particolare, ringrazio Marco, mio gemello, senza di lui non avrei molto probabilmente nemmeno valutato l'idea di iscrivermi all'università ed arrivare a conseguire il titolo in dottore di ricerca; mio padre, Andrea, venuto a mancare ormai 13 anni fa; a lui dedico questo mio risultato, vorrei tanto che fosse ancora qui per vedere la persona che sono diventato.
Ringrazio i miei amici, quelli più vecchi, che sono ancora lì, con loro ho condiviso i momenti più belli e divertenti della mia vita.
Infine, ringrazio Giulia, la mia ragazza, la persona che a più di ogni altro devo questo momento. Senza di lei, senza i suoi incoraggiamenti a suon di "scrivi", "lavora", "applicati", non sarei mai arrivato alla fine di questo mio percorso. In lei ho trovato la motivazione e la forza di volontà necessarie per affrontare le mie insicurezze e per non mollare, quando questo mi sembrava l'unica alternativa. Grazie davvero!