Alma Mater Studiorum – Università di Bologna
in collaborazione con Fondazione Edmund Mach

DOTTORATO DI RICERCA IN

Scienze e Tecnologie Agrarie, Ambientali e Alimentari

Ciclo 29

Settore Concorsuale di afferenza: 07/E1

Settore Scientifico disciplinare: AGR/07

TITOLO TESI

GENOMIC RESOURCES DEVELOPMENT AND FUNCTIONAL CHARACTERIZATION OF ARUNDO DONAX L. UNDER STRESS CONDITIONS

Presentata da: **Poli Michele**

Coordinatore Dottorato

Prof. Dinelli Giovanni

Relatore

Prof. Salvi Silvio

Relatore

Dr. Varotto Claudio

Esame finale anno 2017

Contents

Abstract

Global warming is becoming a major threat for human and wildlife on Earth and scientists agree in considering anthropogenic greenhouse gasses emission as the main factor contributing to atmosphere and sea temperature increase. *Arundo donax* L. is the most promising species for second generation biofuel production in Mediterranean areas but up to now little was known about its genetics. In this study, we first explored the whole transcriptome of young shoot of *Arundo donax* under simulated drought stress through NGS technology. This allowed us to understand the general molecular mechanisms of early plant responses to osmotic stimulus. Through comparative analyses with major Poaceae species, we identified a set of 53 orthologs that can be considered as a core of evolutionary conserved genes important to mediate water stress responses in the family. Leveraging on the availability of this transcriptome, we developed a set of of reliable reference genes with high stability across different stress and/or tissues, to enable further functional studies in this species. We selected a candidate, named *AdDWD1*, that have the potential to play an important role in stress response by targeted protein degradation. Our results indicate that *AdDWD1* is upregulated under osmotic and salt response in *A. donax* and its overexpression in *Arabidopsis* brought to a significant decrease in germination under salt and a growth retardation in ABA-containing media. Overexpression of *AdDWD1* caused downregulation of *DREB2A* and *SOS3*, while it did not affect other stress-related genes, pinpointing a possible pathway-specific regulatory role of the gene. Taken together, these results suggest a strong relation of *AdDWD1* with salt and osmotic stress response and an important role in the signalling pathway during early stress stages.

Keywords: Arundo donax, abiotic stress, salt, osmotic, RNa-seq, qRT-PCR, overexpression, reference gene, Arabidopsis, WD40, DWD, AdDWD1, fossil fuel, biomass, unigene, ABA, RefFinder, geNorm, NormFinder, cloning, expression pattern, transcriptome, Poaceae.

Chapter 1.

Overview and project background.

1.1 Introduction

Global warming is becoming a major treat for human and in general for every kind of life on Earth. Scientists agree in consider anthropogenic greenhouse gasses (GHG) emission as the main contribution to the atmosphere and sea temperature raising (Peck et al., 2016). Several political actions were adopted worldwide to reduce the impact of human activities on the environment and reduction of the equivalent CO_2 emission is one of the key parameters to measure the progresses on this issue. Recently, the conference held in Paris by United Nations Framework Convention on Climate Change (UNFCCC) ended up with an international agreement in which every country proposed the intention to keep the global average temperature below 2 °C above pre-industrial level and possibly below 1.5 °C. Even though the "Paris agreement" is the most important document on climate change until now, several countries have made effort to pursue limitation of GHG emission from the early 2000. Especially, European Union put much effort in climate policy with the so called 20-20-20 programme that aims to reduce of 20% the overall European CO_2 emission and increase of 20% both renewable energy and efficiency by 2020. This goals was recently further moved up to 40% of emission cut and 27% of renewable energy consumption and efficiency before 2030.

One branch of the CO_2 reduction policy focus on road transportation and especially aims to mix the fossil fuels with those coming from processing plant biomasses (e.g. bioethanol). At the beginning, thanks to advantageous incentives, the selected crops were chosen to give a high yield and easy to digest raw products (such as oil rich seeds) so that species like corn, sugar cane and soybean were used extensively for biofuel conversion (so called first generation biofuels), competing for good quality lands against food-crops (Naik et al., 2010). It was easily argued that the climate benefits (if any) could not justify these losses, however this step allowed us to further develop the technologies for crop conversion into fuels. Second generation biofuel refers largely to lignocellulosic materials such as food crop straw or dedicated non-food crops (Figure 1.1). Because

food crops have mainly been selected to produce more and bigger seeds/fruits, we can expect that straw amount would be relatively low and its usage could be only a side input of the total biomass (Doebley et al., 2006). Moreover, it has been demonstrated that straw removal worsens soil quality by depletion of organic carbon that will require the addition of fertilizer in a long-term plantation (Lal, 2009). Dedicated non-food crops are the most promising way for biofuel conversion even if some drawback still need to be solved. There are two main problems: first the biomass yield needs to be improved for an easier digestion without loss of quantity and second, in some cases they are still competitor for land with food-crop. In the last decades in Europe and US, dozens of species were evaluated for their yield, chemical composition and environmental requirement, finally selecting four species with high potential: miscanthus (*Miscanthus* spp.), reed canarygrass (*Phalaris arundinacea*), giant reed (*Arundo donax*) and switchgrass (*Panicum virgatum*) (Lewandowski et al., 2003). The selection of these plants in Europe were made also based on different climates that make more suitable one species respect to another: for example reed canary grass can be used in northern countries thanks to its C_3 photosynthetic pathway that improve biomass quantity and quality in cold climate; *miscanthus* and switchgrass are adapt to central Europe because they can reach high biomass yield with high nutrient and water usage efficiency (C_4 pathway) and high persistence; giant reed is suitable for southern region where it is already present and perform well also under drought stress (Lewandowski et al., 2003).

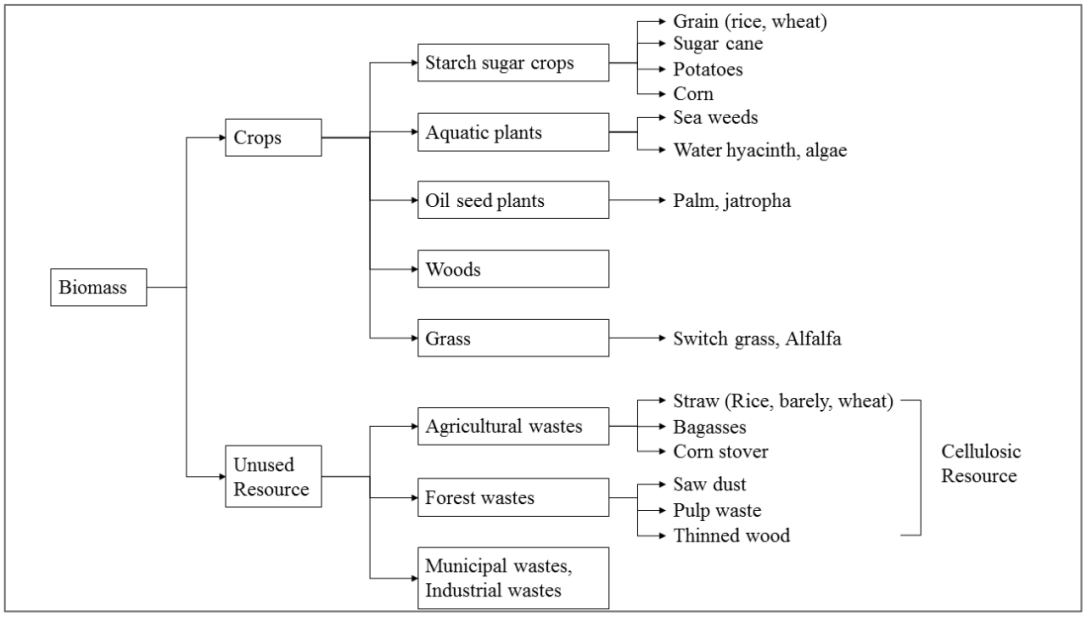


Figure 1.1. **Biomass as renewable feed stock for biorefineries**. From: Naik, S. N., Goud, V. V., Rout, P. K. & Dalai, A. K. Production of first and second generation biofuels: A comprehensive review. Renew. Sustain. Energy Rev. 14, 578–597 (2010).

1.2 *Arundo donax L.* as an important crop for biofuel production

In recent years, much effort was made to find the best biofuel crop. The characteristics of the ideal species are:

- High biomass production;
- High biomass quality (e.g. high lignin and low water and hash content in dry matter);
- Low input requirements (especially fertilization and water);
- Capability to cope with stress conditions and resistance to natural pests;
- Low cropping cost (e.g. minimum tillage) and easy to stock;
- Genetic variability for ecotype selection and improvement.

As expected these requirements are extremely difficult to find in one species so that the selected one will be a trade-off between benefits and drawbacks. Interestingly, *A. donax* meets most of the feature to be a good candidate as biofuel crop. First of all, it as a high production of biomass that can reach 40 to 50 t ha^{-1} yr^{-1} in favourable conditions. This is surprising because *A. donax* has C_3 photosynthetic pathway that is normally considered less productive in terms of biomass yield compared to C_4 species that can fix atmosphere carbon dioxide more efficiently, particularly under high temperature and light conditions (Byrt et al., 2011). In *A. donax*, experimental evidences showed that the photosynthetic pathways is not saturated by the light and in this way the growth can be continuous, having a photosynthetic rate of 37 µmol m^{-2} s^{-1} (Rossa et al., 1998). Thanks to this adaptation, giant reed can reach 8-10 meter height and its growing rate can be very high during specific climate conditions (up to 5 cm d^{-1}) (Perdue, 1958). The stems continuously grow for the whole vegetative season from a fibrous rhizome that lays close to the soil surface and from which roots arise and deeply penetrate the soil. The hollow culms (1-5 cm of diameter) are often branched from the second year and an elongated 30–70 cm long, 5-7 cm broad leaf insist on each node. The biomass consists in stems and leaves that are usually collected during fall or winter and its quality (water content, N, K, S and hash) greatly depends on the field management (e.g. abundant fertilization implies higher final nitrogen content) although, in general, *A. donax* shows higher water content compare to other species at harvesting time (Lewandowski et al., 2003) making the conversion process less efficient.

Arundo donax is highly resistant to drought, salt (Sánchez et al., 2015) and heavy metals (chromium, cadmium, arsenic, nickel) in soil (Kausar et al., 2012; Mirza et al., 2010; Papazoglou et al., 2005; Sagehashi et al., 2011) but it is sensitive to cold, especially freezing event in late spring during sprout emergence. Moreover, it can grow in every kind of soil, can be irrigated with wastewater or salty water (Sánchez et al., 2015) and only few pests are reported in literature: the

galling wasp (*Tetramesa romana*), the shoot-fly *Cryptonevra* sp., and an aphid, *Melanaphis donacis* (Minogue and Wright, 2016). These features make it a perfect candidate to be cultivated on marginal or contaminated lands, where food crop cannot be planted. Abiotic and biotic resistance is one of the key factors for crop selection because it allows to produce biomass without any or little maintenance costs (irrigation, pesticides) and with obvious environmental benefits.

The main hindrance to its extensive use is probably the lack of sexual reproduction. *A. donax* produces no viable seeds due to the failure of megaspore division and therefore propagation via vegetative reproduction (through fragmented stems or rhizomes) is required for field planting, making the first cropping year more expensive than other species (Mariani et al., 2010). More important, the incapability of producing viable seeds of the plant is a limiting factor on genetic selection and improvement of the species. The absence on sexual reproduction is also demonstrated by several studies that show a low if absent genetic intraspecific variability by genetic fingerprinting (see paragraph 1.4).

1.3 Problems related to extensive *A. donax* cultivation.

As the attention for giant reed increases, some issues arise for its use in open field cultivation. Two are the main concerns related to this plant: a) the elevated invasiveness and b) possible air pollution caused by extensive field. The first point is related to the ability of *A. donax* to spread quickly in suitable conditions by vegetative propagation that can occur by rhizome or stem fragmentation. Even though the plant did not originate from Europe, it is considered a naturalized species of the Mediterranean basin because its presence is testified since antiquity. In north America instead, it was intentionally introduced. A. donax reached California from the Mediterranean in the 1820's, in the Los Angeles area, where it was introduced as an erosion-control agent in drainage canals or thatching for roofs of sheds, barns, and other buildings (Bell, 1997). The most vulnerable ecosystem has been the riparian habitat along rivers where giant reed could spread easily during flood events and strongly competes with native plant species (e.g. willow, cottonwood), consequently reducing food and nesting sites for animals (Bell, 1997). Two different studies on the genetic variability of giant reed in US showed moderate (Khudamrongsawat et al., 2004) or low genetic diversity (Ahmad et al., 2008) in the collected samples using isozyme + RAPD and SRAP + TE markers respectively. These results are in accordance with those of related invasive clonal species such as *Alternanthera philoxeroides* (Li and Ye, 2006), *Eichhornia crassipes* (Li et al., 2006) and *Fallopia japonica* (Hollingsworth and Bailey, 2000). More recently though,

microsatellites brought the discovery of multiple introduction in north America of different clones of *A. donax* and suggested Spain as the most probable location from which the predominant clones was imported (Tarin et al., 2013). For managing its invasiveness, many methods were proposed, from classical herbicide to mechanical removal of rhizome but they imply high costs especially in heavily affected areas therefore making biological control the most suitable long-term control agent. On the other hand, an assessment of the risk of cultivating *A. donax* in open field was made in Australia on the basis of SA Weed Risk Management System (SAWRMS) resulting, as expected, in an elevated danger along riparian habitat and a very low risk in terrestrial fields (Virtue et al., 2010). This can lead to the conclusion that giant reed can be safely cultivated in specific areas even where it is not native.

Isoprene (2-methyl-1,3-butadiene) is a non-methane volatile organic compound (NMVOC) emitted by anthropogenic and natural sources and have an impact on the oxidative photochemistry of the tropospheric system (Monson, 2002). Legislative actions have reduced human NMVOC emission but isoprene is still increasing in some areas due to urban and suburban forests where it has consequences at local and a regional level (Monson et al., 2007). Isoprene oxidation in the low atmosphere can produce derived compounds such as ozone (O_3), organic nitrates and organic acids which are noxious to human health (Fuentes et al., 2000). Many plants can emit isoprene from leaves and this appears to have a protective role from abiotic stress, especially high temperature (Sharkey et al., 2008) and from reactive oxygen species (ROS). Not all the plants, though, emit isoprene but its distribution along taxonomy suggests independent origins with several episodes of loss and gain of function (Harley et al., 2004). Evidence in Arundineae tribe show similar results, with *Arundo donax* and *Phramites australis* being strong emitters, while *Molinia caerulea* and *Hakonechloa macra* are respectively low and no emitters (Ahrar et al., 2015). Worth to notice, a case study on the effect of a giant reed plantation on quality air demonstrated that the cultivation site can affect regional troposphere therefore suggesting to considering also air impact (together with other common factors like energy density, growth rate, fertilizer and water requirements and ecological impacts) for the field selection (Porter et al., 2012).

1.4 Arundinoideae classification and phylogenies.

Arundo donax L. belongs to the Poaceae family, one of the largest plant family worldwide comprising some of the most important food crop like maize, wheat, rice, barley, and millet but also bamboo used as building material and several species cultivated for staple forage. This family is

mostly dominant in harsh habitat like savannah and prairie, however Poaceaes are present almost everywhere and they are adapted to extraordinary different environmental conditions (from rain forest to dry desert to cold climate) making them one of the most evolutionary successful plant family. Giant reed belongs to Arundinoideae subfamily which counts about 40 species and to Arundineae tribe in which *Arundo* is present together with *Amphipogon* and *Monachather* genus. It is interesting to notice that, differently from most of the PACMAD (Panicoideae, Arundinoideae, Chloridoideae, Micrairoideae, Aristidoideae, and Danthonioideae) clade, this subfamily counts only C_3 photosynthesis species. The growing interest on this species enhanced the effort to resolve its origin in order to obtain more information also about possible ecotypes that meet the criteria of a biofuel crop. Although *Arundo donax* has been extensively cultivated from ancient time by human in Asia, southern Europe, north Africa and middle East, its origin is still not clear. The first extensive report of this plant was done by Perdue (1958) that clearly explained many of its aspects such as biology, morphology, ecology, distribution and usage. Also, the taxonomy review in Perdue's paper insert giant reed in Festucae tribe with other 5 species, named *A. pliniana* Turra (now *plinii*, native of Mediterranean countries), *A. formosana* Hack. (native of Formosa island) and *A. conspicua* Forst., *A. fulvida* J. Bueh. and *A. richardi* Endl. which are native of New Zealand. After Arundo taxa was classified in a Mediterranean flower encyclopaedia (Polunin and Huxley; 1987), Danin begun to deeply study the European species describing also a new one, named *A. hellenica* (Danin et al., 2002). Again, Danin (Danin, 2004), aware of the confusion in the Arundo genus tried to give a first clear classification in the Mediterranean region based on phenotypic observation finally dividing *A. plinii* in 3 species: *Arundo plinii* s. str. Turra, *Arundo mediterranea* Danin, *Arundo collina* Tenore (before called *A. hellenica*). As explained in his paper, more research and a wider collection of plants was still needed to resolve the *Arundo* genus phylogeny. In the following decade, due to the interest in giant reed several studies were accomplished to definitely resolve the taxa. Molecular analysis of intra- and interspecific variation in Arundo genus from Mediterranean region with AFLP (Amplified Fragment Length Polymorphism) and ISSR (Inter-Simple Sequence Repeats) supported the presence of a monophyletic origin of giant reed without hybridization event with other species (*A. plinii, A. formosana, A. micrantha*) and a marked difference with Asian species (Mariani et al., 2010). The same paper, elucidate clearly seed formation in three Arundo species demonstrating that in giant reed, unlike the other two fertile species, the megaspore development produces a proliferation of undifferentiated cells eventually ending in ovule collapse. Other similar studies were conducted on giant reed in US and Australia with different findings: a low variability in US (Ahmad et al., 2008; Khudamrongsawat et al., 2004) and a higher one in Australia probably due to multiple introductions or the use of different

molecular markers (Inter Simple Sequence Repeats, ISSR) (Haddadchi et al., 2013). In 2012, two studies set up a new and definitive (until now) Arundo nomenclature by reconsidering the name *A. micrantha* Lam. in substitution of *A. mauritanica* Desf. and *A. mediterranea* Danin (Hardion et al., 2012b) and systematics through morphology and AFLP (Amplified Fragment Length Polymorphism) markers that revealed a division of *A. plinii* s.l. in *A. plinii* s. str. (locus classicus in Bologna, Italy), *A. donaciformis* (present in south France and Liguria, Italy) and *A. micrantha* (Hardion et al., 2012a). Recently, the origin of giant reed was extensively studied by Hardion (Hardion et al., 2014) on a wide-range herbarium specimens collection in order to shed light on the evolutionary history of the plant. The results obtained by morphometric analysis and 5 plastid DNA intergenic spacers (trnT-trnL, trnCF-rpoB, psaA-ORF170, rbcL-psaI, trnS(GCU)-psbD) confirm a genetic uniformity of *A. donax* in Europe and move the origin to Middle East countries from where it was imported to the Mediterranean basin in antiquity (Figure 1.2).

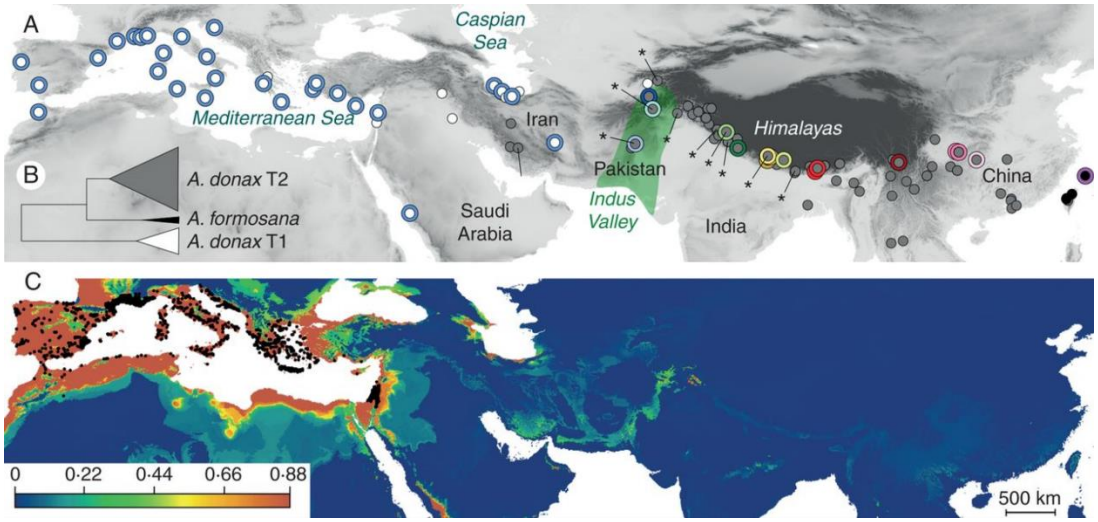


Figure 1.2. ***A. donax* origin**. (A) Geographical distribution of plastid DNA haplotypes and morphotypes. White circles, *A. donax* morphotype T1; grey circles, *A. donax* morphotype T2; black circles, *A. formosana* morphotype; *, seed occurrence. Coloured rings correspond to plastid DNA haplotypes (Fig. 1). (B) UPGMA tree based on morphological data. (C) Ecological niche modelling of *A. donax* calibrated on 1221 Mediterranean occurrences (black dots) and projected on sub-tropical Eurasia using MaxEnt. From: Origin of the invasive Arundo donax (Poaceae): a trans-Asian expedition in herbaria. Ann Bot. 2014;114(3):455-462. doi:10.1093/aob/mcu143.

1.5 Next Generation Sequencing for transcriptome analysis.

Thanks to the advent of new generation sequencing (NGS), it is now possible to obtain huge quantities of data faster and with much lower cost compare to the classic Sanger sequencing. These new technologies are also highly adaptable to different research studies, from whole genome or transcriptome sequencing to detection of microbial diversity in a given matter. It is also noteworthy

that some NGS setup can be applied to every biological form without any previous knowledge (da Fonseca et al., 2016). This is of an extraordinary importance because it allowed to exponentially increase the availability of genomic data in all the disciplines of biological field. Before the advent of NGS sequencing, efforts brought to publish the first genome of the model plant species *Arabidopsis thaliana* in 2000 ("The Arabidopsis Genome Initiative"). Later on, other economic important species like rice (2002), papaya (2008) and maize (2009) were sequenced, became a their own a model for related plants (Unamba et al., 2015). Even though it was important to have some model species, the biological features of all the others (some of which extremely important for food or energy production) were far to be discovered and this gap is now being filled thanks to these high-throughput technologies. Nevertheless, genome investigation by NGS has still some shortcomings that impose limits to its usage such as higher error rate (~0.1–15%) and shorter read length (35–700 bp), which make more difficult downstream steps such as data evaluation and analysis (Goodwin et al., 2016). Moreover, genomic sequence of non-model species is hindered by other factors like genome duplication, heterozygosity, ploidy level and repetitive sequence composition that can only be overcome by the use of multiple approaches (Unamba et al., 2015).

Gene identification and expression analysis have always been one of the key feature of molecular and evolutionary biology. To assess these tasks, the most widely used techniques were Northern blotting in which specific bands of RNA are hybridized with ^{32}P-labeled DNA to be subsequently detected on a photographic film (Alwine et al., 1977) and quantitative real-time PCR (qRT-PCR) which monitors the amplification of a specific target sequence with fluorescent technology (Valasek and Repa, 2005). Recently, qRT-PCR has become an invaluable tool for many scientist in gene expression analysis thanks to its high accuracy and simplicity. Its major drawback is the necessity of gene sequence information in order to design the PCR primers. In this context, RNA sequencing (RNA-seq) through NGS technologies brought a rapid generation of large expression datasets for gene discovery and expression analysis in non-model species (Marguerat and Bähler, 2010). Unlike microarray technique that requires a reference genome, RNA-seq can detect the whole transcriptome in a precise moment and/or tissue obtaining a *de novo* assembly. In recent year, also crops for energy production received great attention and transcriptome data from many species are now available, for example *Sorghum bicolor* (Dugas et al., 2011),*Camelina sativa* (Kagale et al., 2016), , sugarcane (Cardoso-Silva et al., 2014), *miscanthus* (Swaminathan et al., 2012), reed canary grass (Baillie et al., 2017). In 2014, the first whole transcriptome of four tissues (leaf, culm, bud and root) of *A. donax* was released giving a boost to the knowledge of this plant and possibly to its improvement for biofuel production (Sablok et al., 2014). Thanks to this effort, it was possible to identify several genes related to lignin, cellulose, starch, lipid metabolism but also

others related to abiotic stresses that are interesting targets for further characterization. The amount of data retrieved from this and other studies will be essential for new gene discoveries opening the way to thousands of possibilities for breeders and genetic engineers to improve plant species.

This four-year work aimed to further improve our knowledge about *Arundo donax* species, especially the aspects related to abiotic stress response. The mechanisms behind the capability of giant reed to cope with long drought period were the specific objectives of this thesis. We have first carried out experiments that dig into the molecular mechanisms undergoing the response to osmotic stress through the analysis of the whole transcriptome and then we have functionally explored one specific gene that has never been characterized even in other model species such as Arabidopsis or rice.

1.6 References

Ahmad, R., Liow, P.-S., Spencer, D.F., Jasieniuk, M., 2008. Molecular evidence for a single genetic clone of invasive Arundo donax in the United States. Aquat. Bot. 88, 113–120. doi:10.1016/j.aquabot.2007.08.015

Ahrar, M., Doneva, D., Koleva, D., Romano, A., Rodeghiero, M., Tsonev, T., Biasioli, F., Stefanova, M., Peeva, V., Wohlfahrt, G., Loreto, F., Varotto, C., Velikova, V., 2015. Isoprene emission in the monocot Arundineae tribe in relation to functional and structural organization of the photosynthetic apparatus. Environ. Exp. Bot. 119, 87–95. doi:10.1016/j.envexpbot.2015.04.010

Alwine, J.C., Kemp, D.J., Stark, G.R., 1977. Method for detection of specific RNAs in agarose gels by transfer to diazobenzyloxymethyl-paper and hybridization with DNA probes. Proc. Natl. Acad. Sci. U. S. A. 74, 5350–4. doi:10.1073/pnas.74.12.5350

Baillie, R., Drayton, M., Pembleton, L., Kaur, S., Culvenor, R., Smith, K., Spangenberg, G., Forster, J., Cogan, N., 2017. Generation and Characterisation of a Reference Transcriptome for Phalaris (Phalaris aquatica L.). Agronomy 7, 14. doi:10.3390/agronomy7010014

Bell, G.P., 1997. Ecology and Management of Arundo donax, and Approaches to Riparian Habitat Restoration in Southern California., in: Brock, J. (Ed.), Plant Invasions Studies from North America and Europe. Backhuys, pp. 103–113.

Byrt, C.S., Grof, C.P.L., Furbank, R.T., 2011. C4 plants as biofuel feedstocks: optimising biomass production and feedstock quality from a lignocellulosic perspective. J. Integr. Plant Biol. 53, 120–35. doi:10.1111/j.1744-7909.2010.01023.x

Cardoso-Silva, C.B., Costa, E.A., Mancini, M.C., Balsalobre, T.W.A., Costa Canesin, L.E., Pinto, L.R., Carneiro, M.S., Garcia, A.A.F., De Souza, A.P., Vicentini, R., 2014. De novo assembly and transcriptome analysis of contrasting sugarcane varieties. PLoS One 9. doi:10.1371/journal.pone.0088462

da Fonseca, R.R., Albrechtsen, A., Themudo, G.E., Ramos-Madrigal, J., Sibbesen, J.A., Maretty, L., Zepeda-Mendoza, M.L., Campos, P.F., Heller, R., Pereira, R.J., 2016. Next-generation biology: Sequencing and data analysis approaches for non-model organisms. Mar. Genomics 30, 1–11. doi:10.1016/j.margen.2016.04.012

Danin, A., 2004. Arundo (Gramineae) in the Mediterranean reconsidered 34204, 361–369. doi:10.3372/wi.34.34204

Danin, A., Raus, T., Scholz, H., 2002. Contribution to the Flora of Greece: A New Species of Arundo (Poaceae). Willdenowia 32, 191–194.

Doebley, J., Gaut, B., Smith, B., 2006. The molecular genetics of crop domestication. Cell 1309–1321.

Dugas, D. V, Monaco, M.K., Olsen, A., Klein, R.R., Kumari, S., Ware, D., Klein, P.E., 2011. Functional annotation of the transcriptome of Sorghum bicolor in response to osmotic stress and abscisic acid. BMC Genomics 12, 514. doi:10.1186/1471-2164-12-514

Fuentes, J.D., Lerdau, M., Atkinson, R., Baldocchi, D., Bottenheim, J.W., Ciccioli, P., Lamb, B., Geron, C., Gu, L., Guenther, A., Sharkey, T.D., Stockwell, W., 2000. Biogenic Hydrocarbons in the Atmospheric Boundary Layer: A Review. Bull. Am. Meteorol. Soc. 81, 1537–1575. doi:10.1175/1520-0477(2000)081<1537:BHITAB>2.3.CO;2

Goodwin, S., McPherson, J.D., McCombie, W.R., 2016. Coming of age: ten years of next-generation sequencing technologies. Nat Rev Genet 17, 333–351. doi:10.1038/nrg.2016.49

Haddadchi, A., Gross, C.L., Fatemi, M., 2013. The expansion of sterile Arundo donax (Poaceae) in southeastern Australia is accompanied by genotypic variation. Aquat. Bot. 104, 153–161. doi:10.1016/j.aquabot.2012.07.006

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

Hardion, L., Verlaque, R., Callmander, M.M.W., Vila, B., 2012b. Arundo micrantha Lam.(Poaceae), the correct name for Arundo mauritanica Desf. and Arundo mediterranea Danin. Candollea 67, 131–135.

Hardion, L., Verlaque, R., Saltonstall, K., Leriche, A., Vila, B., 2014. Origin of the invasive Arundo donax (Poaceae): a trans-Asian expedition in herbaria. Ann. Bot. mcu143-. doi:10.1093/aob/mcu143

Harley, P., Vasconcellos, P., Vierling, L., Pinheiro, C.C.D.S., Greenberg, J., Guenther, A., Klinger, L., De Almeida, S.S., Neill, D., Baker, T., Phillips, O., Malhi, Y., 2004. Variation in potential for isoprene emissions among Neotropical forest sites. Glob. Chang. Biol. 10, 630–650. doi:10.1111/j.1529-8817.2003.00760.x

Hollingsworth, M.L., Bailey, J.P., 2000. Evidence for massive clonal growth in the invasive weed Fallopia japonica (Japanese Knotweed). Bot. J. Linn. Soc. 133, 463–472. doi:10.1006/bojl.2000.0359

Kagale, S., Nixon, J., Khedikar, Y., Pasha, A., Provart, N.J., Clarke, W.E., Bollina, V., Robinson, S.J., Coutu, C., Hegedus, D.D., Sharpe, A.G., Parkin, I.A.P., 2016. The developmental transcriptome atlas of the biofuel crop Camelina sativa. Plant J. 879–894. doi:10.1111/tpj.13302

Kausar, S., Mahmood, Q., Raja, I.A., Khan, A., Sultan, S., Gilani, M.A., Shujaat, S., 2012. Potential of Arundo donax to treat chromium contamination. Ecol. Eng. 42, 256–259. doi:10.1016/j.ecoleng.2012.02.019

Khudamrongsawat, J., Tayyar, R., Holt, J.S., 2004. Genetic diversity of giant reed (Arundo donax) in the Santa Ana River, California. Weed Sci. doi:10.1614/WS-03-120R1

Lal, R., 2009. Soil quality impacts of residue removal for bioethanol production. Soil Tillage Res. 102, 233–241. doi:10.1016/j.still.2008.07.003

Lewandowski, I., Scurlock, J.M.O.O., Lindvall, E., Christou, M., 2003. The development and current status of perennial rhizomatous grasses as energy crops in the US and Europe. Biomass and Bioenergy 25, 335–361. doi:10.1016/S0961-9534(03)00030-8

Li, J., Ye, W.H., 2006. Genetic diversity of alligator weed ecotypes is not the reason for their different responses to biological control. Aquat. Bot. 85, 155–158. doi:10.1016/j.aquabot.2006.02.006

Li, W., Wang, B., Wang, J., 2006. Lack of genetic variation of an invasive clonal plant Eichhornia crassipes in China revealed by RAPD and ISSR markers. Aquat. Bot. 84, 176–180. doi:10.1016/j.aquabot.2005.09.008

Marguerat, S., Bähler, J., 2010. RNA-seq: from technology to biology. Cell. Mol. Life Sci. 67, 569–79. doi:10.1007/s00018-009-0180-6

Mariani, C., Cabrini, R., Danin, a., Piffanelli, P., Fricano, a., Gomarasca, S., Dicandilo, M., Grassi, F., Soave, C., 2010. Origin, diffusion and reproduction of the giant reed (Arundo donax L.): a promising weedy energy crop. Ann. Appl. Biol. 157, 191–202. doi:10.1111/j.1744-7348.2010.00419.x

Minogue, P., Wright, S., 2016. Biology , Control and Invasive Potential of Giant Reed 2–6.

Mirza, N., Mahmood, Q., Pervez, A., Ahmad, R., Farooq, R., Shah, M.M., Azim, M.R., 2010. Phytoremediation potential of Arundo donax in arsenic-contaminated synthetic wastewater. Bioresour. Technol. 101, 5815–9. doi:10.1016/j.biortech.2010.03.012

Monson, R.K., 2002. Volatile organic compound emissions from terrestrial ecosystems: A primary biological control over atmospheric chemistry. Isr. J. Chem. 42, 29–42. doi:10.1560/0jjc-xqaa-jx0g-fxjg

Monson, R.K., Trahan, N., Rosenstiel, T.N., Veres, P., Moore, D., Wilkinson, M., Norby, R.J., Volder, A., Tjoelker, M.G., Briske, D.D., Karnosky, D.F., Fall, R., 2007. Isoprene emission from terrestrial ecosystems in response to global change: minding the gap between models and observations. Philos. Trans. A. Math. Phys. Eng. Sci. 365, 1677–95. doi:10.1098/rsta.2007.2038

Naik, S.N., Goud, V. V., Rout, P.K., Dalai, A.K., 2010. Production of first and second generation biofuels: A

comprehensive review. Renew. Sustain. Energy Rev. 14, 578–597. doi:10.1016/j.rser.2009.10.003

Papazoglou, E.G., Karantounias, G.A., Vemmos, S.N., Bouranis, D.L., 2005. Photosynthesis and growth responses of giant reed (Arundo donax L.) to the heavy metals Cd and Ni. Environ. Int. 31, 243–9. doi:10.1016/j.envint.2004.09.022

Peck, S.C., Teisberg, T.J., Peck, S.C., Teisberg, T.J., 2016. CETA : A Model for Carbon Emissions Trajectory Assessment Published by : International Association for Energy Economics Stable URL : http://www.jstor.org/stable/41322454 Accessed : 15-04-2016 05 : 36 UTC CETA : A Model for Carbon Emissions Trajectory Asse 13, 55–77.

Perdue, R.E., 1958. Arundo donax—Source of musical reeds and industrial cellulose. Econ. Bot. 12, 368–404. doi:10.1007/BF02860024

Porter, W.C., Barsanti, K.C., Baughman, E.C., Rosenstiel, T.N., 2012. Considering the air quality impacts of bioenergy crop production: a case study involving Arundo donax. Environ. Sci. Technol. 46, 9777–84. doi:10.1021/es3013084

Polunin, O. and Huxley, A.J. (1965) ARUNDO. In Flowers of the Mediterranean, pp. 196, 199. London: Chatto and Windus.

Rossa, B., Tuffers, a V, Naidoo, G., von Willert, D.J., 1998. Arundo donax L. (Poaceae) - a C3 species with unusually high photosynthetic capacity. Bot. Acta 111, 216–221. doi:10.1111/j.1438-8677.1998.tb00698.x

Sablok, G., Fu, Y., Bobbio, V., Laura, M., Rotino, G.L., Bagnaresi, P., Allavena, A., Velikova, V., Viola, R., Loreto, F., Li, M., Varotto, C., 2014. Fuelling genetic and metabolic exploration of C 3 bioenergy crops through the first reference transcriptome of Arundo donax L. Plant Biotechnol. J. n/a-n/a. doi:10.1111/pbi.12159

Sagehashi, M., Liu, C., Fujii, T., Fujita, H., Sakai, Y., Hu, H.-Y., Sakoda, A., 2011. Cadmium Removal by the Hydroponic Culture of Giant Reed (Arundo donax) and Its Concentration in the Plant. J. Water Environ. Technol. 9, 121–127. doi:10.2965/jwet.2011.121

Sánchez, E., Scordia, D., Lino, G., Arias, C., Cosentino, S.L., Nogués, S., 2015. Salinity and Water Stress Effects on Biomass Production in Different Arundo donax L. Clones. BioEnergy Res. doi:10.1007/s12155-015-9652-8

Sharkey, T.D., Wiberley, A.E., Donohue, A.R., 2008. Isoprene emission from plants: why and how. Ann. Bot. 101, 5–18. doi:10.1093/aob/mcm240

Swaminathan, K., Chae, W.B., Mitros, T., Varala, K., Xie, L., Barling, A., Glowacka, K., Hall, M., Jezowski, S., Ming, R., Hudson, M., Juvik, J.A., Rokhsar, D.S., Moose, S.P., 2012. A framework genetic map for Miscanthus sinensis from RNAseq-based markers shows recent tetraploidy. BMC Genomics 13, 142. doi:10.1186/1471-2164-13-142

Tarin, D., Pepper, A.E., Goolsby, J.A., Moran, P.J., Arquieta, A.C., Kirk, A.E., Manhart, J.R., 2013. Microsatellites Uncover Multiple Introductions of Clonal Giant Reed (Arundo donax). Invasive Plant Sci. Manag. 6, 328–338. doi:10.1614/IPSM-D-12-00085.1

Unamba, C.I.N., Nag, A., Sharma, R.K., 2015. Next Generation Sequencing Technologies: The Doorway to the Unexplored Genomics of Non-Model Plants. Front. Plant Sci. 6, 1074. doi:10.3389/fpls.2015.01074

Valasek, M.A., Repa, J.J., 2005. The power of real-time PCR. AJP Adv. Physiol. Educ. 29, 151–159. doi:10.1152/advan.00019.2005.

Virtue, J., Reynolds, T., Malone, J., 2010. Managing the weed risk of cultivated Arundo donax L., in: Seventeenth Australasian Weeds Conference.

Chapter 2.

Arundo donax L. transcriptome under osmotic stress.

Dissection of early transcriptional responses to water stress in *Arundo donax* L. by unigene-based RNA-Seq

Yuan Yuan Fu[1,2]*, Michele Poli[1,3]*, Gaurav Sablok[1], Bo Wang[1,4], Yanchun Liang[5], Nicola La Porta[1], Violeta Velikova[1,6], Francesco Loreto[7], Mingai Li[1], Claudio Varotto[1§]

Affiliation:

[1] Department of Biodiversity and Molecular Ecology, Research and Innovation Centre, Fondazione Edmund Mach, Via E. Mach 1, 38010 S. Michele all'Adige (TN), Italy

[2] Dipartimento di Biotecnologie, Università degli Studi di Verona, Verona, Italy

[3] Dipartimento di Scienze Agrarie, Università di Bologna, Bologna, Italy

[4] Centro di Biologia Integrata (CIBIO), University of Trento, Trento, Italy

[5] College of Computer Science and Technology, Jilin University, Changchun, China

[6] Institute of Plant Physiology and Genetics, Bulgarian Academy of Sciences, Bulgaria

[7] The National Research Council of Italy (CNR), Department of Biology, Agriculture and Food Sciences, Rome, Italy

2.1 Introduction

Among the different sources of renewable energy, biomass is interesting because it has a nearly neutral carbon balance and the ethanol produced by its fermentation can be blended with petrol-derived fuels giving an important contribution in reducing transport-related CO2 emissions (Acres, 2007). So-called second generation bioethanol (i.e., the one not produced from edible parts of crops) can be obtained from food crop straw, but the yields of such biomass source are expected

to be low, as food crops were intentionally selected to maximize photosynthate allocation to edible parts (Doebley, Gaut, & Smith, 2006). In alternative, plant species specifically dedicated to energy production (called bioenergy crops) are normally better biomass producers than food crops, resulting in higher ethanol yields per unit of cultivated area. Arundo donax has been identified among bioenergy crops as the most promising species for the Mediterranean area (Lewandowski et al. 2003). *A. donax,* commonly called giant reed, is a perennial C3, polyploid, bamboo-like grass of the Poaceae family. It favors well-drained soils with abundant moisture, where it can form dense stands up to 6–10 m high with yields of up to 40 tons per hectare each year (comparable, or even exceeding, those of some C4 species) (Byrt et al, 2011). The origins of the giant reed are still debated, but the latest evidences from plastid DNA sequencing and morphometric parameters data collected from 127 herbarium specimens support a Middle-East origin of *A. donax* (Hardion et al, 2014). Despite the production of panicle-like flowers, no viable seeds from Mediterranean ecotypes have been reported so far (Balogh et al, 2012). Natural propagation exclusively occurs vegetatively by rooting of rhizome and stem fragments originating as a consequence of flooding, followed by a slow colonization through rhizome expansion (Di Tomaso et al, 2003). Consistently, genetic diversity in *A. donax* has been reported to be low, but, possibly due to somatic mutation, detectable (Haddadchi et al, 2013). Possibly because of its high ploidy, the low intraspecific diversity of *A. donax* does not seem to be associated to fitness tradeoffs, as indicated by its high resistance to biotic and abiotic stresses (Mariani et al, 2010). If on one hand this resistance causes the high invasiveness of this plant, on the other hand it makes *A. donax* an excellent bioenergy crop, which can grow with very low management input (e.g., pesticides, fertilization, irrigation) even in marginal lands or in fields irrigated with waste or salty water (Mavrogianopoulos et al, 2002).

The recent advent of Next Generation Sequencing (NGS) has made the development of genomic resources progressively simpler and cheaper (Liu et al. 2012). RNA sequencing (RNA-Seq) is to date by far the most powerful tool for the rapid and inexpensive development of genetic resources for any species of interest. In addition, RNA-Seq allows at once the quantitative determination of the expression levels of virtually all transcribed genes in a specific organ, thus providing an extremely powerful tool for the identification of transcripts differentially expressed in response to the abiotic and biotic stresses which negatively impact crop growth and productivity (Martin, Fei, Giovannoni, & Rose, 2013).

It is widely accepted that global warming will increase the duration and frequency of drought periods over the 21th century (Dai 2012). Many countries already started to develop mitigation strategies to avoid this major threat, which could potentially offset the productivity gains

expected from advances in both agricultural and crop breeding techniques (Shanker et al. 2014). Drought is one of the extreme environmental conditions that curtail agricultural crop productivity (Bruce, Edmeades, and Barker 2002). The first response of plants to water limitation is usually avoidance, a strategy that aims at maintaining a neutral balance between water gained from the root system and lost by transpiration through the stomata. In case of short-term or relatively mild water stress, avoidance can maintain performance and prevent negative effects on plant growth. From a physiological point of view, this is usually achieved by increasing the osmotic potential of root cells, increasing root growth as well as reducing water loss by modulation of stomatal conductance (Verslues et al, 2003). These physiological adjustments are the consequence of complex cellular changes like: (1) the reprogramming of the cellular metabolism, which shifts to polysaccharide degradation and aminoacid biosynthesis to allow for the accumulation of solutes with an osmotic function (e.g., glycine–betaine, proline, mannitol, etc.), (2) the production of abscisic acid (ABA, a phytohormone mainly associated to seed dormancy and water stress, which causes a reduction of stomatal conductance through closure of stomata) and other phytohormones, and (3) an increased synthesis of proteins for cellular protection/detoxification (late-embryogenesis-abundant, LEA; chaperones and heat stress-proteins necessary for proper protein folding), (4) extensive modulation of ribosomal activity to support active cell growth and division in the root system (Verslues et al, 2003 and Claeys and Inzé 2013). When avoidance strategies are not sufficient alone to prevent the onset of water stress, either because of the excessive length or magnitude of the water deficit, tolerance responses become progressively more relevant to limit the damages caused by the reduced availability of water. The same physiological and molecular changes are, however, often shared between the two types of responses, so that a clear-cut distinction between them is not always possible. The medium to long-term adjustments associated to tolerance encompass the development of, e.g., thicker epicuticular waxes to limit water evaporation through epidermal cells, the further decrease of the shoot/root biomass ratio and the allocation of resources to long-term survival organs (e.g., tubers or rhizomes), the enhancement of antioxidant capacity to detoxify the reactive oxygen species (ROS) consequent to photosynthetic limitation, the thickening of xylematic cell walls to prevent collapsing of vasculature, etc. (Verslues et al, 2003 and Claeys and Inzé 2013).

This complex series of cellular responses to water limitation obviously requires also a profound reprogramming of gene expression. Our understanding of the genetic bases of drought resistance largely benefitted from forward and genetic screens in model or crop species (e.g. *Arabidopsis thaliana*, rice, maize, wheat; reviewed by Claeys and Inzé 2013). In addition, several studies devoted to the dissection of the transcriptional responses to drought stress or water deficit conditions have been carried out for the most common cereal crop species (e.g. rice, Wang et al.

2011; maize, Kakumanu et al. 2012; foxtail millet, Qi et al. 2013; sorghum, Dugas et al. 2011). More recently also Poaceae species used exclusively or partly as energy crops (switchgrass, Xie et al. 2013; miscanthus Lewandowski et al., 2000; sugarcane, Kido et al. 2012) have been object of transcriptomics studies which could provide a robust comparative basis in poorly characterized species like *A. donax*. Unfortunately, the high spatio-temporal complexity of the physiological adaptations to drought and the large number of variables used in different experimental protocols for the application of water stress (methods for induction of water deprivation, combination with other stresses, length of treatment, type of plant materials and their developmental stages) limit the depth of result comparisons across studies (Deyholos, 2010). Polyethylene glycol (PEG) is a high-molecular weight polymer which can be used to induce controlled water deficits in plants by modifying the osmotic potential of water in hydroponic growth media without being absorbed by the root system (Lagerwerff et al., 1961), thus providing an ideal method for water deprivation in RNA-Seq experiments addressing short-term responses of plants to water stress.

Arundo donax is one of the most promising biomass resources for biofuel development but, up to now, little is known at the molecular level on this species' ability to cope with abiotic stresses in general and in particular with water limitation. Leveraging on the recent obtainment of the first reference transcriptome of *A. donax* by RNA-Seq (Sablok et al. 2014) and on the existing knowledge of the genetics of drought responses in plants, in this study we report the characterization of early transcriptional responses to two levels of PEG-induced water deficit in cohorts of young giant reed cuttings. In particular, we addressed the main questions: (1) How many/which genes are differentially expressed during the early phases of water stress in *A. donax*? (2) What are the main biological functions involved? (3) Which are the transcription factors associated to such transcriptional reprogramming? (4) Are the transcriptional responses of *A. donax* conserved/comparable to those of other monocot species, and in particular of rice? The set of about 3000 early-responsive genes to water stress identified in this study are promising reporters of the physiological status of *A. donax* plantations for the improvement of its management and for a deeper understanding of its biology.

2.2 Results and Discussion

Despite the ability of *A. donax* to withstand prolonged periods of drought, its productivity under water limitation is negatively affected (Lewandowski, Scurlock, Lindvall, & Christou, 2003).

Especially during the first year of establishment, *A. donax* growth can be severely retarded and plants damaged for lack of soil moisture (Perdue 1958). An in-depth understanding of the mechanisms involved in water limitation responses in this species is, therefore, an important pre-requisite to improve its management, but till now no characterization of the transcriptional variations of *A. donax* tissues associated to water deprivation is available. To fill this gap of knowledge, we carried out by RNA-Seq a comprehensive identification of early transcriptional responses of shoots and roots to two different levels of PEG-induced water limitation in *A. donax* (details of the experimental design can be found in Materials and Methods and in Supplementary File 2.1).

Following assembly, we obtained 111,749 transcripts covering 45,821 components. Given the high ploidy of *A. donax*, we chose to use a relatively high Kmer coverage during assembly (min-kmer_cov = 5) to minimize the formation of transcripts with retained introns (Gruenheit, et al. 2012). The observed N50 of the assembled transcriptome is 1826 bp, in line with our previous N50 reports (Sablok et al. 2014), indicating that a good coverage of the transcriptome has been achieved. To eliminate redundant transcripts, we further clustered the transcripts using the CD-HIT software resulting in a total of 80,962 transcripts. The non-redundant transcript set was further assembled into unigenes with MIRA to remove spurious transcripts, resulting in a final set of 80,335 unigenes with an N50 of 1570 bp. Summary statistics results for transcriptome assembly are provided in (Table 2.1).

Assembly *	**Summary Statistics**
Total trinity transcripts:	111749
Total trinity components:	45821
Contig N50	1826
MIRA Unigenes	80355
Total length of sequence:	75960964 bp
N50 stats:	1570 bp
GC %:	47.70%

Table 2.1. **Summary statistics of the sequencing reads and the corresponding assemblies.** Summary statistics of A. donax whole drought transcriptome obtained by Trinity and MIRA. *Trinity assembly: K 25, Kmer coverage 5

Identification of differentially expressed genes (DEGs) by RNA-Seq

As a first step in the characterization of *A. donax* transcriptional responses to water stress, we carried out the identification of the unigenes whose expression level significantly changed upon PEG-treatments. A total of 3034 genes showed differential expression in at least one of the two stress conditions (mild water stress vs. control, severe stress vs. control and severe vs. mild stress), with roughly the same number of genes being differentially expressed in shoots and roots (1684 and 1712 DEGs, respectively). Validation of expression levels for ten selected DEG candidates was carried out by real-time qRT-PCR, (Supplementary File 2.2). The high congruence between RNA-Seq and real time PCR results (coefficient of determination R^2= 0.94), indicates the reliability of RNA-Seq quantification of gene expression. Therefore, the selected genes could also constitute useful markers of early water deficit in *A. donax.* DEG identified in biological replicates clustered together in both organs, indicating good reproducibility of treatments. In addition, the heat maps qualitatively indicated the closer similarity of control and mild water stress between each other as compared to severe water stress (Supplementary File 2.3).

A detailed assessment of the number and the identity of the DEGs between conditions for each organ confirmed this observation: in shoots, only 98 genes were differentially expressed between control condition and mild water stress, *versus* 1572 between control and severe water stress, and 831 between mild and severe water stress. A similar trend, but less marked, characterized also root DEGs (Figure 2.1a), indicating the successful induction of varying degrees of water stress as a function of PEG concentration (Lagerwerff et al., 1961).

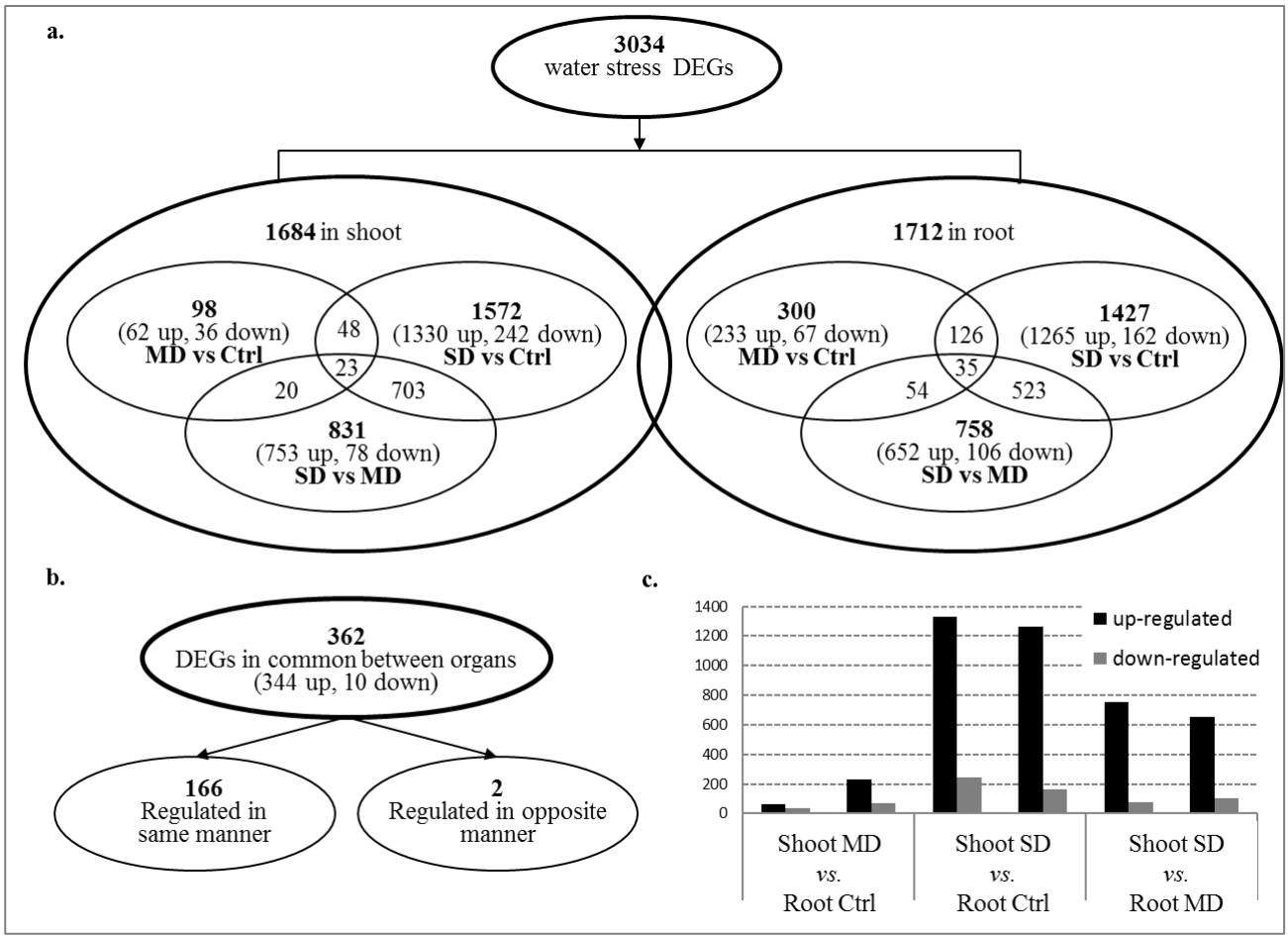


Figure 2.1. **Summary of DEGs in shoots and roots of A. donax upon drought stress**. a) Number of genes up-/down-regulated by drought stress under different conditions (MC: mild water stress vs control, SC: severe water stress vs control; SM: severe water stress vs mild water stress.) in root and shoot. b. Total number of DEGs in common between root and shoot. c. Number of regulated genes between different conditions. Grey bar: down-regulated genes; black bar: up-regulated genes.

By comparing the 362 DEGs in common between organs, we further observed a general conservation of expression patterns, with 166 of the genes being regulated in the same way in shoot and root and only two genes displaying opposite regulation (Figure 2.1b). In addition, the overall direction of expression variation resulted to be conserved between organs, with the large majority of DEGs being up- rather than down-regulated (Figure 2.1c). A closer analysis of the absolute numbers of DEGs in the two organs, however, highlighted a relatively large difference in gene up-regulation upon mild stress in roots as compared to shoots (300 DEGs in root *vs* 98 in shoot; Figure 2.1a). Given the application of the PEG directly to the root system and the sampling of only one time point, it is possible that, at least in part, these differences could stem from a faster onset of the water stress in roots compared to shoots. These results, however, are also in line with a transcriptionally higher responsivity of the root system compared to shoots, as previously reported, e.g., in the case of poplar (Cohen, 2010), which could indicate tissue-specific responses.

Functional classification of transcriptional responses to water stress in *A. donax*

To identify organ-specific differences, stress-related genes were identified based on curated homology searches against genes experimentally characterized in previous studies. The majority of stress-related genes belonged to categories "salt," "oxidative," "dehydration," and "osmotic." This is expected, as water limitation is known to cause reduced turgor and integrity of membranes, increase of intracellular ionic and non-ionic solute concentrations and enhanced production of reactive oxygen species (ROS) that cross-trigger responses to high-salinity, oxidative and osmotic stresses (Huang 2012) (Figure 2.2; Supplementary File 2.4). Worth of note, the two differentially expressed categories encompassing the largest differences in number of genes between organs are "dehydration" and "osmotic". Both categories are more abundant in shoot than root, but the highest shoot/root ratio (eight times) is found for dehydration-related genes (Figure 2.2).

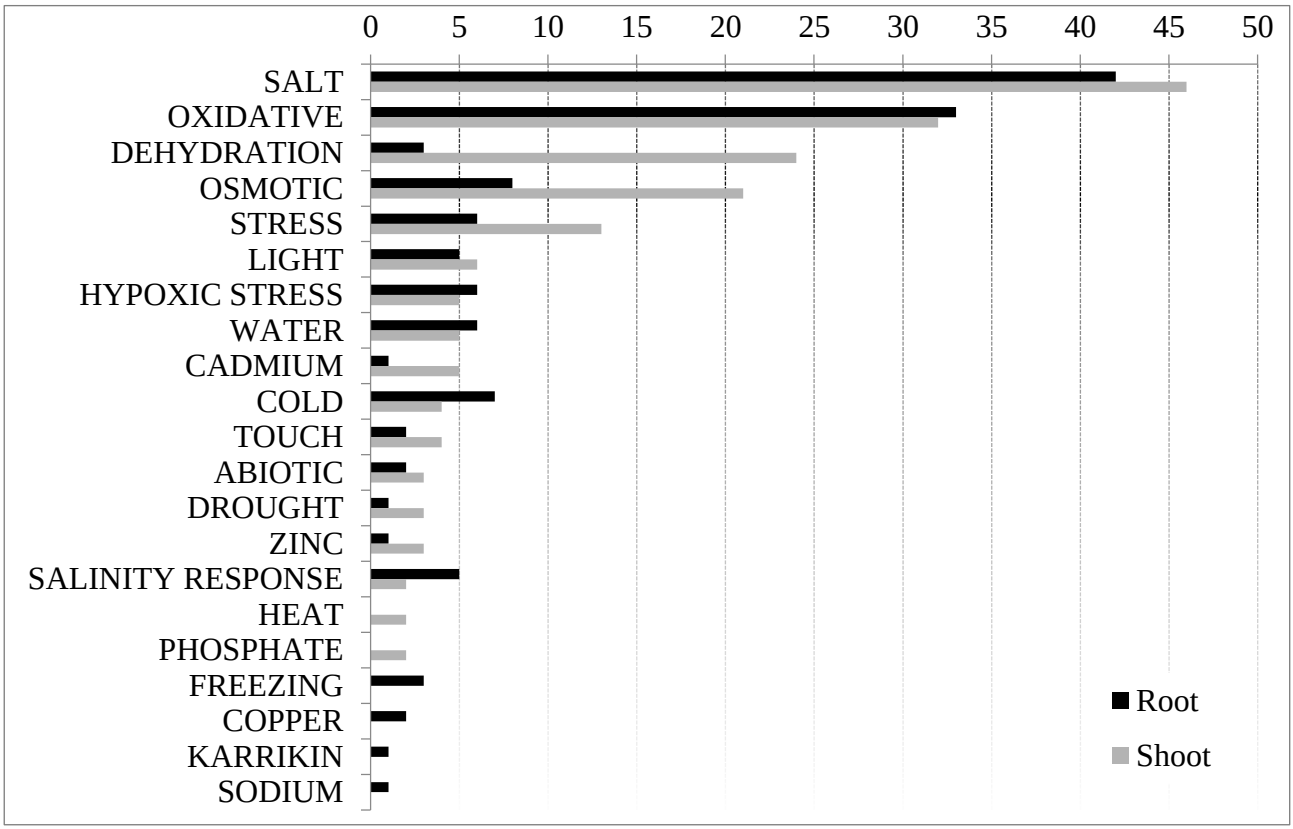


Figure 2.2. **Distribution of stress-related functional categories of DEGs**. Stress-related functional categories are identified by annotation of A. donax putative homologs in of Arabidopsis genes from ASPRGDB. Data are sorted by number of shoot DEGs. Black bar: root DEGs; grey bar: shoot DEGs.

We next carried out a homology-based annotation specifically for all the 3034 DEGs identified upon PEG treatments, identifying at the same time the GO terms associated to this dataset. (Supplementary File 2.5). Based on BLASTN searches using a 70 % identity and 50 %

query coverage cutoff, we identified a total of 214 differentially expressed *A. donax* unigenes in shoots and 642 in roots, which were not present in the reference transcriptome assembly (Supplementary 2.4) (Sablok et al, 2014), thus contributing to ongoing gene discovery and functional annotation in this poorly characterized species.

To determine the gene functional classes which were chiefly involved in the response to water stress, we carried out an analysis of over/under-representation of GO terms associated to DEGs. A total of three contrasts were carried out: mild water stress vs. control (MC), severe water stress vs. control (SC) and severe water stress vs. mild water stress (SM). We further selected the most significantly enriched GOs using REVIGO (Supek et al, 2011) (Supplementary 2.6) and analyzed the number of GO terms in common between contrasts to pinpoint differences and similarities between organs and conditions (Fig. 3: shoot/SC; root/SC; root/SM for both (1) molecular function and (2) biological process terms).

The seven molecular function and nine biological process GOs consistently over-represented in SC shoot and root contrasts, respectively (see, e.g., GO:0003700 sequence-specific DNA binding transcription factor activity, GO:0030528 transcription regulator activity; GO:0004722 protein serine/threonine phosphatase activity, Supplementary File 2.6; Figure 2.3), delineated an ongoing reprogramming of cellular transcription and post translational protein modification related to drought recovery and osmotic adjustment in both organs, in line with the severity of the water stress applied (Verslues et al,2006). In particular, the changes observed in proline metabolism in plants (GO:0004657 proline dehydrogenase activity, GO:0006562 proline catabolic process) are a well-known response to a multiplicity of abiotic stresses, including drought (Verbruggen & Hermans, 2008). The majority of GO terms were, however, specific for shoot and root, providing compelling evidence of qualitative and quantitative differences in the responses of these organs to water deprivation: despite the highest responsivity of the root transcriptome observed above, in fact, the transcriptional response in the shoot involved 2-3 times more functions than in root (Figure 2.3). Also, the GO terms differed significantly in the different organs: in root several DEGs were associated to functions related to polysaccharide catabolism (e.g. GO:0000272, polysaccharide catabolic process; GO:0016161, beta-amylase activity), indicating extensive osmotic adjustment to reduce the water potential and limit cellular damage (Verslues et al,2006). The enrichment of terms related to biotic stress (e.g. GO:0009816, defense response to bacterium incompatible interaction; GO:0009607, response to biotic stimulus been reported, e.g., in sorghum tissue treated with PEG or ABA (Dugas et al. 2011), thus supporting the conservation of the cross talk between biotic and

abiotic stress responses in Poaceae. Also, enrichment of gene functions related to translation was observed (e.g., GO:0003735, structural constituent of ribosome; GO:0006412, translation), possibly as a response to the extensive transcriptional reprogramming observed in roots and/or to root cell growth and division. The extension of the root apparatus is indeed a common response in plants to water stress which maximizes the chance of reaching the moisture available in deeper layers of soil (Verslues et al, 2006). Worth of note, in root the only biological function specific to the milder PEG treatment (SM contrast, GO:0009685 gibberellin metabolic process; Figure 2.3b) indicates a possible involvement of gibberellins (GA) in the control of this trait through root growth. Based on the comparison of emmer wheat susceptible and resistant varieties, GA signaling and biosynthesis genes have been associated to resistance to drought in roots (Krugman et al, 2011). These results are consistent with a role of GA in the maintenance of root growth as part of the developmental decrease of the shoot/root biomass ratio usually observed in plants growing under water stress (Colebrook et al, 2014). It is thus possible that the enrichment of functions related to GA observed also in *A. donax* could contribute to the onset of the developmental changes triggered by mild water stress to increase accessibility of roots to soil with higher moisture.

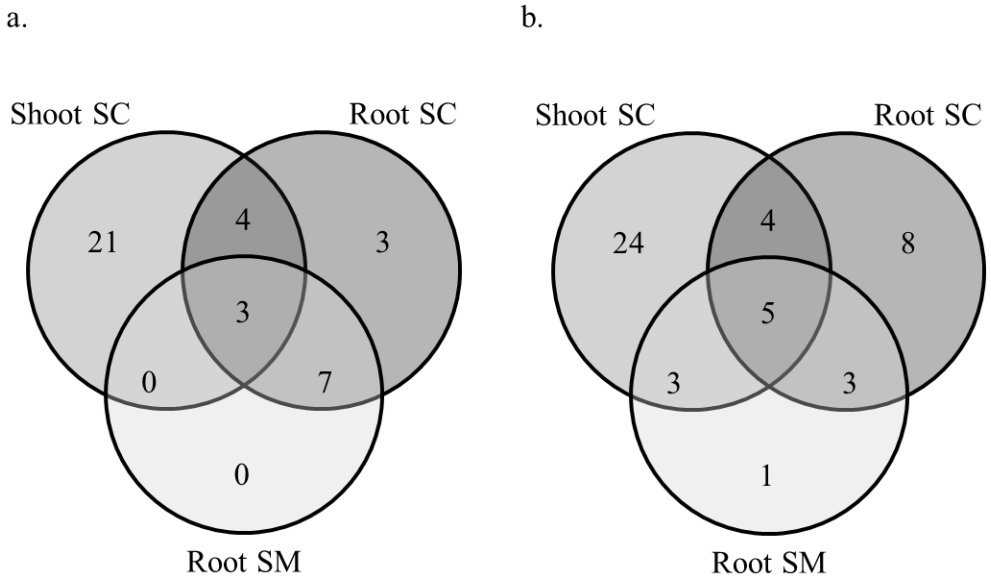


Figure 2.3. **Venn diagram of significantly enriched GOs.** The GO terms which were overrepresented under different conditions have been slimmed by REVIGO, and compared by category: a molecular function; b biological process. MC mild water stress vs. control, SC severe water stress vs control; SM severe water stress vs. mild water stress.

Compared to root, in shoot the pattern of GO terms enrichment in response to water limitation was dominated by functions related to signal transduction and protein modification associated to phosphorylation (Supplementary File 2.6; Figure 2.3). This result mirrors the dramatic

increase in post-translational phosphorylation levels observed in wheat leaves under drought stress (Zhang et al. 2014). Interestingly, several other functional classes specifically enriched in *A. donax* shoot transcriptome corresponded to those of proteins undergoing phosphorylation in wheat (e.g., GO:0009405pathogenesis, GO:0008643carbohydrate transport, GO:0005509calcium ion binding, GO:0015291secondary active transmembrane transporter activity), indicating that a synergistic effect between transcriptional and post-translation reprogramming may take place in Poaceae shoots during water stress (Zhang et al. 2014).

Taken together, the identified DEGs indicate major differences between organs in the transcriptional responses to water stress: Roots experienced a seemingly more severe/earlier stress, whereas in shoots the transcriptional response was still mainly at the level of signal transduction. Time course analyses will be required to precisely define the relative contribution of stress induction kinetics versus organ-specificity to the patterns of differential expression observed in this study. Given the relevance that the root system plays in both acclimation and adaptation of plants to water stress (Lynch et al 2014), several of the early-responsive genes identified could constitute suitable markers for the detection of early water stress in *A. donax*.

Metabolic pathways related to water stress in *A. donax*.

The set of 3034 DEGs was mapped onto KEGG pathways in *Arabidopsis thaliana* and *Oryza sativa*, highlighting the involvement of several drought-related pathways (Figure 2.4). 'Plant hormone signal transduction' (ko04075), comprising 11 DEGs in roots and 12 DEGs in shoots, was overrepresented. In this pathway, for both shoots and roots, the transcripts of several hormone-responsive proteins involved in regulation and signal transduction were up-regulated. Plant hormones play crucial roles in a diverse set of developmental processes, as well as in the response to biotic and abiotic stresses (Bari and Jones, 2009). For example, MYC2 is known to function as an activator in ABA signaling and its overexpression in Arabidopsis confers increased tolerance to drought (Abe et al, 2003). As discussed in more detail below, also ABA-activated SnRK2 is required for dehydration stress signaling in Arabidopsis (Yoshida et al, 2002). Previous studies also suggested that in rice OsJAZ1 could connect the jasmonate and drought stress signaling cascades by functionally interacting with OsbHLH148 and OsCOI1 (Seo et al, 2011). Taken together these results confirm the pivotal role played in water stress response by the differential regulation of genes involved in hormone signal transduction (Huang et al, 2012).

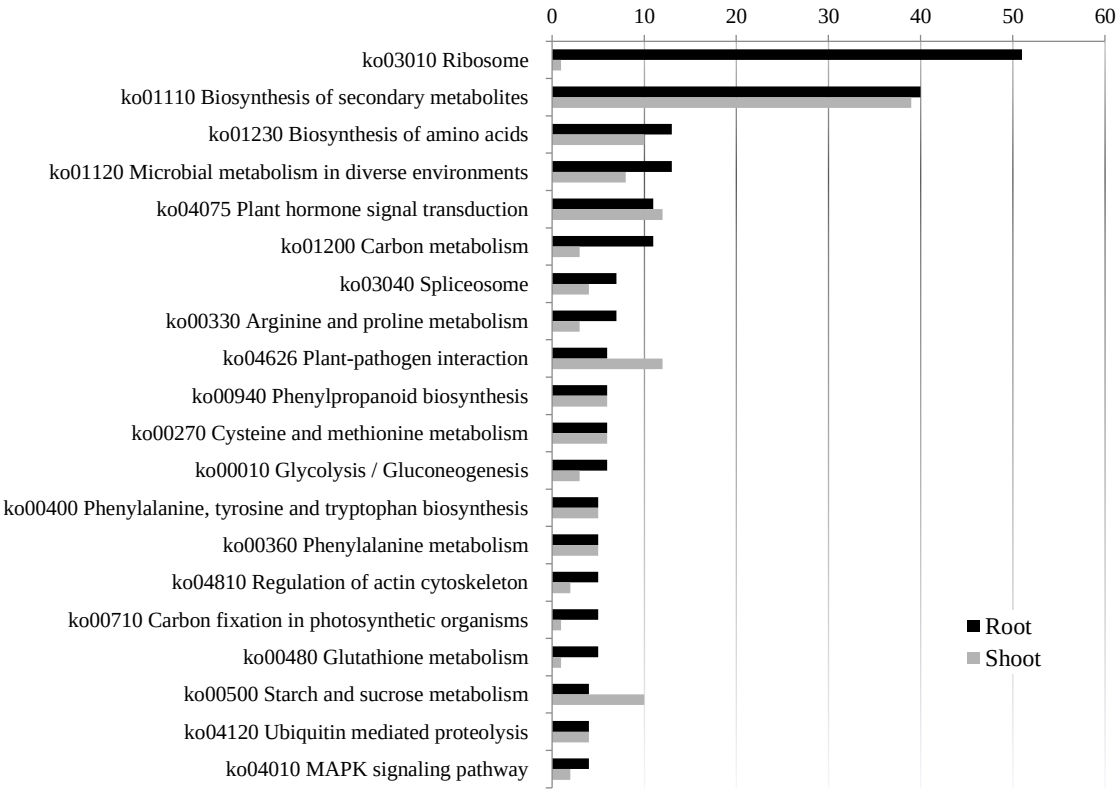


Figure 2.4. **Distribution KEGG Pathways for DEGs in shoot and root**. Data are sorted by number of root DEGs mapping to KEGG pathways. Black bar root DEGs; gray bar shoot DEGs

Two other important pathways, including 'Phenylalanine metabolism' (ko00360) and 'Plant-pathogen interaction' (ko04626), were also found in our study to be regulated by water stress (Figure 2.4). *CYP73A* (*trans-cinnamate 4-monooxygenase*), *4-coumarate--CoA ligase* and *peroxidase*, associated with 'phenylalanine metabolism', were all highly accumulated in response to water stress, in agreement to the relevance of this pathway in plant responses to drought (Gholizadeh A et al, 2011). Additionally, transcripts from *CALM* (*calmodulin*), *CML* (*calcium-binding protein: CaM-like protein*), *MYC2*, *RBOH* (*respiratory burst oxidase*), *PR1* (*pathogenesis-related protein 1*) and *JAZ* members of the 'Plant-pathogen interaction' pathway, were also induced by water stress. All these genes have been reported to be involved in response to several stresses. For example, as calcium is one of the most important signaling molecules in plants, the expression of CALM and CMLs is well regulated due to different environmental requirements in Arabidopsis (Fuchs et al. 2011). Finally, *RBOH* genes are also commonly expressed in many plants in response to biotic and abiotic stresses (Marino et al., 2012).

Other examples of relevant pathways which are known to be involved in responses to abiotic stresses in general or specifically to drought were 'Starch and sucrose metabolism' (ko00500),

‘Arginine and proline metabolism’ (ko00330), and ‘MAPK signaling pathway’ (ko04010) (Huang et al, 2012, Mohammadkhani and Heidari, 2008 and Yoshiba Y et al, 1997).

Strikingly, the biggest difference observed between root and shoot was related to ribosomal DEGs (ko03010 ribosome; Figure 2.4), thus identifying reprogramming of ribosomal translation as one of the largest responses of the root system during the early stages of water stress in *A. donax*. As noted above, it is likely that such large effect to translation could represent the early phases of the modulation of shoot/root resource allocation preluding to root cell growth and division, a typical avoidance responses of the root system during the early phases of water stress. Given the high number of ribosomal subunit genes and the complexity of their regulation as a function of water stress intensity/duration as well as species- and even genotype-dependent variation (Benešová et al, 2012), the detailed dissection of ribosome-related pathway reprogramming will be relevant for the elucidation of root-specific responses to early water stress in *A. donax*.

Identification of transcription factors responsive to water stress in *A. donax*

Transcription factors (TF) have been identified among the most promising targets for the improvement of plant performance under drought stress. Mining of DEGs for putative TFs and their interactors led to the identification of 238 *A. donax* unigenes, corresponding to 136 high confidence rice homologs previously identified as drought-responsive genes from 37 TF families (Priya & Jain 2013; Supplementary File 2.7). Because of the altered water potential under salt stress (Huang 2012), the majority of the genes (108) are also responsive to salinity. A total of 18 genes, are, however, specifically responding to drought (Supplementary File 2.7). The most represented *A. donax* differentially expressed families, constituting alone the majority of the genes, were those of NAC, WRKY, AP2-EREBP, bHLH, bZIP and AUX/IAA, which are known to mediate water stress responses in plants (Hadiarto & Tran 2011). The majority of these families were also among the most represented in drought-stressed rice (Wang et al. 2011). *A. donax* unigenes from the NAC family are the most common among differentially expressed TF genes (36 in total), matching a total of 14 different rice loci. Six of them (Os03g60080/SNAC1; Os01g66120/SNAC2/OsNAC6; Os11g08210/OsNAC5; Os11g03300/OsNAC10; Os08g06140; Os05g34830) have been previously identified as drought-responsive (Nuruzzaman M et al. 2013). Four of them have been characterized in depth through functional analyses, confirming their pivotal role in water stress-related trasncriptional reprogramming in rice. In particular, all of them have been demonstrated to be ABA-responsive (Hu et al. 2006; Hu et al. 2008; Sperotto et al. 2009; Jeong et al. 2010.), in agreement with the activation of the ABA signal translation cascade observed above.

A total of 67 out of the 150 rice homologs to differentially expressed *A. donax* TF unigenes (45%) were consistently found to be differentially expressed also in rice (Ray et al. 2011; Supplementary File 2.7.). Not all the TF families, however, were equally represented in both *A. donax* organs, indicating that part of the differences observed between shoot and root transcriptional responses may be mediated by members of these groups. In root among the families encompassing more than 5 differentially expressed unigenes, we found twice as many AP2-EREBP, AUX/IAA and MYB unigenes than in shoot. AP2-EREBP is a superfamily of transcription factors composed by the ERF, AP2 and RAV families (Riechmann & Meyerowitz 1998; Rashid et al. 2012) (Figure 2.5).

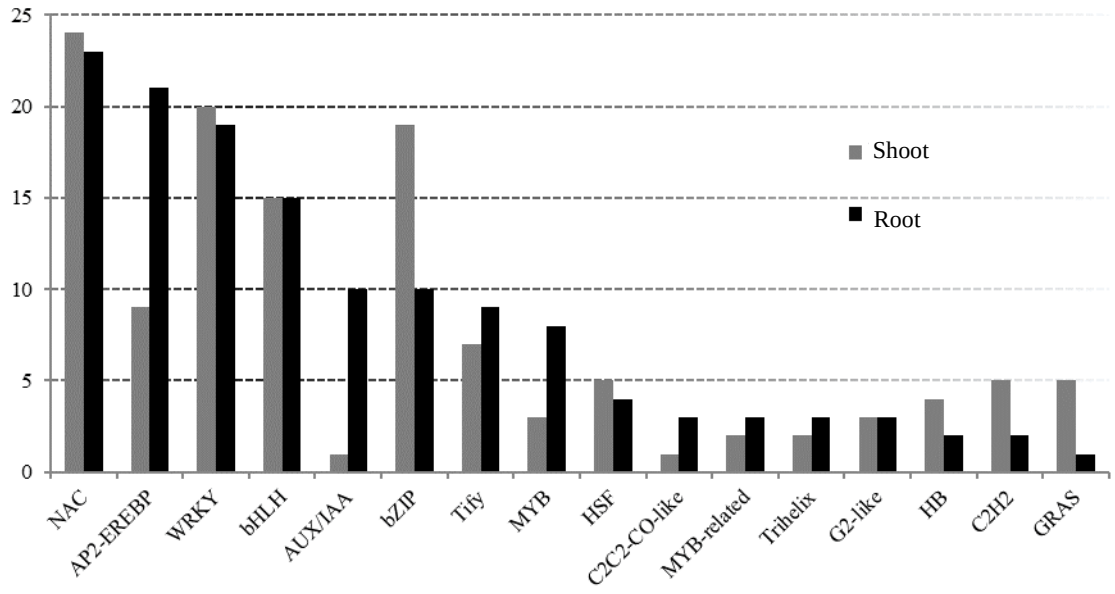


Figure 2.5. **Distribution of transcription factors responsive to water stress in *A. donax*.** Data are sorted by number of root DEGs. Only categories with more than 3 DEGs identified as transcription factors are shown. Black bar root DEGs; gray bar shoot DEGs

AP2/EREBP TFs are involved in many fundamental biological processes, ranging from development to response to biotic and abiotic stresses (Nakano et al. 2006). Five of the *A. donax* unigenes, homologous to rice genes Os02g51670 (DREB2B), Os09g20350 (DREBF1), Os04g55520 (DREB2F) and Os06g03670 (OsDREB1C/CBF), belong to the DREB subfamily of ERF TFs, which are known to control expression of several genes in response to dehydration and low temperature (Sakuma Y et al. 2002). All the four rice homologs have been directly involved in the responses to water stress, thus confirming their relevance towards this stress also in *A. donax* (Matsukura et al. 2010; Wang et al. 2008; Dubouzet et al. 2003; Moumeni A et al. 2011). Other *A. donax* unigenes from the AP2/EREBP super-family were homologs of 4 rice ERF genes

(Os03g09170; Os08g31580; Os06g10780; Os05g41780). Unlike DREB genes, ERF TFs have been associated mainly to biotic stress responses mediated by ethylene (Ohme-Takagi & Shinshi, 1995). It is, therefore, likely that the *A. donax* ERF TFs identified in this study are at least in part responsible for the enrichment of GO terms related to biotic stresses observed above, possibly with the participation of members of the NAC and WRKY families (Narsai et al. 2013).

The second TF family with most unigenes in roots compared to shoot was that of AUX/IAA. About ten times more *A. donax* unigenes from this family were differentially expressed in root upon water stress compared to shoot (Figure 2.5). AUX-IAA proteins interact with TFs of the ARF (Auxin-Responsive Factors) family, repressing root growth as consequence of the increase of the intracellular levels of the auxin plant hormone indoleacetic acid (IAA; Overvoorde et al. 2010). Since *A. donax* is a perennial plant whose large rhizomes serve as a long-term survival organ, the AUX-IAA genes specifically induced in roots are interesting candidates to dissect the coordination of *A. donax* root and shoot growth under water deprivation.

MYB and MYB-related transcription factors also were more represented in roots as compared to shoots (Figure 2.5). Two of the *A. donax* MYB unigenes were homologs to rice Os12g37690, which had been previously associated to differences among drought-sensitive and drought-tolerant rice cultivars (Degenkolbe et al. 2009). The gene is also upregulated in response to oxidative stress during the early response of japonica rice to chilling (Yun et al. 2010). Interestingly, another *A. donax* unigene was homolog to Os12g37970, a rice MYB TF involved in the coordinate regulation of cellulose and lignin biosynthesis (Ambavaram et al. 2011), indicating that it could contribute to alter the structure of cell walls in response to water stress.

Among the largest TF families displaying a higher number of differentially expressed unigenes in *A. donax* shoots compared to roots we found the bZIP, C2H2 and GRAS families (Figure 2.5). Strikingly, nine of the 21 differentially identified bZIP unigenes were homologous to rice Os02g52780. This rice gene, also called OsbZIP23 (Nijhawan et al. 2008), has been functionally demonstrated to have a relevant role among rice bZIP genes in conferring ABA-dependent drought and salinity tolerance (Xiang et al. 2008). Somehow unexpectedly, none of the rice genes homologs to differentially expressed *A. donax* unigenes (LOC_Os07g39470, LOC_Os01g62460, LOC_Os01g71970, LOC_Os07g36170, LOC_Os11g47870) have been functionally characterized, leaving open their specific role in drought responses. Two of them (Os01g71970 and Os07g36170), however, had already been identified among the few GRAS TFs differentially expressed in rice upon drought stress (Ray et al. 2011). More recently, additional evidences for the involvement of members of the GRAS family in the responses to water

deprivation have been reported for rice (Xu et al. 2015). Consistently with our results in *A. donax*, expression of *OsGRAS23* was significantly induced in rice leaves following treatments with PEG, dehydration, salt, gibberellins and jasmonic acid. In particular, transgenic rice overexpressing *OsGRAS23* (LOC_Os04g50060) was more resistant to drought and tolerant to oxidative stress compared with wild-type, thanks to the upregulation of genes involved in anti-oxidant functions (Xu et al. 2015). Taken together, these results indicate that GRAS genes in general and in particular those identified in *A. donax* represent interesting candidates for increasing water stress tolerance in monocots. Also, the majority of the rice homologs of differentially expressed *A. donax* unigenes from the C2H2 family (LOC_Os03g55540, LOC_Os03g13600, LOC_Os03g60570, LOC_Os09g38340) were have been previously identified as drought responsive (Ray et al. 2011), indicating their conserved role in Poaceae. Two among them and an additional C2H2 gene not previously identified (LOC_Os03g10140, LOC_Os09g38340, LOC_Os09g38790) are known to control the vegetative to floral phase transition in monocots (Colasanti et al. 2006; Higgins et al. 2010), indicating that responsiveness of C2H2 genes to water deprivation may be relictual in *A. donax*: while other species from the *Arundo* genus are fertile and could benefit from accelerating seed setting as a drought-escape strategy, *A. donax* is fully sterile (Hardion et al. 2015) and no clear selective advantage seems to be associated to this trait. Therefore, loss of function mutations of C2H2 or other flowering time TFs could be interesting candidates to extend the vegetative phase and, thus, biomass accumulation in *A. donax* (Sablok et al. 2014).

Characterization of co-regulated gene expression network in *A. donax*

We compared the distribution of both differentially and non-differentially expressed *A. donax* genes with the 15 drought-responsive modules of rice orthologs recently identified (Zhang et al, 2012). Only Module 7 and Module 14 were over-represented in both shoots and roots, while Module 10 was over-represented only in shoots (Table 2.2).

Rice module	Rice genes	Putative orthologs in *A. donax*	Putative orthologs in *A. donax* shoot DEGs (*p* value)	Putative orthologs in *A. donax* root DEGs (*p* value)	Putative module function
Module 1	303	149	2	8	
Module 2	213	155	5	4	
Module 3	141	61	1	0	
Module 4	134	35	2	3	
Module 5	117	71	2	2	

Module 6	90	29	4	0	
Module 7	77	46	22 (4.89E−15)	12 (3.66E−07)	Hormonal signal transduction
Module 8	48	18	0	0	
Module 9	47	16	3	2	
Module 10	47	27	8 (3.22E−4)	0	Post-translational protein modification
Module 11	46	11	0	0	
Module 12	42	11	0	0	
Module 13	38	13	0	0	
Module 14	28	17	6 (5.42E−4)	8 (3.53E−07)	Stomatal closure
Module 15	21	11	1	0	

Table 2.2. ***Comparison between*** **A. donax** ***water stress response genes and rice drought response network****. A total of 56 (in shoot) and 39 (in root)* A. donax *DEGs for which rice orthologs could be identified are mapped onto the 15 co-expression modules previously identified in rice (Zhang et al, 2012). Significance levels for over- and under-representation as compared to rice (p-value) are provided. There are 1392 rice genes in the 15 modules. Based on Blastp reciprocal best hits method, a total of 56 and 39 differentially expressed genes were identified as putative orthologs of rice genes in shoot and root of* A. donax*, respectively.*

Module 10 had been identified as a post-translational drought-related signaling/regulation cascade (genes involved in protein amino acid phosphorylation processes), further confirming the results from GO enrichments discussed above. The functions of Module 7 and Module 14, however, were not reported. Based on the functional mining of rice and *A. donax* orthologs in each module (Supplementary File 2.8), we found that Module 7 might be related to hormonal signal transduction, since these genes are mapped on JAZ, CML, PTC2_3, and ABF, which all belong to the 'Plant hormone signal transduction' pathway. The observation that PSY (phytoene synthase), which controls metabolic flux through the pathway supplying carotenoid precursors for ABA biosynthesis, is also part of this module further strengthens the identification of Module 7 as likely ABA-related co-expression module. Additionally, promoters of genes from Module 7 in rice were found to be enriched in S-BOX motif, which is the ABI4 binding site. ABI4 is known to be an important link between ABA hormone and glucose signaling pathway, and it has been proposed that in some species carbohydrate metabolism might be the initial response to drought (Acevedo-Hernández, León and Herrera-Estrella, 2005;). Thus, genes belonging to Module 7 likely play a conserved role

in Poaceae in the ABA-mediated modulation of carbohydrate metabolism in response to water stress, as observed also in some dicotyledonous species (Pinheiro et al, 2011). Meanwhile, Module 14 comprises several genes related to the ABA hormone: PP2C, which upon inhibition by ABA, is a fundamental trigger in stress-related ABA signaling cascade (Umezawa et al, 2010); raffinose synthase [EC:2.4.1.82], which is involved in the biosynthesis of raffinose, an osmoprotectant associated to drought tolerance (Taji et al, 2002). Taken together, these results support the view that Module 14 is likely involved in a plant hormone transduction pathway related to ABA, necessary for the early onset of stomatal closure. Recent physiological analyses indicate that *A. donax* can fix CO_2 at soil water contents close to wilting point, thanks to its ability to effectively control stomatal regulation in relation to soil water content (Cosentino et al, 2016). The association in our transcriptomics data of Module 7 and 14 to ABA-related pathways confirms and further extends this observation, indicating that in *A. donax* such regulation can be activated as early as 1 h after the onset of PEG-induced water stress and that it may contribute to the high adaptability of this species to resource-poor habitats and marginal soils (Lewandowski et al, 2003). Despite the incomplete understanding of the functions of the genes comprised in these clusters, selected members of both Module 7 and 14 could constitute, on one hand, suitable markers to dissect early stress responses and, on the other hand, promising targets to modulate drought tolerance in *A. donax*.

Interestingly, in the afore-mentioned study, only Modules 4, 7 and 14 are significantly associated to rice early responses to drought, where an experimental design similar to ours (3 hours treatment, two tissues) has been used (Zhang et al, 2012). This match supports the conservation of early drought response networks between *A. donax* and rice, two species associated to water-rich environments. A closer examination of the genes belonging to the latter three modules extends the possible conservation of drought-related regulatory networks even further: Among the 53 drought response genes in common among *A. donax*, rice, sorghum and foxtail (see next paragraph), 13.2% of genes (7 genes) are from Module 14. Considering that there are only 28 genes in Module 14, much less than the others, especially this module seems to capture a particularly important drought-related mechanism across Poaceae species.

Identification of a core set of Poaceae genes differentially regulated upon water stress

The comparison of transcriptomes across different species can provide information about conservation of gene functions over evolutionary time. We, therefore, identified the subsets of water stress-related DEGs in common between *A. donax*, foxtail, sorghum, and rice. When the *A. donax* DEGs were compared with drought-responsive genes reported in previous studies (Qi et al. 2013;

Dugas et al. 2011; Zhang et al. 2012), a total of 343, 496 and 143 putative orthologs were identified from foxtail, sorghum and rice, respectively (Figure 2.6; Supplementary File 2.9). In total 53 groups of putative orthologs present in all species were identified, which constitute a core of evolutionarily conserved genes associated to early responses to water deprivation.

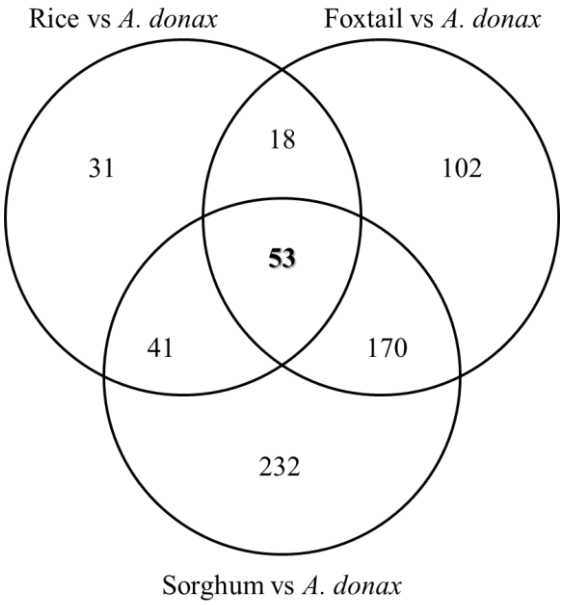


Figure 2.6. **Drought response genes comparison across *A. donax*, rice, foxtail, and sorghum**. The Venn diagram represents putative orthologs of *A. donax* stress-responsive genes identified by OrthoMCL in at least two species

Some of these genes are involved in the quality control and targeted degradation of proteins. Possibly the most striking example among them is a putative ortholog of *Arabidopsis* AT5G51070 gene, which codes for the ClpD subunit of the plastidial Hsp100/Clp complex, a caseinolytic protease (Clp) necessary for chloroplast biogenesis and protein homeostasis (Supplementary File 2.9) (Bruch et al. 2012). This gene, also known as *Early Responsive to Dehydration 1* (*ERD1*), is an ATP-dependent molecular chaperone that likely directs unfolded polypeptides to the Clp complex for degradation (Colombo et al. 2014), suggesting that also in *A. donax* it could help eliminate damaged/misfolded proteins and aid proteome reprogramming upon water stress. This observation is supported by the fact that in rice the ClpD protein has been reported to be preferentially upregulated along with several other proteases in response to conditions of active drought signaling but water availability, which in shoots mimicks the early water stress of our study (Mirzaei et al. 2014).

Several of the other conserved DEGs participate in the biosynthesis of different metabolites, ranging from sugars to hormones, lipids and flavonoids (Supplementary File 2.9). Among the most interesting *A. donax* candidates involved in sugar metabolism there is a putative raffinose synthase

protein homolog to *Arabidopsis* AT5G40390, the only isoform reported to be responsive to a wide array of abiotic stresses, including water deficit (Egert et al. 2013). Knockout mutants of AT5G40390 have reduced amounts of verbascose, sucrose and mannitol, but, surprisingly, their shoots are slightly less susceptible to prolonged drought than WT (Anderson et al. 2001). The *Arundo* homolog of this gene, however, is upregulated in roots, indicating that significant regulatory differences exist between species and suggesting that in the latter species this gene could contribute to short-term osmotic adjustment in roots.

In agreement with previous studies, also a certain number of membrane transporters are among the conserved genes involved in the early response to water stress (Supplementary File 2.9). For instance, a homolog of *Arabidopsis* gene AT3G20300 is upregulated in both shoots and roots. In *Arabidopsis,* this poorly characterized gene belongs to the monosaccharide transporter-like (MST-like) superfamily and codes for a predicted polyol/cyclitol/monosaccharide-H+-symporter of the mitochondrion, indicating that in *A. donax* its ortholog could be involved in the response to water deprivation by redistributing small organic solutes between cytosol and mitochondria. Another transporter which was upregulated in both shoots and roots is the homolog of *Arabidopsis* gene AT1G15520, coding for ABCG40, a plasma membrane ABA uptake transporter. In *Arabidopsis*, stomata of *abcg40* mutants respond more slowly to ABA and are less drought tolerant than WT plants (Kang et al. 2010). Interestingly, the putative ortholog of AT1G78390, nine-cis-epoxycarotenoid dioxygenase 9, a key enzyme in ABA biosynthesis (Lefebvre et al. 2006), is strongly upregulated in both *A. donax* shoots and roots during the early responses to water stress. Taken together, these results are in line with the established role of ABA as the main plant hormone in the early responses to water stress (Huang 2012). The identification of several differentially expressed ABA-responsive kinases and phosphatases allowed also the definition of a conserved core of signalling genes shared between *A. donax* and the other Poaceae considered. The putative ortholog of *Arabidopsis* AT4G33950 gene is strongly upregulated in water-stressed *A. donax* shoots. This gene is a member of SNF1-related protein kinases (SnRK2) responsive to both ionic and non-ionic osmotic stresses. Among the SnRK2 paralogs, AT4G33950 (also called SnRK2.6) is the most important for overall stomatal control (Virlouvet and Fromm 2015). In *Arabidopsis*, loss of function mutations of this gene completely abolishes ABA-mediated stomatal responses, leaving unaffected the ABA-independent reactions and resulting in increased drought susceptibility (Mustilli et al. 2002). Given the proposed involvement of this gene in the early phases of ABA perception before the development of reactive oxygen species associated to cell damage (Mustilli et al. 2002), the ortholog of AT4G33950 could constitute an interesting candidate to modulate the responsiveness to water stress responses in *A. donax* and be used as a sensitive marker for shoot

drought stress. Two additional kinases are specifically upregulated in *A. donax* leaves subjected to water stress. The first one, homologous to gene AT1G70520, encodes a cystein-rich receptor-like protein kinase which in *Arabidopsis* has been shown to respond only weakly to ABA and other hormone treatments, but to be upregulated shortly after ozone treatment (Wrzaczek et al. 2010). The second one is a poorly characterized protein kinase homolog to AT1G56130, one of the four paralogous loci present as tandem duplications in the *Arabidopsis* genome. Possibly due to redundancy, very limited functional information is available about this small gene family. However, the early response and high levels of upregulation in water-stressed shoots of *A. donax* makes it an interesting candidate deserving further characterization. In addition to protein kinases, also a phosphatase homologous to *Arabidopsis* AT2G29380 gene, also called Highly ABA-Induced1 (*HAI1*), is among the conserved Poaceae DEGs. Like several other water-stress clade A protein phosphatase 2Cs (PP2Cs), *HAI1* acts as a negative regulator of osmoregulatory solute accumulation. Unlike the majority of its closest paralogs, however, the HAI1 protein is largely insensitive to inhibition by members of the ABA receptors family (Bhaskara et al. 2012). The concomitant expression of *A. donax* homologs of HAI1 and SnRK2.6 (the latter acting downstream of the other ABA-receptor repressible PP2Cs; Soon et al, 2012) raises the interesting possibility that HAI1 may act antagonistically to SnRK2.6 to prevent excessive osmoregulatory solute accumulation. This hypothesis is supported by the fact that *Arabidopsis hai1* mutants accumulate higher amounts of proline and other osmoregulatory solutes than wild type plants (Bhaskara et al. 2012). Strikingly, among the conserved early water stress DEGs, the only transcription factor is a homolog of *ATHB7* (Arabidopsis gene AT2G46680), a member of class I plant-specific homeodomain-leucine zipper family (Johannesson et al. 2001). In *Arabidopsis*, *ATHB7* and its paralog *ATHB12* modulate abscisic acid signalling by regulating protein phosphatase 2C and abscisic acid receptor gene activities (Valdés et al. 2012). Despite direct regulation of *HAI1* has not been tested, several other clades A *PP2Cs* are under positive transcriptional regulation by *ATHB7* / *ATHB12*, which at the same time repress transcription of genes from the PYR/PYL family of ABA receptors (Valdés et al. 2012). As, both in *Arabidopsis* and rice, paralogs with different tissue specific and developmental expression patterns have been implicated in different aspects of ABA-mediated growth responses to water stress (Harris et al. 2011), the characterization of Arundo's homeodomain-leucine zipper family members seems to be a promising starting point to dissect the details of abscisic acid signalling modulation and stomatal control in this species.

2.3 Conclusion

The lack of information available about the molecular mechanisms involved in stress responses in *A. donax* is currently a major constraint for the further development of this semi-wild species into a fully-fledged bioenergy crop. To fill at least in part this gap, we hereby provided the first characterization of *A. donax* shoot and root transcriptomes in response to water stress, one of the factors of highest concern for its productivity. Given the commonality of the responses to water limitation and other stresses, in addition to providing a general overview of the early transcriptional responses to simulated drought, our results shed also light at the molecular level on the general mechanisms of stress response and adaptation in *A. donax*. Upon functional validation, thus, many of the unigenes identified in the present study have the potential to be used for the development of novel *A. donax* varieties with improved productivity and stress tolerance. In addition, the inventory of early-responsive genes to water stress provided in this study could constitute useful markers of the physiological status of *A. donax* plantations to deepen our understanding of its productivity under water limitation.

2.4 Material and Methods

Plant material and application of water limitation stress

In the present study, we applied a water stress by treating cohorts of *A. donax* cuttings (collected in Sesto Fiorentino, Florence, Italy 43°49'01.8"N 11°11'57.0"E) with two different concentrations of polyethylene glycol 6000 (PEG; 10% and 20% w/w, referred to as mild and severe water stress conditions, corresponding to osmotic potentials of -1.54 bars and -5.04 bars, respectively; Michel & Kaufmann 1973). Briefly, *A. donax* cuttings were let rooting in tap water, then transferred to 1% Hoagland solution and grown in a growth chamber with day-length of 16 hrs, light intensity of 150 µmol of photons m^{-2} s^{-1}, 24°C and 60% RH. At the three-leaves stage, two cohorts of cuttings were transferred to 1% Hoagland solution containing either 10% or 20% PEG, while a third cohort used as control was transferred to 1% Hoagland solution without PEG. After 1 hr the tissue above the third leaf and roots were separately collected from each cohort, quickly rinsed in distilled water and snap-frozen in liquid nitrogen. A total of 18 samples (three biological replicates from both shoot and root for each of the three conditions) were sampled.

Next generation RNA sequencing

Frozen root and the shoot tissues were grounded in liquid nitrogen with pre-cooled mortars and pestles. RNA isolation was carried out using the Spectrum Plant Total RNA Extraction Kit (Sigma) and Rneasy® Plant Mini Kit, respectively, for shoots and roots according to the manufacturer's instructions. The quantity and the quality of the isolated RNA was evaluated using agarose gel electrophoresis and spectrophotometric measurements. Additionally, the isolated RNA was subjected to quality checks using the RNA 6000 Pico kit and the Agilent Bioanalyser 2100 (Agilent). Paired-end RNA-Seq libraries were prepared using the TruSeq RNA Sample Prep V2 kit (Illumina, San Diego, CA), pooled in equimolar ratio and sequenced on an Illumina HiSeq2000 sequencer (CIBIO NGS Facility, Povo (TN), Italy).

A minimum of 694 million reads were obtained from each of the 18 libraries sequenced (Supplementary 2.1). RNA-Seq data are available in the ArrayExpress database (www.ebi.ac.uk/arrayexpress) under accession number [ArrayExpress:E-MTAB-3769]. Assessment of read quality metrics was carried out using the FastQC software (available at http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) after which stringent quality filtering, removal of reads containing Ns and de-duplication was carried out as previously described (Sablok et al, 2014). For transcript reconstruction, we concatenated all read pairs passing the quality checks and assembled them using the Trinity software with-group-pair-distance=500 and −min-cov=5 to limit intron retention in reconstructed transcripts (Grabherr et al, 2011). After discarding transcripts shorter than 200 bp, transcript redundancy was reduced using CD-HIT-EST with 95 % identity and a word size of 8 (Li and Godzik, 2006). The resulting non-redundant transcript dataset was further assembled into unigenes using the Overlap-Layout-Consensus assembler MIRA (parameters: job = denovo, est, accurate, 454 using the notraceinfo option) (Chevreux et al, 2004). All final *A. donax* unigenes are available for homology searches and download through a dedicated web Blast server (http://ecogenomics.fmach.it/arundo/) (Priyam et al, 2015).

Water stress transcriptome annotation

Following the assembly, transcriptome curation was carried out by perfoming BLASTx searchers with E-value threshold $1E^{-5}$ against NCBI non-redundant (www.ncbi.nlm.nih.gov), UniProt (www.uniprot.org) and TrEMBL (Bairoch A, and Apweiler R., 2000) plant databases. Additionally, we also curated the unigenes using the FastAnnotator program, which integrates the functionality of BLAST2GO, PRIAM, domain identification and Gene Ontology classification

(Chen et al. 2011). We further slimmed the obtained Gene Ontology (GO) categories using the PlantGO Slim categories available from the Gene Ontology consortium (www.geneontology.org) by collapsing the small child categories into broader classification of gene ontologies. Protein domains of transcriptome unigenes were identified using InterPro (https://www.ebi.ac.uk/interpro/). To identify the putative homologs of stress responsive genes characterized so far in *Arabidopsis thaliana*, BLASTx searches were performed with an E-value threshold of $1E^{-5}$ against the ASPRGDB database (Subhomoi Borkotoky et al, 2013), retaining only hits with query sequence coverage and identity higher than 50%. Additionally, we mined functionally relevant genes involved in drought stress by creating a customized, manually curated database from *Sorghum bicolor*, *Zea mays*, *Arabidopsis thaliana*, and *Oryza sativa*.

Identification and functional classification of differentially expressed genes

To identity genes which are differentially expressed upon water stress, reads from each of the 18 libraries were individually mapped on the unigene assembly and fragments per kilobase of exon per million fragments mapped (FPKM) values were estimated as a measure of the expression using RSEM (Bo and Li, 2011). For the identification of differentially expressed genes, we used EdgeR (R version: 3.0.1, edgeR version: 3.4.2; Robinson et al, 2010), implementing the Generalized Linear Model (GLM) (Smyth and Verbyla, 1996) approach. For the normalization of the read count, we applied the trimmed mean of M-values (TMM) normalization method (Robinson and Oshlack, 2010). Additionally contrasts were made to identify the set of differentially expressed genes between mild and severe water stress among the induced treatments. A false discovery rate (FDR) cutoff of 0.001 and a log-fold change (LogFC) threshold of 2 was implemented to filter the statistically significant up- and down-regulated genes between the treatment and the control. The genes with logFC>=2 and logFC<=-2 with a FDR cutoff of FDR=0.001 between two treatment conditions were determined to be up-regulated and down-regulated, respectively. All the statistically significant up- and down-regulated differentially expressed genes were custom annotated against the functionally identified drought- responsive genes in model grass clade. To select the most interesting candidates for functional studies, genes were first ranked all DEGs according to absolute difference of FPKM values and logFC between control and each of the stress conditions. To further select genes with the highest consistency in DE among biological replicates, only genes with CV < 0.7 and |logFC| >= 2 were retained, and the top 20 genes from either top- or down-regulated lists were highlighted in the results. For the identification of transcription factors responsive to water stress in *A. donax* we mined the Stress Responsive Transcriptio Factor Database

of rice (SRTFDB; Priya et al, 2013) by Blastn searches with an E-value cutoff of 1e-5. For the identification of the subsets of water stress-related DEGs in common between *A. donax*, foxtail, sorghum and rice, OrthoMCL software V5 was used with default settings (Li et al, 2003). For the comparative study between *A. donax* and rice co-regulation network, we identificated putative orthologs between the *A. donax* and rice with Blastp Reciprocal Best hits method (RBH, E-value $1e^{-6}$; Moreno-Hagelsieb and Latimer, 2008).

GO enrichment

Blast2GO was also used for a GO functional enrichment analysis of certain genes, by performing Fisher's exact test with a robust FDR (<0.05) correction to obtain an adjusted p-value between certain test gene groups and the whole transcriptome annotation. To provide a more comprehensive interpretation and of GO data we have used the freely available web-based software REVIGO (Supek et al, 2011). We uploaded the lists of over-represented GO ID along with the p-value from the result of the fisher's exact test. The analysis was run by selecting the small size of the resulting list, with the numbers associated to GO categories p-values, with the Oryza sativa database, and the SimRel as the semantic similarity measure. The GO terms which were over-represented under different conditions were slimmed by REVIGO (Supplementary File 2.6). The number of over-represented terms in common between conditions was displayed as Venn diagram (Figure 2.3: a. molecular function; b. biological process. MC: mild water stress *vs* control, SC: severe water stress *vs* control; SM: severe water stress *vs* mild water stress).

Pathway enrichment

To identify functionally relevant patterns associated to water stress in shoot and root DEGs, we created a unigene dataset for each organ discarding genes with FPKM ≤1. Each dataset was subsequently used as background to identify over- and under-represented GO categories among DEGs using the fischer's test with a p-value cut-off of 0.05. In addition, pathway enrichment analysis of DEGs were carried out with the KOBAS software (Xie et al, 2011) using BLASTx searches against the *Oryza sativa* var. *japonica* proteins.

Real-time validation of selected DEG candidates using qRT-PCR:

Each RNA sample was treated with DNase I (Sigma-Aldrich) and 1 µg of total RNA was reversed transcribed using the SuperScript® III Reverse Transcriptase (Life Technologies),

according to the manufacturer's instructions. Real time qRT-PCR was performed for a total of 10 DEGs with Platinum® SYBR® Green qPCR SuperMix-UDG and carried out in the Bio-Rad C1000 Thermal Cycler detection system according to the manufacturer's instructions. All the genes were normalized with putative *A. donax* actin protein with highest homology to sorghum AC1 gene (GenBank accession no. P53504). Each PCR reaction (12,5 μL) contained 11 μL real-time PCR Mix, 0.25 μM of each primer and 1 μl of a 1:5 dilution of cDNA. The thermal cycling conditions were 50 °C for 2 min, 95 °C for 2 min, followed by 40 cycles of 15 s at 95 °C and 30 s at 60 °C. All reactions were performed in triplicate and fold change measurements calculated with the $2^{-\Delta\Delta CT}$ method (Livak & Schmittgen, 2001). Sequences of primers used for real time PCR are provided in Supplementary File 2.2.

2.5 References

Abe H, Urao T, Ito T, Seki M, Shinozaki K. Transcriptional activators in abscisic acid signaling. Plant Cell. 2003;15:63–78.

Acevedo-Hernández GJ, León P, Herrera-Estrella LR. Sugar and ABA responsiveness of a minimal RBCS light-responsive unit is mediated by direct binding of ABI4. Plant J. 2005;43:506–19.

Acres G. Alcoholic Fuels. Platin Met Rev. 2007;51:34–5.

Ambavaram MMR, Krishnan A, Trijatmiko KR, Pereira A. Coordinated activation of cellulose and repression of lignin biosynthesis pathways in rice. Plant Physiol. 2011;155:916–31.

Anderson CM, Kohorn BD: Inactivation of Arabidopsis SIP1 leads to reduced levels of sugars and drought tolerance. J Plant Physiol 2001, 158:1215–9.

Bairoch A, Apweiler R. The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. Nucleic Acids Res. 2000;28:45–8.

Balogh E, Herr JM, Czakó M, Márton L. Defective development of male and female gametophytes in Arundo donax L. (POACEAE). Biomass and Bioenergy. 2012;45:265–9.

Bari R, Jones JDG. Role of plant hormones in plant defence responses. Plant Mol Biol. 2009;69:473–88.

Benešová M, Holá D, Fischer L, Jedelský PL, Hnilička F, Wilhelmová N, Rothová O, Kočová M, Procházková D, Honnerová J, Fridrichová L, Hniličková H. The physiology and proteomics of drought tolerance in maize: early stomatal closure as a cause of lower tolerance to short-term dehydration? PLoS One. 2012;7:e38017.

Bhaskara GB, Nguyen TT, Verslues PE. Unique drought resistance functions of the highly ABA-induced clade A Protein Phosphatase 2Cs. Plant Physiol. 2012;160:379–95.

Borkotoky S, Saravanan V, Jaiswal A, Das B, Selvaraj S, Murali A, Lakshmi PT V. The Arabidopsis stress responsive gene database. Int J Plant Genomics. 2013;2013:949564.

Bruce WB, Edmeades GO, Barker TC. Molecular and physiological approaches to maize improvement for drought tolerance. J Exp Bot. 2002;53:13–25.

Bruch EM, Rosano GL, Ceccarelli EA. Chloroplastic Hsp100 chaperones ClpC2 and ClpD interact in vitro with a transit peptide only when it is located at the N-terminus of a protein. BMC Plant Biol. 2012;12:57.

Bucci A, Cassani E, Landoni M, Cantaluppi E, Pilu R. Analysis of chromosome number and speculations on the origin of Arundo donax L. (Giant Reed). Cytol Genet. 2013;47:237–41.

Byrt CS, Grof CPL, Furbank RT. C4 plants as biofuel feedstocks: optimising biomass production and feedstock quality from a lignocellulosic perspective. J Integr Plant Biol. 2011;53:120–35.

Chen TW, Gan RCR, Wu TH, Huang PJ, Lee CY, Chen YYM, Chen CC, Tang P. FastAnnotator--an efficient transcript

annotation web tool. BMC Genomics. 2012;13:S9.

Chevreux B, Pfisterer T, Drescher B, Driesel AJ, Müller WEG, Wetter T, Suhai S. Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs. Genome Res. 2004;14:1147–59.

Claeys H, Inzé D. The agony of choice: how plants balance growth and survival under water-limiting conditions. Plant Physiol. 2013;162:1768–79.

Cohen D, Bogeat-Triboulot MB, Tisserant E, Balzergue S, Martin-Magniette ML, Lelandais G, Ningre N, Renou J-P, Tamby J-P, Le Thiec D, Hummel I. Comparative transcriptomics of drought responses in Populus: a meta-analysis of genome-wide expression profiling in mature leaves and root apices across two genotypes. BMC Genomics. 2010;11:630.

Colasanti J, Tremblay R, Wong AYM, Coneva V, Kozaki A, Mable BK. The maize INDETERMINATE1 flowering time regulator defines a highly conserved zinc finger protein family in higher plants. BMC Genomics. 2006;7:158.

Colombo C V, Ceccarelli E a, Rosano GL. Characterization of the accessory protein ClpT1 from Arabidopsis thaliana: oligomerization status and interaction with Hsp100 chaperones. BMC Plant Biol. 2014;14:228.

Dai A. Increasing drought under global warming in observations and models. Nat Clim Chang. 2012;3:52–8.

Daszkowska-Golec A, Szarejko I. Open or close the gate - stomata action under the control of phytohormones in drought stress conditions. Front Plant Sci. 2013;4:138.

Davidson RM, Gowda M, Moghe G, Lin H, Vaillancourt B, Shiu S-H, Jiang N, Robin Buell C. Comparative transcriptomics of three Poaceae species reveals patterns of gene expression evolution. Plant J. 2012;71:492–502.

Degenkolbe T, Do PT, Zuther E, Repsilber D, Walther D, Hincha DK, Köhl KI. Expression profiling of rice cultivars differing in their tolerance to long-term drought stress. Plant Mol Biol. 2009;69:133–53.

Deyholos MK. Making the most of drought and salinity transcriptomics. Plant, Cell Environ. 2010;33:648–54.

Di Tomaso JM., Healey EA. Aquatic and riparian weeds of the west. University of California – Agriculture and Natural Resources Publication; 2003. p. 254–62.

Doebley JF, Gaut BS, Smith BD. The molecular genetics of crop domestication. Cell. 2006;1309–21.

Dubouzet JG, Sakuma Y, Ito Y, Kasuga M, Dubouzet EG, Miura S, Seki M, Shinozaki K, Yamaguchi-Shinozaki K. OsDREB genes in rice, Oryza sativa L., encode transcription activators that function in drought-, high-salt- and cold-responsive gene expression. Plant J. 2003;33:751–63.

Dugas DV, Monaco MK, Olsen A, Klein RR, Kumari S, Ware D, Klein PE. Functional annotation of the transcriptome of Sorghum bicolor in response to osmotic stress and abscisic acid. BMC Genomics. 2011;12:514.

Edenhofer O, Pichs-Madruga R, Sokona Y, Farahani E, Kadner S, Seyboth K, Adler A, Baum I, Brunner S, Eickemeier P, Kriemann B, Savolainen J, Schlömer S, von Stechow C, Zwickel T, Minx JC. Climate change 2014: Mitigation of climate change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press; 2014.

Egert A, Keller F, Peters S: Abiotic stress-induced accumulation of raffinose in Arabidopsis leaves is mediated by a single raffinose synthase (RS5, At5g40390). BMC Plant Biol 2013, 13:218.

European Commission (2009) Directive 2009/28/EC of the European Parliament and of the Council of 23 April 2009 on the promotion of the use of energy from renewable sources and amending and subsequently repealing directives 201/77/EC and 2003/30/EC (Text with EEA relevance). In European Parliament, C.o.t.M.S (ed) European Union

Ferreira de Carvalho J, Poulain J, Da Silva C, Wincker P, Michon-Coudouel S, Dheilly A, Naquin D, Boutte J, Salmon A, Ainouche M. Transcriptome de novo assembly from next-generation sequencing and comparative analyses in the hexaploid salt marsh species Spartina maritima and Spartina alterniflora (Poaceae). Heredity (Edinb). 2013;110:181–93.

Fuchs M, Vothknecht UC. Calmodulin and Calmodulin-like proteins in Arabidopsis thaliana. J Endocytobiosis Cell Res. 2011;21:114–21.

Gholizadeh A. Effects of drought on the activity of Phenylalanine ammonia lyase in the leaves and roots of maize inbreds. Aust J Basic Appl Sci. 2011;5: 952–6.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat

Biotechnol. 2011;29:644–52.

Gruenheit N, Deusch O, Esser C, Becker M, Voelckel C, Lockhart P. Cutoffs and k-mers: implications from a transcriptome study in allopolyploid plants. BMC Genomics. 2012;13:92.

Guerra D, Crosatti C, Khoshro HH, Mastrangelo AM, Mica E, Mazzucotelli E. Post-transcriptional and post-translational regulations of drought and heat response in plants: a spider's web of mechanisms. Front Plant Sci. 2015;6:57.

Haddadchi A, Gross CL, Fatemi M. The expansion of sterile Arundo donax (Poaceae) in southeastern Australia is accompanied by genotypic variation. Aquat Bot. 2013;104:153–61.

Hadiarto T, Tran L-SP. Progress studies of drought-responsive genes in rice. Plant Cell Rep. 2011;30:297–310.

Hardion L, Verlaque R, Rosato M, Rosselló JA, Vila B. Impact of polyploidy on fertility variation of Mediterranean Arundo L. (Poaceae). C R Biol. 2015;338:298–306.

Hardion L, Verlaque R, Saltonstall K, Leriche A, Vila B. Origin of the invasive Arundo donax (Poaceae): a trans-Asian expedition in herbaria. Ann Bot. 2014;114:455–62.

Harris JC, Hrmova M, Lopato S, Langridge P. Modulation of plant growth by HD-Zip class I and II transcription factors in response to environmental stimuli. New Phytol. 2011;190:823–37.

Higgins JA, Bailey PC, Laurie DA. Comparative genomics of flowering time pathways using Brachypodium distachyon as a model for the temperate grasses. PLoS One. 2010;5:e10065.

Hu H, Dai M, Yao J, Xiao B, Li X, Zhang Q, Xiong L. Overexpressing a NAM, ATAF, and CUC (NAC) transcription factor enhances drought resistance and salt tolerance in rice. Proc Natl Acad Sci U S A. 2006;103:12987–92.

Hu H, You J, Fang Y, Zhu X, Qi Z, Xiong L. Characterization of transcription factor gene SNAC2 conferring cold and salt tolerance in rice. Plant Mol Biol. 2008;67:169–81.

Huang GT, Ma SL, Bai LP, Zhang L, Ma H, Jia P, Liu J, Zhong M, Guo Z-F. Signal transduction during cold, salt, and drought stresses in plants. Mol Biol Rep. 2012;39:969–87.

Jeong JS, Kim YS, Baek KH, Jung H, Ha S-H, Do Choi Y, Kim M, Reuzeau C, Kim J-K. Root-specific expression of OsNAC10 improves drought tolerance and grain yield in rice under field drought conditions. Plant Physiol. 2010;153:185–97.

Johannesson H, Wang Y, Engström P. DNA-binding and dimerization preferences of Arabidopsis homeodomain-leucine zipper transcription factors in vitro. Plant Mol Biol. 2001;45:63–73.

Jørgensen U. Benefits versus risks of growing biofuel crops: the case of Miscanthus. Curr Opin Environ Sustain. 2011;3:24–30.

Kakumanu A, Ambavaram MMR, Klumas C, Krishnan A, Batlang U, Myers E, Grene R, Pereira A. Effects of drought on gene expression in maize reproductive and leaf meristem tissue revealed by RNA-Seq. Plant Physiol. 2012;160:846–67.

Kang J, Hwang JU, Lee M, Kim YY, Assmann SM, Martinoia E, Lee Y. PDR-type ABC transporter mediates cellular uptake of the phytohormone abscisic acid. Proc Natl Acad Sci U S A. 2010;107:2355–60.

Kido ÉA, Neto JRCF, Silva RLD, Pandolfi V, Guimarães ACR, Veiga DT, Chabregas SM, Crovella S, Benko-Iseppon AM. New insights in the sugarcane transcriptome responding to drought stress as revealed by superSAGE. ScientificWorld Journal. 2012;2012:821062.

Lagerwerff JV, Ogata G, Eagle HE. Control of osmotic pressure of culture solutions with polyethylene glycol. Science. 1961;133:1486–7.

Lal R. Soil quality impacts of residue removal for bioethanol production. Soil Tillage Res. 2009;102:233–41.

Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009;10:R25.

Lefebvre V, North H, Frey A, Sotta B, Seo M, Okamoto M, Nambara E, Marion-Poll A. Functional analysis of Arabidopsis NCED6 and NCED9 genes indicates that ABA synthesized in the endosperm is involved in the induction of seed dormancy. Plant J. 2006;45:309–19.

Lewandowski I, Clifton-Brown JC. Miscanthus: European experience with a novel energy crop. Biomass and Bioenergy. 2000;19:209–27.

Lewandowski I, Scurlock JMO, Lindvall E, Christou M. The development and current status of perennial rhizomatous grasses as energy crops in the US and Europe. Biomass and Bioenergy. 2003;25:335–61.

Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-seq data with or without a reference genome. BMC Bioinformatics. 2011;12:323.

Li F, Tsfadia O, Wurtzel ET. The phytoene synthase gene family in the Grasses: subfunctionalization provides tissue-specific control of carotenogenesis. Plant Signal Behav. 2009;4:208–11.

Li L, Stoeckert CJ, Roos DS. OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Res. 2003;13:2178–89.

Li W, Godzik A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics. 2006;22:1658–9.

Liu L, Li Y, Li S, Hu N, He Y, Pong R, Lin D, Lu L, Law M. Comparison of next-generation sequencing systems. J Biomed Biotechnol. 2012;2012:251364.

Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods San Diego Calif. 2001;25:402–8.

Lynch JP, Chimungu JG, Brown KM. Root anatomical phenes associated with water acquisition from drying soil: targets for crop improvement. J Exp Bot. 2014;65:6155–66.

Mariani C, Cabrini R, Danin A, Piffanelli P, Fricano A, Gomarasca S, Dicandilo M, Grassi F, Soave C. Origin, diffusion and reproduction of the giant reed (Arundo donax L.): a promising weedy energy crop. Ann Appl Biol. 2010;157:191–202.

Marino D, Dunand C, Puppo A, Pauly N. A burst of plant NADPH oxidases. Trends Plant Sci. 2012;17:9–15.

Martin LBB, Fei Z, Giovannoni JJ, Rose JKC. Catalyzing plant science research with RNA-seq. Front Plant Sci. 2013;4:66.

Matsukura S, Mizoi J, Yoshida T, Todaka D, Ito Y, Maruyama K, Shinozaki K, Yamaguchi-Shinozaki K. Comprehensive analysis of rice DREB2-type genes that encode transcription factors involved in the expression of abiotic stress-responsive genes. Mol Genet Genomics. 2010;283:185–96.

Mavrogianopoulos G, Vogli V, Kyritsis S. Use of wastewater as a nutrient solution in a closed gravel hydroponic culture of giant reed (Arundo donax). Bioresour Technol. 2002;82:103–7.

Michel BE, Kaufmann MR. The osmotic potential of polyethylene glycol 6000. Plant Physiol. 1973;51:914–6.

Mirzaei M, Soltani N, Sarhadi E, George IS, Neilson KA, Pascovici D, Shahbazian S, Haynes PA, Atwell BJ, Salekdeh GH. Manipulating root water supply elicits major shifts in the shoot proteome. J Proteome Res. 2014;13:517–26.

Mohammadkhani N, Heidari R. Drought-induced accumulation of soluble sugars and proline in two maize varieties. World Appl Sci J. 2008;3:448–53.

Moreno-Hagelsieb G, Latimer K. Choosing BLAST options for better detection of orthologs as reciprocal best hits. Bioinformatics. 2008;24:319–24.

Moumeni A, Satoh K, Kondoh H, Asano T, Hosaka A, Venuprasad R, Serraj R, Kumar A, Leung H, Kikuchi S. Comparative analysis of root transcriptome profiles of two pairs of drought-tolerant and susceptible rice near-isogenic lines under different drought stress. BMC Plant Biol. 2011;11:174.

Mustilli A-C, Merlot S, Vavasseur A, Fenzi F, Giraudat J. Arabidopsis OST1 protein kinase mediates the regulation of stomatal aperture by abscisic acid and acts upstream of reactive oxygen species production. Plant Cell. 2002;14:3089–99.

Nakano T, Suzuki K, Fujimura T, Shinshi H. Genome-wide analysis of the ERF gene family in Arabidopsis and rice. Plant Physiol. 2006;140:411–32.

Narsai R, Wang C, Chen J, Wu J, Shou H, Whelan J. Antagonistic, overlapping and distinct responses to biotic stress in rice (Oryza sativa) and interactions with abiotic stress. BMC Genomics. 2013;14:93.

Nijhawan A, Jain M, Tyagi AK, Khurana JP. Genomic survey and gene expression analysis of the basic leucine zipper transcription factor family in rice. Plant Physiol. 2008;146:333–50.

Nuruzzaman M, Sharoni AM, Kikuchi S. Roles of NAC transcription factors in the regulation of biotic and abiotic stress responses in plants. Front Microbiol. 2013;4:248.

Ohme-Takagi M, Shinshi H. Ethylene-inducible DNA binding proteins that interact with an ethylene-responsive element. Plant Cell. 1995;7:173–82.

Overvoorde P, Fukaki H, Beeckman T. Auxin control of root development. Cold Spring Harb Perspect Biol. 2010;2:a001537.

Perdue R. Arundo donax—source of musical reeds and industrial cellulose. Econ Bot. 1958;12:368–404.

Pilu R, Cassani E, Landoni M, Badone FC, Passera A, Cantaluppi E, Corno L, Adani F. Genetic characterization of an Italian Giant Reed (Arundo donax L.) clones collection: exploiting clonal selection. Euphytica. 2013;196:169–81.

Pinheiro C, António C, Ortuño MF, Dobrev PI, Hartung W, Thomas-Oates J, Ricardo CP, Vanková R, Chaves MM, Wilson JC. Initial water deficit effects on Lupinus albus photosynthetic performance, carbon metabolism, and hormonal balance: metabolic reorganization prior to early stress responses. J Exp Bot. 2011;62:4965–74.

Polunin O, Huxley A. Flowers of the Mediterranean. London: Hogarth Press; 1987.

Priya P, Jain M. RiceSRTFDB: a database of rice transcription factors containing comprehensive expression, cis-regulatory element and mutant information to facilitate gene function analysis. Database (Oxford). 2013;2013:bat027.

Priyam A, Woodcroft BJ, Rai V & Wurm Y (in prep) SequenceServer: BLAST search made easy. http://sequenceserver.com

Qi X, Xie S, Liu Y, Yi F, Yu J. Genome-wide annotation of genes and noncoding RNAs of foxtail millet in response to simulated drought stress by deep sequencing. Plant Mol Biol. 2013;83:459–73.

Rashid M, Guangyuan H, Guangxiao Y, Hussain J, Xu Y. AP2/ERF Transcription Factor in Rice: Genome-Wide Canvas and Syntenic Relationships between Monocots and Eudicots. Evol Bioinform Online. 2012;8:321–55.

Ray S, Dansana PK, Giri J, Deveshwar P, Arora R, Agarwal P, Khurana JP, Kapoor S, Tyagi AK. Modulation of transcription factor and metabolic pathway genes in response to water-deficit stress in rice. Funct Integr Genomics. 2011;11:157–78.

Riechmann JL, Meyerowitz EM. The AP2/EREBP family of plant transcription factors. Biol Chem. 1998;379:633–46.

Robinson MD, McCarthy DJ and Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26:139-40.

Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol. 2010;11:R25.

Sablok G, Fu Y, Bobbio V, Laura M, Rotino GL, Bagnaresi P, Allavena A, Velikova V, Viola R, Loreto F, Li M, Varotto C. Fuelling genetic and metabolic exploration of C3 bioenergy crops through the first reference transcriptome of Arundo donax L. Plant Biotechnol J. 2014;12:554–67.

Sakuma Y, Liu Q, Dubouzet JG, Abe H, Shinozaki K, Yamaguchi-Shinozaki K. DNA-binding specificity of the ERF/AP2 domain of Arabidopsis DREBs, transcription factors involved in dehydration- and cold-inducible gene expression. Biochem Biophys Res Commun. 2002;290:998–1009.

Seo JS, Joo J, Kim MJ, Kim YK, Nahm BH, Song SI, Cheong J-J, Lee JS, Kim J-K, Choi Y Do. OsbHLH148, a basic helix-loop-helix protein, interacts with OsJAZ proteins in a jasmonate signaling pathway leading to drought tolerance in rice. Plant J. 2011;65:907–21.

Shanker AK, Maheswari M, Yadav SK, Desai S, Bhanu D, Attal NB, Venkateswarlu B. Drought stress responses in crops. Funct Integr Genomics. 2014;14:11–22.

Smyth GK, Verbyla AP. A conditional likelihood approach to REML in Generalized Linear Models. J R Stat Soc Ser B. 1996;58:565–72.

Soon FF, Ng LM, Zhou XE, West GM, Kovach A, Tan MH, Suino-Powell KM, He Y, Xu Y, Chalmers MJ, Brunzelle JS, Zhang H, Yang H, Jiang H, Li J, Yong EL, Cutler S, Zhu JK, Griffin PR, Melcher K, Xu HE. Molecular mimicry regulates ABA signaling by SnRK2 kinases and PP2C phosphatases. Science 2012;335:85-8.

Sperotto RA, Ricachenevsky FK, Duarte GL, Boff T, Lopes KL, Sperb ER, Grusak MA, Fett JP. Identification of up-regulated genes in flag leaves during rice grain filling and characterization of OsNAC5, a new ABA-dependent transcription factor. Planta. 2009;230:985–1002.

Supek F, Bosnjak M, Skunca N, Smuc T. REVIGO summarizes and visualizes long lists of Gene Ontology terms. PLoS One. 2011;6: e21800.

Taji T, Ohsumi C, Iuchi S, Seki M, Kasuga M, Kobayashi M, Yamaguchi-Shinozaki K, Shinozaki K. Important roles of drought- and cold-inducible genes for galactinol synthase in stress tolerance in Arabidopsis thaliana. Plant J. 2002;29:417–26.

Umezawa T, Nakashima K, Miyakawa T, Kuromori T, Tanokura M, Shinozaki K, Yamaguchi-Shinozaki K. Molecular basis of the core regulatory network in ABA responses: sensing, signaling and transport. Plant Cell Physiol. 2010;51:1821–39.

Valdés AE, Övernäs E, Johansson H, Rada-Iglesias A, Engström P. The homeodomain-leucine zipper (HD-Zip) class I transcription factors ATHB7 and ATHB12 modulate abscisic acid signalling by regulating protein phosphatase 2C and abscisic acid receptor gene activities. Plant Mol Biol. 2012;80:405–18.

Valliyodan B, Nguyen HT. Understanding regulatory networks and engineering for enhanced drought tolerance in plants. Curr Opin Plant Biol. 2006;9:189–95.

Verbruggen N, Hermans C. Proline accumulation in plants: A review. Amino Acids. 2008;35:753–9.

Verslues PE, Agarwal M, Katiyar-Agarwal S, Zhu J, Zhu JK. Methods and concepts in quantifying resistance to drought, salt and freezing, abiotic stresses that affect plant water status. Plant J. 2006;45:523–39.

Virlouvet L, Fromm M. Physiological and transcriptional memory in guard cells during repetitive dehydration stress. New Phytol. 2015;205:596–607.

Wang D, Pan Y, Zhao X, Zhu L, Fu B, Li Z. Genome-wide temporal-spatial gene expression profiling of drought responsiveness in rice. BMC Genomics. 2011;12:149.

Wang Q, Guan Y, Wu Y, Chen H, Chen F, Chu C. Overexpression of a rice OsDREB1F gene increases salt, drought, and low temperature tolerance in both Arabidopsis and rice. Plant Mol Biol. 2008;67:589–602.

Wrzaczek M, Brosché M, Salojärvi J, Kangasjärvi S, Idänheimo N, Mersmann S, Robatzek S, Karpiński S, Karpińska B, Kangasjärvi J. Transcriptional regulation of the CRK/DUF26 group of receptor-like protein kinases by ozone and plant hormones in Arabidopsis. BMC Plant Biol. 2010;10:95.

Xiang Y, Tang N, Du H, Ye H, Xiong L. Characterization of OsbZIP23 as a key player of the basic leucine zipper transcription factor family for conferring abscisic acid sensitivity and salinity and drought tolerance in rice. Plant Physiol. 2008;148:1938–52.

Xie C, Mao X, Huang J, Ding Y, Wu J, Dong S, Kong L, Gao G, Li CY, Wei L. KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. Nucleic Acids Res. 2011;39:W316–22.

Xie F, Stewart CN, Taki F a, He Q, Liu H, Zhang B. High-throughput deep sequencing shows that microRNAs play important roles in switchgrass responses to drought and salinity stress. Plant Biotechnol J. 2014;12:354–66.

Xu K, Chen S, Li T, Ma X, Liang X, Ding X, Liu H, Luo L. OsGRAS23, a rice GRAS transcription factor gene, is involved in drought stress response through regulating expression of stress-responsive genes. BMC Plant Biol. 2015;15:141.

Yadav NR, Taunk J, Rani A, Aneja B, Yadav RC. Role of Transcription Factors in Abiotic Stress Tolerance in Crop Plants. In: Tuteja N, Gill SS, editors. Climate change and plant abiotic stress tolerance. Weinheim: Wiley-VCH Verlag GmbH & Co. KGaA; 2013..

Yoshiba Y, Kiyosue T, Nakashima K, Yamaguchi- Shinozki K, Shinozaki K. Regulation of levels of proline as an osmolytes in plants under water stress. Plant Cell Physiol.1997;38:1095-102.

Yoshida R, Hobo T, Ichimura K, Mizoguchi T, Takahashi F, Aronso J, Ecker JR, Shinozaki K. ABA-activated SnRK2 protein kinase is required for dehydration stress signaling in Arabidopsis. Plant Cell Physiol. 2002;43:1473–83.

Yun KY, Park MR, Mohanty B, Herath V, Xu F, Mauleon R, Wijaya E, Bajic VB, Bruskiewich R, de Los Reyes BG. Transcriptional regulatory network triggered by oxidative signals configures the early response mechanisms of japonica rice to chilling stress. BMC Plant Biol. 2010;10:16.

Zhang L, Yu S, Zuo K, Luo L, Tang K. Identification of gene modules associated with drought response in rice by network-based analysis. PLoS One. 2012;7:e33748.

Zhang M, Lv D, Ge P, Bian Y, Chen G, Zhu G, Li X, Yan Y. Phosphoproteome analysis reveals new drought response and defense mechanisms of seedling leaves in bread wheat (Triticum aestivum L.). J Proteomics. 2014;109:290–308.

Zhang Q. Strategies for developing Green Super Rice. Proc Natl Acad Sci U S A. 2007;104:16402–9.

Zheng B, Halperin T, Hruskova-Heidingsfeldova O, Adam Z, Clarke AK. Characterization of Chloroplast Clp proteins in Arabidopsis: Localization, tissue specificity and stress responses. Physiol Plant. 2002;114:92–101.

Chapter 3.

Characterization of a stress-related Arundo gene.

The *Arundo donax* L. *AdDWD1* gene, belonging to the WD40 family, is involved in osmotic and salt stress response.

Michele Poli[1,2], Silvio Salvi[2], Mingai Li[1§], Claudio Varotto[1§]

Affiliation:

[1] Department of Biodiversity and Molecular Ecology, Research and Innovation Centre, Fondazione Edmund Mach, Via E. Mach 1, 38010 S. Michele all'Adige (TN), Italy.

[2] Department of Agricultural Sciences, University of Bologna, Bologna, Italy

3.1 Introduction

Abiotic stresses are among the main hazards for plant growth and survival. Especially drought and high salinity are responsible for productivity and quality loss of crop plants worldwide. Therefore, the improvement of crop species is a key factor to guarantee an adequate yield also in light of the climate change (Tester and Langridge, 2010). To cope with drought and salt stresses, plants activate molecular, chemical and physiological responses that allow them to survive or adapt to new environmental conditions (Osakabe et al., 2014; Tuteja and Sopory, 2008). At molecular level the early response to these stresses through regulation of the cell osmotic potential is particularly important to avoid permanent damages of the plant cells (Adamec, 1984). In the last decades, many genes have been found to have a role in defence mechanisms which involved several different pathways. The main pathways that have been recognized to participate in drought and/or salinity response include ABA-dependent, Salt Overly Sensitive (SOS) and dehydration-responsive

element-binding protein (DREB) mediated pathways (Ji et al., 2013; Yoshida et al., 2014). In particular, some genes are considered to be key regulators in the respective pathways: AREB/ABFs (ABRE- binding protein/ABRE-binding factors) in ABA-dependent, DREB2A (Dehydration-responsive element-binding protein 2A) in ABA-independent (Yoshida et al., 2014) and SOS1-3 in SOS pathway.

WD40 proteins contain 4 to 8 repeats of 44-60 amino acid typically ending with Tryptophan-Aspartate dipeptide (WD) and have been found in many organisms, from bacteria to animals and plants (Stirnimann et al., 2010). Structurally, these proteins consist of a four-stranded anti-parallel β-sheet that forms each of the seven blades composing the β-propeller fold that is usually considered to facilitate protein-protein interaction. To date, in plants 237, 200 and 225 potential WD40 genes have been identified in Arabidopsis (van Nocker and Ludwig, 2003), rice (Ouyang et al., 2012) and foxtail millet (Mishra et al., 2014), respectively. The high number of these genes indicates an important role in the plant system. Despite many progresses have been already made towards elucidation of WD40 gene functions, many members of this large family remain uncharacterized. WD40 proteins have been found to function in cytoplasm and nucleoplasm and they play important roles in several biological processes like signal transduction, regulation of cell division, chromatin modification, transcription regulation, plant immunity and photomorphogenesis (Miller et al., 2015; Schroeder et al., 2016; van Nocker and Ludwig, 2003). DWD (DDB1 binding WD40) proteins are a subgroup of the WD40 superfamily recognizable from a peculiar 16 amino acid sequence at the end of one repeat consisting of several highly conserved residues: Asp/Glu_{7}, Trp/Tyr_{13}, Asp/Glu_{14}, Arg/Lys_{14} (Lee et al., 2008). Several DWD genes have been implicated in abiotic stress response or environmental adaptation. For example, in Arabidopsis HTD1 (heat stress tolerant DWD1) protein is involved in modulation of the thermotolerance process (Kim et al., 2014) while ABD1 (ABA-hypersensitive DCAF1) targets ABI5 (ABA-insensitive 5) for degradation, therefore regulating ABA-mediated stress response (Seo et al., 2014); in rice five WD40 genes (SRWD1-5) have been characterized to be responsive to salt stress (Huang et al., 2008); in foxtail millet SiWD40 is induced by various abiotic stresses and may be subjected to regulation by DRE (dehydration-responsive elements) genes (Mishra et al., 2012a); in *Triticum aestivum* TaWD40D is involved either in ABA-dependent (through ABI2 regulation) and ABA-independent stress response pathway (Kong et al., 2014). The cullin-based E3 ligases are complexes that serve as a scaffold for assembling the ubiquitination machinery. It as been demonstrated that Cullin 4 utilizes damaged DNA binding 1 (DDB1) protein as an adaptor to assemble the E3 ligase complex (Lee and Kim, 2011). Strong evidence also suggests that DWD proteins bind to the DDB1

adaptor by means of the WDxR motif (Angers et al., 2006) as substrate receptor, thus committing the whole complex for degradation via the 26S proteasome.

Arundo donax L. (also known as giant reed) is a perennial bamboo-like species of the Poaceae family with the potential to play a key role in bioenergy production (Pilu et al., 2012). Its capability to cope with environmental stresses such as drought and salt and its high yield (up to 40 t/ha of dry biomass per year) make it a suitable candidate as non-food energy crop. The agronomical and biological features of *A. donax* have been extensively studied (Mantineo et al., 2009; Nassi o Di Nasso et al., 2010; Pilu et al., 2013b) either for its exploitation in the Mediterranean region where it is native (Hardion et al., 2015) or for its eradication where it is allochthonous (e.g. United States and Australia) (Quinn and Holt, 2007). Until recent years the only genetic studies carried out on *A. donax* comprise phylogenesis (Hardion et al., 2015, 2012) and genetic variation (Pilu et al., 2013a). Recently, though, with the advent of new generation sequencing platforms, it became possible to quickly and cheaply produce extensive genetic resources for virtually any living species, which could support deeper and more specific trait-related analyses. Thus, the whole giant reed transcriptome was released first in 2014 for four organs (Sablok et al., 2014) and subsequently also for young shoots subjected to osmotic stress (Fu et al., 2016). This precious data not only already shed novel light on the genetic response of giant reed to external stimuli, but they are also an important step to select possible key regulator genes from this non-model species for further characterization.

In this study, we select from the *A. donax* water-stressed transcriptome a putative DWD gene involved in osmotic stress response. Quantitative real-time PCR demonstrated a high upregulation both in root and shoot of young giant reed plants grown in hydroponic solution and subjected to osmotic and salt stress. Moreover, the heterologous overexpression of *AdDWD1* shows that *Arabidopsis* seedlings are more sensitive than wild type plants when grown in salt and ABA enriched media. Analysis of 7 known stress-related genes belonging to ABA-dependent, independent and SOS pathways have been checked in transgenic lines, showing a downregulation of DREB2A and SOS3 transcripts . Our results suggest that *AdDWD1* may act as negative regulator of the ABA-independent pathway or may interfere with the function of the CUL4-DDB1 complex.

3.2 Material and Methods

Plant growth and abiotic stress treatments.

Cohorts of *A. donax* cuttings (Sesto Fiorentino, Florence, Italy 43°49'01.8"N 11°11'57.0"E) were grown in a growth chamber with day-length of 16 hrs, light intensity of 150 µmol of photons m^{-2} s^{-1}, 24°C and 60% relative humidity in tap water until 2-leaves stage. The cuttings were then moved to hydroponic solution (renewed every 3 days) and harvested at the 3-leaves stage. The hydroponic solution contained: KNO_3 (6mM), $Ca(NO_3)_2*4H_2O$ (3.5 mM), $NH_4H_2PO_4$ (1mM), $MgSO_4*7H_2O$ (2mM), H_3BO_3 (46.3 µM), $MnSO_4*H_2O$ (12.6 µM), $ZnSO_4*7H_2O$ (0.8 µM), $CuSO_4*5H_2O$ (3 µM), $H_{24}Mo_7N_6O_{24}*4H_2O$ (1 µM), Fe-EDTA (0.0817 mM). For stress treatments, plants were moved from hydroponic solution to fresh solution containing 15% PEG 6000 (osmotic stress), 150 mM NaCl (salt stress), 500 µM $CdSO_4$ (heavy metal stress) or exposed at 4°C in the basic hydroponic solution (cold stress). Sample collection was carried out after 0h (control), 30 min, 1h 30min, 3h, 6h, 11h and 24h. Three biological replicates for each condition were taken, dividing the root and the shoot, snap- frozen in liquid nitrogen and stored at -80 °C until RNA extraction. Five replicas of water loss assay were carried out with 15 leaves each detached from three weeks old plants, let air-dry on the growth-chamber shelf and weighted at time 0, 30 min, 45 min and every hour until 8 time-points were collected. Dry weight was measured after dessication of the samples for 24 hours in oven at 60 °C.

Phylogenetic tree reconstruction.

Protein sequences of rice, sorghum, maize, purple false brome, foxtail millet and switchgrass species were retrieved from Phytozome and reciprocal BLAST hit were used to detect putative orthologs. The Gblocks software (http://molevol.cmima.csic.es/castresana/Gblocks.html) was used to remove from the sequences highly variable regions and CLUSTLW was employed for sequences alignment. The final alignment encompassed a total of 21 amino acid sequences. All positions with less than 80% site coverage were eliminated so that fewer than 20% alignment gaps, missing data, and ambiguous bases were allowed at any position. There were a total of 337 positions in the final dataset. Phylogenetic reconstruction was carried out with the Maximum Likelihood (ML) algorithm implemented in MEGA7 (Kumar et al., 2016). A total of 1000 bootstrap replicates were used to assess reliability of the topology of the resulting tree.

Cloning of *AdDWD1* and creation of transgenic lines.

The gene was selected from the transcriptome analysis of *Arundo donax* under water stress (Fu et al., 2016) and amplified from shoot cDNA with Phusion High-Fidelity DNA. The product has been run on 1% agarose gel, cut and purified with NucleoSpin® Gel and PCR Clean-up (Macherey-Nagel). After quantification, the isolated fragment was cloned using pENTR™/D-

TOPO® Cloning Kit in *E. coli* competent cells by electroporation, following the manufacturer's instructions. The donor vector was recombined with destination vector *pk7WG2* by GATEWAY™ conversion technology (Invitrogen). Transgenic *Arabidopsis* plants carrying the construct under the control of cauliflower mosaic virus 35S promoter (35S::AdDWD1) were then obtain by floral dip method (Clough and Bent, 1999) in a *Agrobacterium tumefaciens* solution and selected on plates with specific antibiotic. The *Arabidopsis* homolog (*At1g78080*) was used to create a construct carrying GUS gene fused with its promoter (1853 bases) in similar way: the region, obtained from TAIR database, was amplified and used in pENTR™/D-TOPO® Cloning system with *pkGWFS7* plasmid as destination vector. *Arabidopsis* plants were then brought to T2 generation for subsequent analysis. Cloning primers are listed in Supplementary Material (Supplementary File 3.5).

RNA extraction, qRT-PCR and gene expression analysis.

About 100 mg of tissues was used for total RNA isolation with Spectrum Plant Total RNA Extraction Kit (Sigma) and Rneasy® Plant Mini Kit (Qiagen) for *Arundo donax* shoot and roots respectively. For stress-related gene experiment, RNA from *Arabidopsis* seedlings was obtained by TRIzol® reagent (Thermo Fisher) following the manufacturer's protocol. RNA was treated with DNase I (Sigma-Aldrich) and 1 µL was run on 1% agarose gel for integrity control. The quality of the RNA was also assessed by spectrophotometer and an absorption of OD260/OD280 between 1.9 and 2.2 was considered as parameter for high quality RNA. Single strand cDNA was reversed transcribed from 1 µg of RNA primed with oligodT primers in a reaction mixture of total 20 µL using SuperScript® III Reverse Transcriptase (Life Technologies). Total 10 µL PCR mixture containing 0,5 µL of 5-time diluted cDNA template, 1 µL Buffer 10X, 1 µL dNTPs (1 mM), 250 nM of each primer and 0.2 µL of *Taq* polymerase was used to check specific amplification and product length on electrophoresis run with 2% agarose gel. PCR run was set as follow: 5 min at 95°C, 30 cycles of 40s at 94 °C, 30s at 60 °C and 20s at 72 °C, with 5 min final extension at 72 °C. Mixture of qRT-PCR contained 1 µL of 5-fold diluted cDNA, 200 nM of each primer and 6.25 µL of Platinum® SYBR® Green qPCR SuperMix-UDG (Invitrogen) in a 12.5 µL total volume. Run for qRT-PCR in Bio-Rad C1000 Thermal Cycler was set as: 2 min at 50 °C, 2 min at 95 °C, 40 cycles of 15s at 95 °C and 30s at 60 °C. Moreover, a melt curve was created by constantly rising the temperature from 65 °C to 90 °C to control PCR product specificity. Each sample has been run in three technical replicas and every plate has one No Template Control (NTC) well for each primer pair used. Primers are listed in Supplementary File 3.5.

Stress assays.

Salt assay with *35S::AdDWD1* overexpressing line was made by sowing 77 seeds of each genotype of *Arabidopsis* (wild type, OE1 and OE2) on MS plate, MS plate + 175mM of NaCl or MS plate + 0,5 µM ABA. Seed germination of three independent experiments (total 231 seeds per line) was recorded until the 5^{th} day after three days of stratification in dark at 4°C. Growing chamber conditions were the same as reported above for giant reed propagation. For stress-related genes experiments, one-week-old seedlings grown in hydroponic solution were moved either to a fresh solution with 150 mM of salt (treatment) or without (control) for 4 hours before collection of 10-14 plants for each of the three biological replicas. The hydroponic medium contained: KNO_3 (5,1 mM), $Ca(NO_3)_2$ * $4H_2O$ (1,01 mM), $NH_4H_2PO_4$ (0,13 mM), $MgSO_4$ * $7H_2O$ (0,498 mM), NaOH (8,89 µM), H_3BO_3 (9,68 µM), $MnCl_2$ * $4H_2O$ (2,03 µM), $ZnSO_4$ * $7H_2O$ (0,314 µM), $CuSO_4$ * $5H_2O$ (0,21 µM), MoO_3 (0,139 µM), $Co(NO_3)_2$ * $6H_2O$ (0,086 µM), NH_4NO_3 (29,3 µM), NaFe(III)EDTA (22,4 µM).

3.3 Results

Cloning, sequence analysis and phylogenetic tree.

After a screening of the differentially expressed genes from recently published *Arundo donax* L. transcriptome (Fu et al., 2016), Unigene032775 was selected to be characterized for its function under abiotic stress conditions in the model plant *Arabidopsis thaliana*. In fact, a primary evaluation of its transcript showed an interesting upregulated pattern both in root and shoot but a research in published papers and open database (GeneBank, Uniprot, Phytozome) gives very little information about it making it a good candidate to be studied. After PCR amplification and confirmation by sequencing, the isolated transcript resulted 1711 bases long. The longest CDS it encoded a 476 amino acid protein with an estimated molecular mass of 52,64 kDa and pI 5,55 A. The overxpression construct 35S::AdDWD1 was transformed in *Arabidopsis* plant and 25 T2 generation lines were checked for transgene expression levels (Supplementary. File 3.1). The two homozygous T3 lines with highest transgene expression (from now indicated as OE1 and OE2) were used for further gene characterization.

The closest homologue of AdDWD1 in Arabidopsis thaliana was identified by BLASTp-based homology searches against the curated TAIR *Arabidopsis* proteome (https://www.arabidopsis.org/Blast/index.jsp). The best hit (65% identities, 80% positives, E value = e-163, Supplementary File 3.6) was *At1g78070*, a member of the transducin/WD40 repeat-like superfamily. A domain search carried out with the SMART software (http://smart.embl.de/)

confirmed the presence of three WD40 domains in *AdDWD1* and its ortholog (Figure 3.1A). However, the presence of a typical DWD-box at the end of the second repeat (residues position 362-377), which make this protein a candidate DCAF (Figure 3.1B) (Biedermann and Hellmann, 2011). Although SMART software recognizes only three WD-40 repeats, it is known that often one or more WD40 repeats remain undetected even if the correct domain is assign to the protein (Stirnimann et al., 2010). For this reason, we employed also other two online tools: the WDSP software (Wang et al., 2013), that predicts 7 repeats from position 154 to 463 and WRDD (Wang et al., 2016), that predicts 6 repeats (Supplementary File 3.2).

To understand the evolutionary history of *AdDWD1*, the transcriptome of giant reed, rice, sorghum, maize, purple false brome, foxtail millet and switchgrass were compared by Reciprocal BLAST Hit method (RBH), with which we selected the 21 most closely related sequences, all containing the DWD-box. As expected by the low level of curation of the database, the Phytozome annotation of these sequences varied among species: "WD40 repeat" for sorghum, maize and foxtail millet proteins, "nucleoside-triphosphatase activity" for *Brachypodium distachyon* and no annotation for *Panicum virgatum*. Phylogenetic reconstruction shows a relatively well resolved and supported topology (Figure 3.1C). *AdDWD1* forms a monophyletic clade with rice *Os03g26870.1*, purple false brome protein *Bradi1g60960.*2, sorghum *Sb01g033820.1*, maize *GRMZM2G078806* and two switchgrass proteins *Pavir.Ia03143.1.p* and *Pavir.Ib01829.1.p* (Clade A1; Fig. 5). The phylogenetic positioning between *A. donax* and the other monocot species included in the analysis are consistent in all the clades identified and reflect the known phylogeny of the Poaceae (Soreng et al., 2015). Worth of note, all orthologs of *AdDWD1* are present in single copy in all species with the exception of *P. virgatum*, where two paralogues were identified, and of foxtail millet, where no homologue was detected. Despite not supported by high bootstrap values, the closest homologue of *AdDWD1* is *A. donax* Unigene053546 (Clade A2), while the paralogous *A. donax* genes Unigene_064478 and Unigene064479 belong to Clade B (Fig. 5).

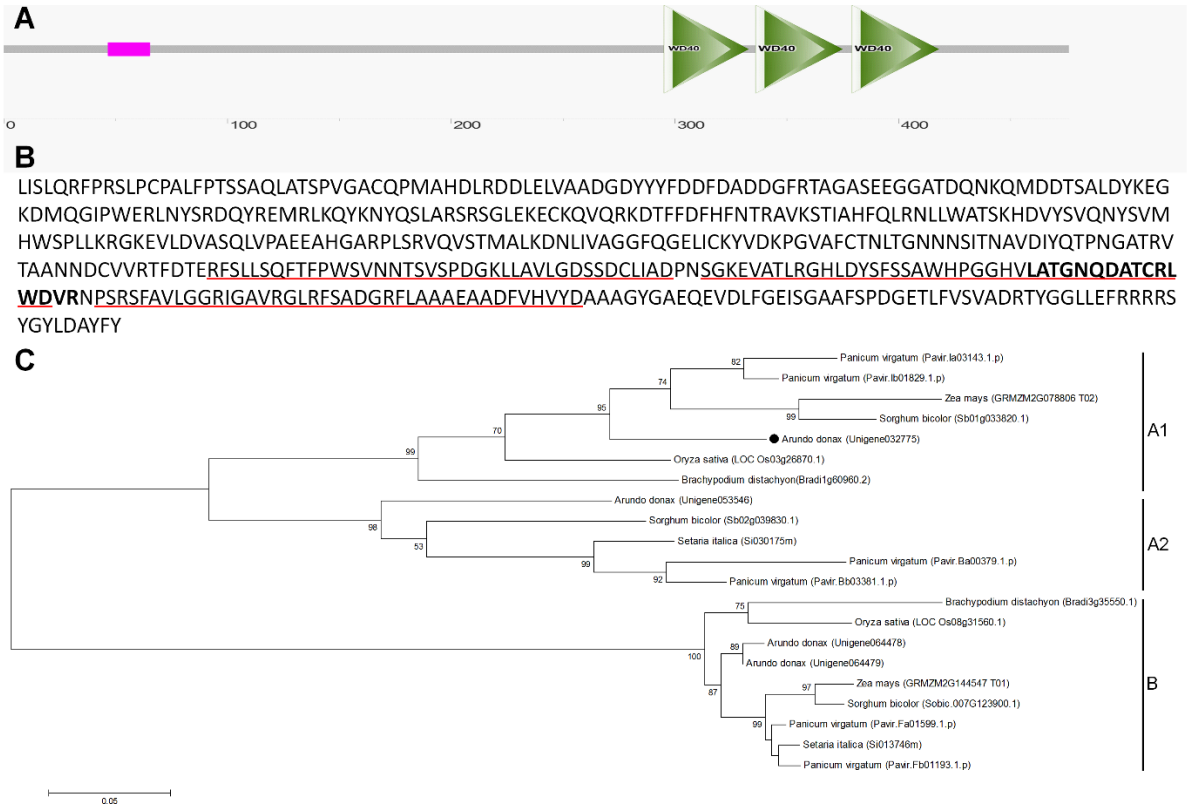


Figure 3.1. **Gene features**. A) SMART domain prediction of *AdDWD1* shows a low complexity region (pink bar) and three WD40 repeats (green arrows). B) Amino acid sequence of *AdDWD1* with SMART WD40 repeats underlined and the 16 amino acid of DWD motif in bold style. C) Maximum Likelihood unrooted phylogenetic tree of *AdDWD1* and its homologues from fully-sequenced monocot species. The tree is drawn to scale, with branch lengths measured as number of substitutions per site. Bootstrap values above 50 are shown at the nodes.

Expression pattern of *AdDWD1*.

We examined *AdDWD1* expression pattern under different stress conditions and in two separate tissues: root and shoot. Figure 3.2 shows that osmotic and salt stresses induce the highest upregulation of this transcript, especially in roots, where the fold-change compared to time 0 is above 400 times and about 300, respectively. As expected, a delay in shoot response is visible from the graph, mainly when osmotic stress was applied. Interestingly, the kinetic of the expression is faster in salt stress where the highest level is reached after 1 hour 30 minutes of treatment, while under osmotic stress the maximum level is reached at 3 hours in root and at 11 hours in shoot. In the treatment with heavy metal and cold, *AdDWD1* was observed to be upregulated only in root, even though the expression value is much lower compared with the other two stresses. On the other hand, in shoot there is a downregulation of the gene in the early stages, with a recover of the expression level from 3 hours on.

To investigate the expression under normal condition in adult plants, semi-quantitative RT-PCR has been used to quantify the presence of *AdDWD1*. As reported in Fig2 I, the expression

levels after 34 cycles of PCR are higher in developed leaves (L1, L4), lower in not emerged leaf (L0), stem (node NO, internode, IN) and bud (MS) and almost absent in the sheath (GS). Moreover, when compared to the reference gene GAPDH it is evident that absolute quantity of the transcript in the lack of stress is low even in the most highly expressing tissues.

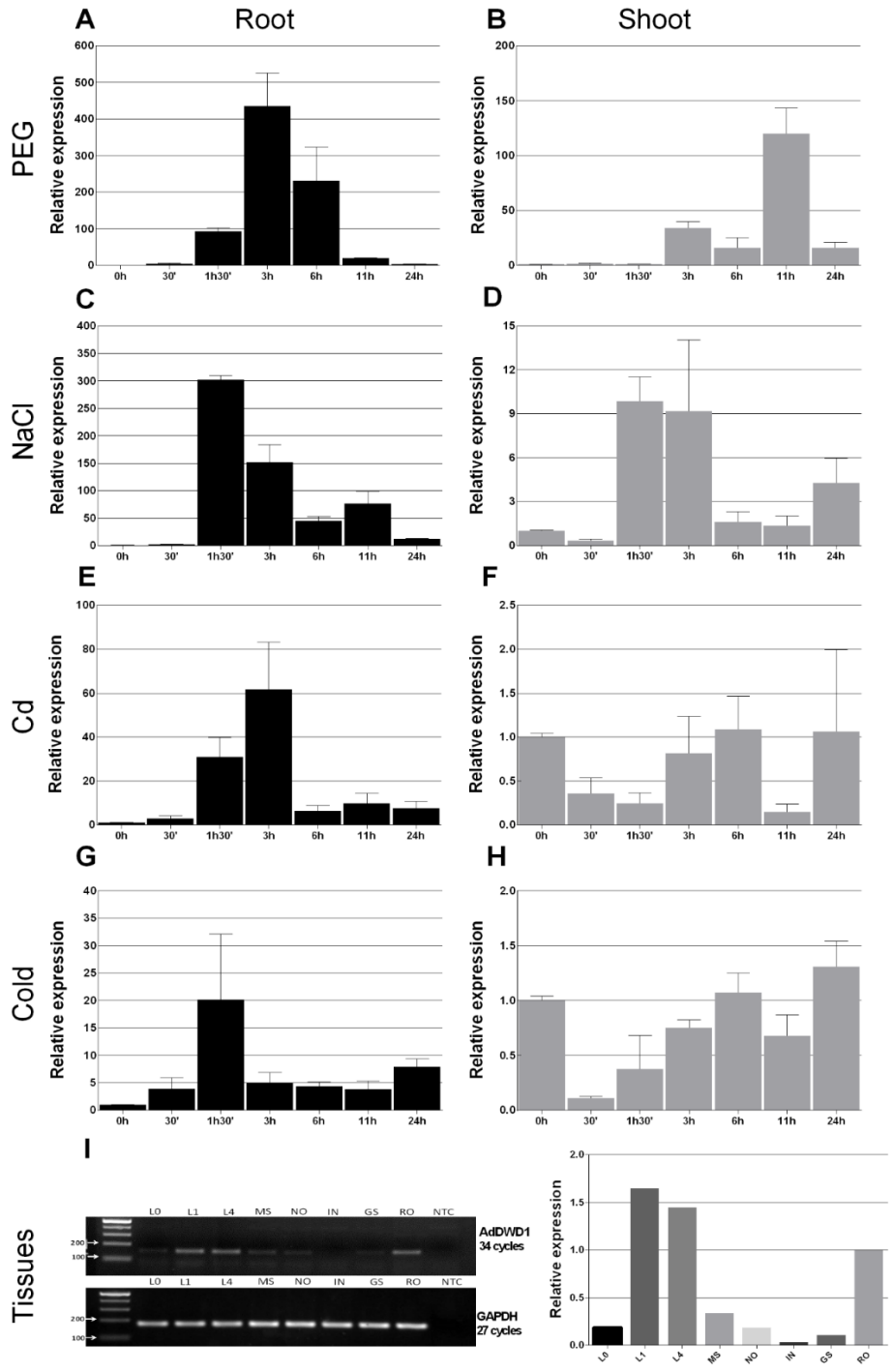


Figure 3.2. **Expression pattern of *AdDWD1* gene under abiotic stress conditions**: osmotic root (A) and shoot (B), salt root (C) and shoot (D), heavy metal root (E) and shoot (F), cold root (G) and shoot (H). Graph (I) is the expression pattern in different tissues: L0, Apical not opened leaf; L1, first fully open leaf from top; L4, fourth leaf from top (middle part); MS, meristem shoot (bud); NO, node of the fourth leaf; IN, internode between 4th and 3th leaf; GS, green sheath; RO, root; NTC, no template control.

AdDWD1 reduces germination and growth under stress in overexpressing lines.

The involvement of AdDWD1 in drought and salt stresses indicated by the expression analyses does not provide information on its actual mechanism of action. We, therefore, carried out the physiological characterization of Arabidopsis *AdDWD1* overexpressing lines. Germination assay in normal (MS) and stressed conditions showed high difference in stress response between wild type (WT) plants and *35S::AdDWD1* lines (OE1 and OE2) (Figure 3.3). When 175 mM of NaCl was added to the medium, the percentage of OE lines germination was significantly lower than that of WT (max 38% versus 70% at 5 days after stratification, respectively; Figure 3.3B, D). Moreover, when the seeds were placed on agar supplemented with 0,5 µM ABA, we noticed a sensitive repression in cotyledon emergence from the third day on (about 40% of the WT against 20% of the OE lines at the fifth day; Figure 3.3C, E). In the light of this results, root growth of young seedlings was also checked with salt and ABA additions but no significant differences with WT plants were detected (Supplementary File 3.3). Also, a water loss assay carried out on detached leaves showed a higher tendency of OE lines to loose water than WT leaves. This became significant at 3 hours and increase up to 6 hours after leaf detachment (Figure 3.3F). On the other hand, when adult plants were exposed to air drying no differences in survival rate were notice between WT and OE (Supplementary File 3.4).

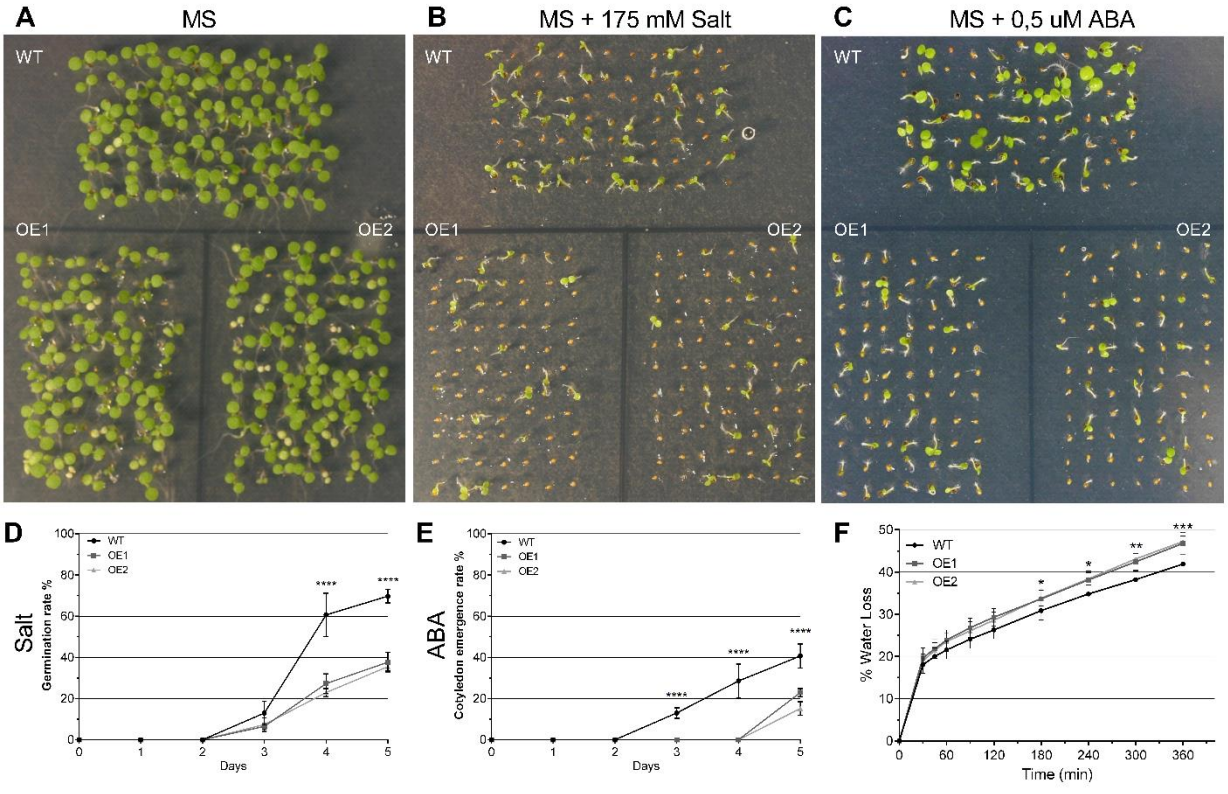


Figure 3.3. **Physiological studies of OE1 and OE2**. Germination assay in MS (A), MS supplemented with 175 mM of NaCl (B) and 0,5 µM ABA (C). The graph shows the percentage of germination rate under salt (D) and cotyledon

emergence under ABA (E) at 5 days for wild type and two 35S::AdDWD1 lines. Experiments were conducted in triplicates. Water loss (F) was calculated by the formula W0-Wt/W0 where W is normalized by final dry weight. One, two or three asterisks indicate P value minor to 0,5, 0,01 and 0,001 respectively. Significance was calculated with 2-way ANOVA with Dunnet correction for multiple comparisons. The experiment consisted of five replicas.

The effect of *AdDWD1* overexpression was also studied by checking seven different stress-related genes: dehydration inducible genes *DREB2A* (Dehydration-responsive element-binding protein 2A) and *RD29A*, salt responsive genes *SOS1*, *SOS3* and *P5CS* and ABA responsive genes *ABI1* and *COR15A* (also drought responsive). Results show that under salt stress, the selected genes are mainly upregulated in WT and overexpressing lines except for SOS3. Only the *DREB2A* and *SOS3* genes expression levels showed a statistically significant reduction of expression in both OE lines compared to WT (Figure 3.4).

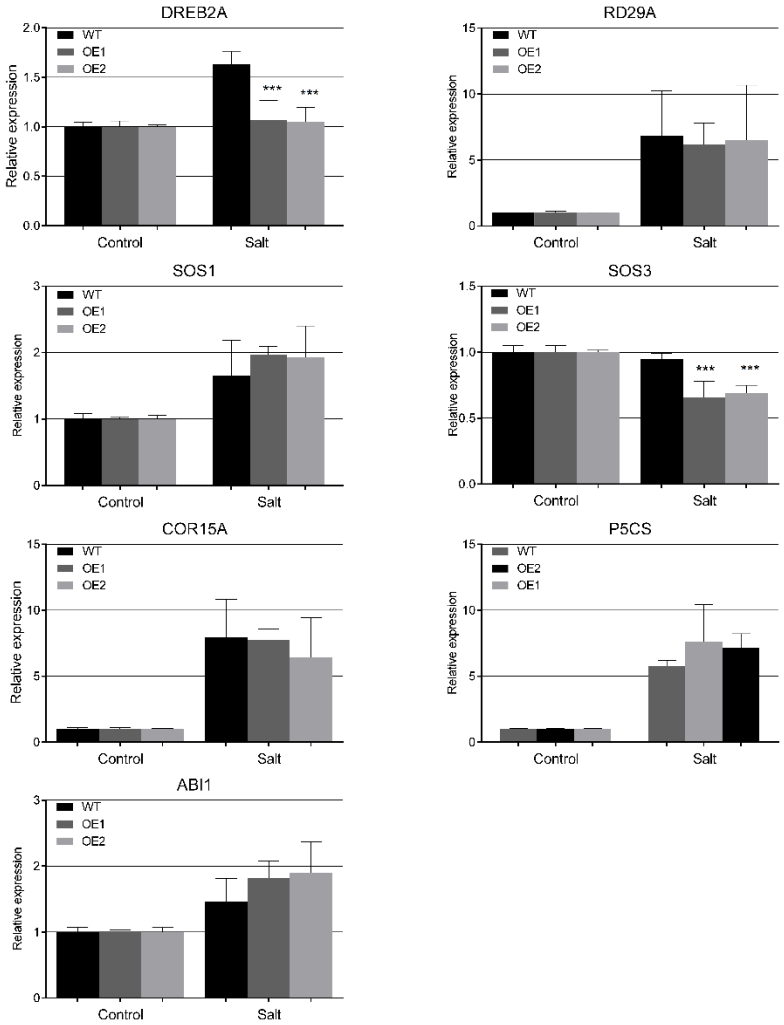


Figure 3.4. **qRT-PCR of target stress-related genes**. One week old seedlings were exposed to 150 mM NaCl for 4 h. Transcript levels of stress-responsive genes were detected with gene-specific primers and Actin II was used as reference in 2(-Delta Delta C(T)) method. Errors bars represent standard deviation among three biological replicas with significant difference marked with asterisks and calculated by one-way ANOVA corrected with Sidak test for multiple comparisons (***P < 0,001).

Expression pattern of *AdDWD1* Arabidopsis homolog.

AdDWD1 Arabidopsis closest homolog (*At1g78070*) retrieved from TAIR database was used in a GUS staining assay. A 1853 bp region before the starting codon was fused with a GUS reporter gene and expression was checked under the microscope for different organs. The level of GUS staining is reported for a strong and a weak line of transformed Arabidopsis in Figure 3.5. The staining in the tissues varies depending on the developmental stage of the plant; in fact a strong staining was detected in one week old seedling with particular emphasis on the cotyledon, leaf vascular tissue, trichomes and main root (Figure 3.4, A and B) and on the primary root of germinated seeds, excluded the root tip (G,H). In adult organs, such as flowers or siliques the staining was very little if absent (Figure 3.5 C, D, F, G).

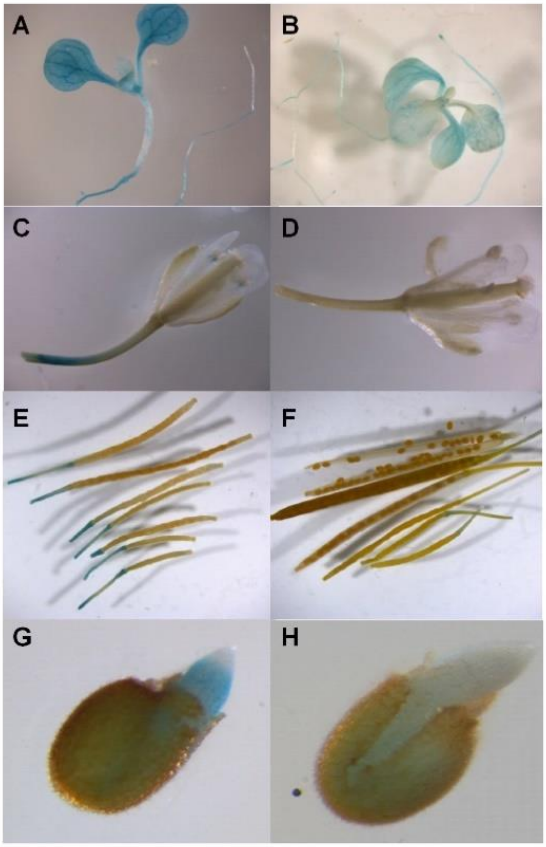


Figure 3.5. **Two GUS reporter lines of the At1g78070 promoter**. A-B) Seedling; C-D) Floral organ; E-F) Siliques; G-H) Seed after primary root emergence.

3.4 Discussion

AdDWD1 is a DWD protein.

DWD proteins (also known as WDXR) constitute a subgroup of the WD40 superfamily in which the key features is a 16 residues sequence at the end of one repeat with typical closing quartet WDXR (He et al., 2006). The proteins having this pattern are demonstrated to link with CUL4-DDB1 complex with particular importance of the Arg at the 16th position that is essential for binding DDB1 (Lee et al., 2008). *AdDWD1* shows the required pattern to be a DWD protein (Figure 3.1B), leads to the hypothesis that this protein participates to protein-protein interaction in the CUL4-DDB1 system. The phylogenetic reconstruction of closely related homologs in Poaceae (Figure 3.1C) reliably identified the orthologs of *AdDWD1*, indicating the evolutionary conservation of this gene across the 107 to 129 million years since the estimate origin of the family (Prasad et al., 2011). The observation that in the majority of the analysed taxa the protein is present in single copy, further suggests its functional specialization. None of the orthologous proteins have been characterized in depth, so *AdDWD1* is the first representative of this subclade of DWD proteins. By contrast to clade A1, where orthologs from all the species except *S. italica* are present, in clade A2 gene loss happened in *O. Sativa*, *B. distachyon* and *Z. mays*, suggesting that subfunctionalization of genes in this clade happened early after the gene duplication at the base of clade A1 and A2. The lineage-specific gene loss observed in *S. italica* further indicate that in this species the function of *AdDWD1* may have been compensated by neofunctionalization of the clade A2 paralogue, Si030175m.

***AdDWD1* is involved in *Arundo donax* osmotic and salt response.**

In the transcriptome released for *Arundo donax* under water stress, the *AdDWD1* transcript (referred to Unigene32775) was upregulated by about 3-fold and 5-fold times after one hour treatment in shoot and root, respectively (Fu et al., 2016). The high responsiveness of *AdDWD1* to osmotic-related stresses (PEG and salt) was confirmed in this study, while the gene transcriptionally reacts much less in the response to the other non-osmotic related stress we tested (cadmium and cold). (Figure 3.2). The involvement of *AdDWD1* in the early stages of osmotic stress response may possibly require the ubiquitin system, which under stress serves for misfolded protein cleavage (Stone, 2014). Several *WD40* and *DWD* proteins are known to be directly involved in abiotic stresses and greatly upregulated during these events (Kong et al., 2014; Luo et al., 2016; Mishra et al., 2014), implying a direct function in defence mechanisms. Worth of note, the general

transcriptional response greatly differed in the two tissues used in this study; in fact, roots reacted faster and stronger, while shoots had a delayed response, if any, in non-osmotic stresses. The reaction responses elicited by the abiotic stresses tested in this study are known to be partly overlapping and share part of the regulatory networks controlling them in time and space. PEG, salt and cold stresses cause all water stress, either by decrease of extracellular water potential (PEG and salt) or by membrane permeabilization caused by decrease fluidity, resulting in increased water loss from the apoplastic compartment. In turn, water stress causes photosynthetic limitation and photodamages consequent to reactive oxygens species (ROS) production (Choudhury et al., 2016). *In-vivo* time-resolved analysis of transcription suggests that *AdDWD1* is primarily involved in the early stage of salt stress sensing/response, as activation of its transcription is observed in both root and shoot as early as 1h 30min after stress application and upregulation in the root is the highest among all types of stress tested. The upregulation in response to heavy metal and cold is most likely the result of the secondary osmotic and oxidative stresses. The responsiveness of *AdDWD1* to osmotic stress is also confirmed by the strong, but slower transcriptional response observed in the roots and the very delayed activation in the shoots, where water stress is effectively buffered by root pressure in grasses (Prieto et al., 2012).

The low abundance of the *AdDWD1* transcript in the majority of unstressed *A. donax* tissues and organs indicates that its expression is tightly controlled, a feature shared by many stress-responsive genes. The highest gene expression was detected in mature leaves and roots (Figure 3.2I). indicating that in organs not actively growing the *AdDWD1* protein may take part in the constitutive process of protein turnover and/or degradation. respect to stem (node and internode) or bud tissues.

Arabidopsis lines 35S::AdDWD1 have reduced fitness under stress during germination.

The physiological analyses carried out in transgenic Arabidopsis overexpressing *AdDWD1* further support the greater involvement of the gene in salt as compared to drought stress. In recent years, much effort was done in understanding the biological role of *WD40* protein finding a large spectre of functions in the plant system such as meristem organization, organ development and apoptosis (van Nocker and Ludwig, 2003). It is also known that in plants, some *DWD* genes are directly involved in abiotic response and specifically in drought. Loss of function mutations *DWD* genes often lead to increased sensitivity to stress (Lee et al., 2010, 2011), while overexpression is often associated to increased tolerance (Kong et al., 2014). However, the overexpression of a foreign gene can also lead to developmental problems especially in the early stages (e.g. Guleria et al., 2014; Zhang and Schroeder, 2010). Our physiological study on overexpressing lines

35S::AdDWD1 shows no differences in growth under normal condition but severe drawbacks in response to stress, especially at the germination stage. In fact, seedling root emergence and growth were greatly reduced under salt and ABA stress and detached leaves of adult plants lost water faster than the WT (Figure 3.3). It is also relevant to notice that the GUS-promoter lines of *Arabidopsis* homologue are expressed especially during seed germination (Figure 3.5G, 3.5H) and seedling stage (Figure 3.5A, 3.5B), that is consistent with a responsive pattern of OE lines at very early stages. It is however possible that lineage-specific differences among monocots and dicots may exist, as suggested by the preferential expression of AdDWD1 in fully developed leaves of *A. donax*. A possible explanation of these results can be made by considering that some *WD40* protein can act also as negative regulator in stresses (Stone, 2014), suppressing other responsive pathways. This is supported by the finding of a lower expression of stress-related genes such as *DREB2A* and *SOS3* that are involved in the ABA-independent pathway (Yoshida et al., 2014) to cope with osmotic and salt stress, respectively (Figure 3.4). In *Arabidopsis* for example, negative regulators of *DREB2A* are the *DREB2A-INTERACTING PROTEIN1 (DRIP1)* and *DRIP2* (Qin et al., 2008). It is demonstrated that *DRIP1* and *DRIP2* inactivate DREB2A protein under non-stress condition and the double mutant *drip1 drip2* has an increased expression of many stress-responsive genes induced by dehydration, especially those regulated by *DREB2A*. Also, *Arabidopsis SOS3* gene have a negative regulator for its expression; in fact transcription factor *AtMyb73* is highly upregulated only under salt stress and its knockout mutant has improved salt tolerance (Kim et al., 2013).

On the other hand, it is possible that *AdDWD1* may interfere in overexpressing *Arabidopsis* lines with formation of the *CUL4-DDB1-WD40* complex. Considering the high similarity with the homolog at the *DWD* box that binds to *DDB1* (He et al., 2006), we can hypothesize that the *CUL4-DDB1* can form a complex with *AdDWD1* protein but this is not able to complete the conjunction with the specific target protein. This may eventually lead to an interruption of the response pathway with decrease stress tolerance of the plant. Further studies to characterize *AdDWD1* targets will be required to elucidate its precise functional role and mechanism of action *in vivo*.

3.5 Conclusions

WD40-repeats proteins are abundant in plants and with high functional differentiation (Mishra et al., 2012b) making them a difficult and interesting target for functional characterization. In this work, we have characterized an *A. donax* gene, which is greatly upregulated in roots and to a lower extent shoots during osmotic stresses (PEG and Salt), while cold and metal treatments

induced expression change only in roots. Moreover, *35S::AdDWD1 Arabidopsis* transgenic lines show a lower germination in response to salt and ABA during the germination stage, probably due to the inability of the giant reed protein to bind the target in the *CUL4-DDB1-WD40*-Target complex. Taken together, these results imply a direct involvement of *AdDWD1* primarily in salt and possibly osmotic stress response, and secondarily in water limitation stress through an ABA-dependent pathway. Its role as possible repressor of salt and osmotic responses in non-stressed condition, however, will require the further elucidation of the protein(s) it likely targets for degradation by the ubiquitin pathway.

3.6 References

Adamec, L., 1984. The effect of plasmolysis and deplasmolysis on the permeability of plant membranes. Biol. Plant. 26, 128–131. doi:10.1007/BF02902278

Angers, S., Li, T., Yi, X., MacCoss, M.J., Moon, R.T., Zheng, N., 2006. Molecular architecture and assembly of the DDB1-CUL4A ubiquitin ligase machinery. Nature 443, 590–593. doi:10.1038/nature05175

Biedermann, S., Hellmann, H., 2011. WD40 and CUL4-based E3 ligases: lubricating all aspects of life. Trends Plant Sci. 16, 38–46. doi:10.1016/j.tplants.2010.09.007

Clough, S.J., Bent, A.F., 1999. Floral dip : a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana 16, 735–743.

Choudhury, F. K., Rivero, R. M., Eduardo Blumwald & Mittler, R. Reactive oxygen species, abiotic stress and stress combination. *Plant J.* 1–12 (2016). doi:10.1111/tpj.13299

Fu, Y., Poli, M., Sablok, G., Wang, B., Liang, Y., La Porta, N., Velikova, V., Loreto, F., Li, M., Varotto, C., 2016. Dissection of early transcriptional responses to water stress in Arundo donax L. by unigene-based RNA-seq. Biotechnol. Biofuels 9, 54. doi:10.1186/s13068-016-0471-8

Guleria, P., Masand, S., Yadav, S.K., 2014. Overexpression of SrUGT85C2 from Stevia reduced growth and yield of transgenic Arabidopsis by influencing plastidial MEP pathway. Gene 539, 250–257. doi:10.1016/j.gene.2014.01.071

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

Hardion, L., Verlaque, R., Rosato, M., Rosselló, J. a., Vila, B., 2015. Impact of polyploidy on fertility variation of Mediterranean Arundo L. (Poaceae). C. R. Biol. doi:10.1016/j.crvi.2015.03.013

He, Y.J., McCall, C.M., Hu, J., Zeng, Y., Xiong, Y., 2006. DDB1 functions as a linker to recruit receptor WD40 proteins to CUL4-ROC1 ubiquitin ligases. Genes Dev. 20, 2949–2954. doi:10.1101/gad.1483206

Huang, J., Wang, M.-M., Bao, Y.-M., Sun, S.-J., Pan, L.-J., Zhang, H.-S., 2008. SRWD: a novel WD40 protein subfamily regulated by salt stress in rice (OryzasativaL.). Gene 424, 71–9. doi:10.1016/j.gene.2008.07.027

Ji, H., Pardo, J.M., Batelli, G., Van Oosten, M.J., Bressan, R.A., Li, X., 2013. The salt overly sensitive (SOS) pathway: Established and emerging roles. Mol. Plant 6, 275–286. doi:10.1093/mp/sst017

Kim, J.H., Nguyen, N.H., Jeong, C.Y., Nguyen, N.T., Hong, S.W., Lee, H., 2013. Loss of the R2R3 MYB, AtMyb73, causes hyper-induction of the SOS1 and SOS3 genes in response to high salinity in Arabidopsis. J. Plant Physiol. 170, 1461–1465. doi:10.1016/j.jplph.2013.05.011

Kim, S., Lee, J., Seo, K., Ryu, B., Sung, Y., Chung, T., Deng, X.W., Lee, J., 2014. Characterization of a Novel DWD Protein that Participates in Heat Stress Response in Arabidopsis. Mol. Cells 37, 833–840. doi:10.14348/molcells.2014.0224

Kong, D., Li, M., Dong, Z., Ji, H., Li, X., 2014. Identification of TaWD40D, a wheat WD40 repeat-containing protein that is associated with plant tolerance to abiotic stresses. Plant Cell Rep. 34, 395–410. doi:10.1007/s00299-014-1717-1

Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33, msw054. doi:10.1093/molbev/msw054

Lee, J.-H., Terzaghi, W., Gusmaroli, G., Charron, J.-B.F., Yoon, H.-J., Chen, H., He, Y.J., Xiong, Y., Deng, X.W., 2008. Characterization of Arabidopsis and rice DWD proteins and their roles as substrate receptors for CUL4-RING E3 ubiquitin ligases. Plant Cell 20, 152–67. doi:10.1105/tpc.107.055418

Lee, J.-H., Yoon, H.-J., Terzaghi, W., Martinez, C., Dai, M., Li, J., Byun, M.-O., Deng, X.W., 2010. DWA1 and DWA2, two Arabidopsis DWD protein components of CUL4-based E3 ligases, act together as negative regulators in ABA signal transduction. Plant Cell 22, 1716–32. doi:10.1105/tpc.109.073783

Lee, J.H., Kim, W.T., 2011. Regulation of abiotic stress signal transduction by E3 ubiquitin ligases in arabidopsis. Mol. Cells 31, 201–208. doi:10.1007/s10059-011-0031-9

Lee, J.H., Terzaghi, W., Deng, X.W., 2011. DWA3, an Arabidopsis DWD protein, acts as a negative regulator in ABA signal transduction. Plant Sci. 180, 352–357. doi:10.1016/j.plantsci.2010.10.008

Luo, Q., Zhao, Z., Li, D.K., Zhang, Y., Xie, L.F., Peng, M.F., Yuan, S., Yang, Y., 2016. Overexpression of NaKR3 enhances salt tolerance in Arabidopsis. Genet. Mol. Res. 15, 1–11. doi:10.4238/gmr.15016378

Mantineo, M., D'Agosta, G.M., Copani, V., Patanè, C., Cosentino, S.L., 2009. Biomass yield and energy balance of three perennial crops for energy use in the semi-arid Mediterranean environment. F. Crop. Res. 114, 204–213. doi:10.1016/j.fcr.2009.07.020

Miller, J.C., Chezem, W.R., Clay, N.K., 2015. Ternary WD40 Repeat-Containing Protein Complexes: Evolution, Composition and Roles in Plant Immunity. Front. Plant Sci. 6, 1108. doi:10.3389/fpls.2015.01108

Mishra, A.K., Muthamilarasan, M., Khan, Y., Parida, S.K., Prasad, M., 2014. Genome-wide investigation and expression analyses of WD40 protein family in the model plant foxtail millet (Setaria italica L.). PLoS One 9, e86852. doi:10.1371/journal.pone.0086852

Mishra, A.K., Puranik, S., Bahadur, R.P., Prasad, M., 2012a. The DNA-binding activity of an AP2 protein is involved in transcriptional regulation of a stress-responsive gene, SiWD40, in foxtail millet. Genomics 100, 252–63. doi:10.1016/j.ygeno.2012.06.012

Mishra, A.K., Puranik, S., Prasad, M., 2012b. Structure and regulatory networks of WD40 protein in plants. J. Plant Biochem. Biotechnol. 21, 32–39. doi:10.1007/s13562-012-0134-1

Nassi o Di Nasso, N., Angelini, L.G., Bonari, E., 2010. Influence of fertilisation and harvest time on fuel quality of giant reed (Arundo donax L.) in central Italy. Eur. J. Agron. 32, 219–227. doi:10.1016/j.eja.2009.12.001

Osakabe, Y., Osakabe, K., Shinozaki, K., Tran, L.-S.P., 2014. Response of plants to water stress. Front. Plant Sci. 5, 86. doi:10.3389/fpls.2014.00086

Ouyang, Y., Huang, X., Lu, Z., Yao, J., 2012. Genomic survey, expression profile and co-expression network analysis of OsWD40 family in rice. BMC Genomics 13, 100. doi:10.1186/1471-2164-13-100

Pilu, R., Bucci, A., Badone, F.C., Landoni, M., 2012. Giant reed (Arundo donax L.): A weed plant or a promising energy crop? AFRICAN J. Biotechnol. 11, 9163–9174. doi:10.5897/AJB11.4182

Pilu, R., Cassani, E., Landoni, M., Badone, F.C., Passera, A., Cantaluppi, E., Corno, L., Adani, F., 2013a. Genetic characterization of an Italian Giant Reed (Arundo donax L.) clones collection: exploiting clonal selection. Euphytica. doi:10.1007/s10681-013-1022-z

Pilu, R., Manca, A., Landoni, M., Agrarie, S., 2013b. Arundo donax as an energy crop: pros and cons of the utilization of this perennial plant. Maydica.

Prasad, V., Strömberg, C. a. E., Leaché, a. D., Samant, B., Patnaik, R., Tang, L., Mohabey, D.M., Ge, S., Sahni, a., 2011. Late Cretaceous origin of the rice tribe provides evidence for early diversification in Poaceae. Nat. Commun. 2, 480. doi:10.1038/ncomms1482

Prieto, I., Armas, C. & Pugnaire, F. I. Water release through plant roots: New insights into its consequences at the plant and ecosystem level. *New Phytol.* **193,** 830–841 (2012).

Qin, F., Sakuma, Y., Tran, L.-S.P., Maruyama, K., Kidokoro, S., Fujita, Y., Fujita, M., Umezawa, T., Sawano, Y., Miyazono, K.-I., Tanokura, M., Shinozaki, K., Yamaguchi-Shinozaki, K., 2008. Arabidopsis DREB2A-interacting proteins function as RING E3 ligases and negatively regulate plant drought stress-responsive gene

expression. Plant Cell 20, 1693–1707. doi:10.1105/tpc.107.057380

Quinn, L.D., Holt, J.S., 2007. Ecological correlates of invasion by Arundo donax in three southern California riparian habitats. Biol. Invasions 10, 591–601. doi:10.1007/s10530-007-9155-4

Sablok, G., Fu, Y., Bobbio, V., Laura, M., Rotino, G.L., Bagnaresi, P., Allavena, A., Velikova, V., Viola, R., Loreto, F., Li, M., Varotto, C., 2014. Fuelling genetic and metabolic exploration of C 3 bioenergy crops through the first reference transcriptome of Arundo donax L. Plant Biotechnol. J. n/a-n/a. doi:10.1111/pbi.12159

Schroeder, D.F., Gahrtz, M., Maxwell, B.B., Cook, R.K., Kan, J.M., Alonso, J.M., Ecker, J.R., Chory, J., 2016. De-Etiolated 1 and Damaged DNA Binding Protein 1 Interact to Regulate Arabidopsis Photomorphogenesis. Curr. Biol. 12, 1462–1472. doi:10.1016/S0960-9822(02)01106-5

Seo, K.-I., Lee, J.-H., Nezames, C.D., Zhong, S., Song, E., Byun, M.-O., Deng, X.W., 2014. ABD1 Is an Arabidopsis DCAF Substrate Receptor for CUL4-DDB1-Based E3 Ligases That Acts as a Negative Regulator of Abscisic Acid Signaling. Plant Cell 26, 695–711. doi:10.1105/tpc.113.119974

Soreng, R.J., Peterson, P.M., Romaschenko, K., Davidse, G., Zuloaga, F.O., Judziewicz, E.J., Filgueiras, T.S., Davis, J.I., Morrone, O., 2015. A worldwide phylogenetic classification of the Poaceae (Gramineae). J. Syst. Evol. 53, 117–137. doi:10.1111/jse.12150

Stirnimann, C.U., Petsalaki, E., Russell, R.B., Müller, C.W., 2010. WD40 proteins propel cellular networks. Trends Biochem. Sci. 35, 565–574. doi:10.1016/j.tibs.2010.04.003

Stone, S.L., 2014. The role of ubiquitin and the 26S proteasome in plant abiotic stress signaling. Front. Plant Sci. 5, 135. doi:10.3389/fpls.2014.00135

Tester, M., Langridge, P., 2010. Breeding Technologies to Increase Crop Production in a Changing World. Science (80-.). 327, 818–822. doi:10.1126/science.1183700

Tuteja, N., Sopory, S.K., 2008. Chemical signaling under abiotic stress environment in plants. Plant Signal. Behav. 3, 525–536. doi:10.4161/psb.3.8.6186

van Nocker, S., Ludwig, P., 2003. The WD-repeat protein superfamily in Arabidopsis: conservation and divergence in structure and function. BMC Genomics 4, 50. doi:10.1186/1471-2164-4-50

Wang, C., Dong, X., Han, L., Su, X.D., Zhang, Z., Li, J., Song, J., 2016. Identification of WD40 repeats by secondary structure-aided profile-profile alignment. J. Theor. Biol. 398, 122–129. doi:10.1016/j.jtbi.2016.03.025

Wang, Y., Jiang, F., Zhuo, Z., Wu, X.H., Wu, Y.D., 2013. A Method for WD40 Repeat Detection and Secondary Structure Prediction. PLoS One 8, 1–13. doi:10.1371/journal.pone.0065705

Yoshida, T., Mogami, J., Yamaguchi-Shinozaki, K., 2014. ABA-dependent and ABA-independent signaling in response to osmotic stress in plants. Curr. Opin. Plant Biol. 21, 133–139. doi:10.1016/j.pbi.2014.07.009

Zhang, Y., Schroeder, D.F., 2010. Effect of overexpression of Arabidopsis Damaged DNA-binding protein 1A on De-etiolated 1. Planta 231, 337–348. doi:10.1007/s00425-009-1056-6

Chapter 4.

Arundo donax *qRT-PCR reference gene selection for abiotic stress.*

Selection of reference genes in the energy crop *Arundo donax L.* under abiotic stresses.

Michele Poli[1,2], Silvio Salvi[2], Mingai Li[1§], Claudio Varotto[1§]

Affiliation:

[1] Department of Biodiversity and Molecular Ecology, Research and Innovation Centre, Fondazione Edmund Mach, Via E. Mach 1, 38010 S. Michele all'Adige (TN), Italy.

[2] Department of Agricultural Sciences, University of Bologna, Bologna, Italy

4.1 Introduction

Reverse transcription real-time quantitative polymerase chain reaction (qRT-PCR) is a well-established and widely used technology for gene expression analyses in many biological fields. Its precision and sensitivity allow to accurately measure the transcriptional variations of a gene among different samples, giving precious information on its function through the characterization of its characteristic tissue- or stress-specific expression pattern. Two methods are used in qRT-PCR experiments: absolute and relative quantification. The absolute quantification identifies the exact copy number of transcribed RNA in a given sample. This method relies on a pre-built calibration curve that associates for each primer combination (specific for the gene of interest) known concentrations of cDNA template standards with the fluorescence data produced during amplification. Fitting the real-time PCR data from unknown samples to the calibration curve, thus,

provides the extrapolated absolute number of copies of target gene (Leong et al., 2007). Relative quantification compares the gene of interest with an internal reference gene to obtain the expression variation (Ginzinger, 2002). This method requires a gene (or multiple genes) with stable expression levels as calibrator to be compared with the gene of interest in order to eliminate possible sources of errors/differences in sample preparation (e.g. RNA extraction, quantification, reverse transcription). Between the two methods, relative quantification is widely used to assess quantitative difference of expression of target genes in test conditions with respect to a control condition (Pfaffl, 2001). Despite many advantages, the most critical aspect of relative quantification is the choice of the internal control, that, ideally, should be expressed at constant level in every condition and tissue/cell type (Bustin et al., 2009). In literature, it has sometimes been implicitly assumed that the most common housekeeping genes can be safely used as references in different conditions and/or species without previous validation (Kozera and Rapacz, 2013). The inherent lack of perfectly stable genes in all tissues and conditions, however, very likely violates this assumption. Thus, preparation of any new qRT-PCR experiment should also include the experimental choice and validation of the best internal controls (Guénin et al., 2009). In plants, several tested candidate genes play a role in basic cellular structure and basic metabolism, like 18S rRNA (18S ribosomal RNA), 28S rRNA (28S ribosomal RNA), *ACT* (Actin), *EF-1α* (Elongation Factor-1α), *GAPDH* (glyceraldehyde-3-phosphate dehydrogenase gene), *TUB- α* (α-tubulin), *TLF* (translation factor), *RPII* (RNApolymerase II). Nevertheless, their basic functions do not exclude possible variability among many conditions or growing stage, so that they need to be proven adequate case by case (Kozera and Rapacz, 2013).

Until recent years only few commonly used plant reference genes were available due to limited sequence information for the majority of non-model species. With the advent of next generation sequencing, however, large-scale transcriptome analysis can provide tens to hundreds of possible reference genes for virtually any species of interest at accessible costs. RNA-Seq technology allows to simultaneously detect virtually all the expressed transcripts of a plant sample, providing at the same time also estimates of their expression levels (Yim et al., 2015). The choice of the best reference, though, is not straightforward and to help in this task many algorithms have been released; among all the most widely used are geNorm (Vandesompele et al., 2002), NormFinder (Andersen et al., 2004), BestKeeper (Pfaffl et al., 2004), RefFinder (Xie et al., 2012) and Delta Ct (Silver et al., 2006; Vandesompele et al., 2002). While most of these tools are equivalent from many points of view, not all of them have been implemented to take into account the same parameters of qRT-PCR. For example, RefFinder algorithm is an online tool that calculates a rank for the other four algorithms but it does not consider the efficiency of each primer pair, assuming it

a priori approximately 100% (De Spiegelaere et al., 2015). Similarly, delta Ct is based on pairwise comparison of raw Cq values among candidates without efficiency correction, which in case of variable amplification efficiencies could lead to biased results.

Arundo donax L. (commonly known as giant reed) is a fast-growing grass that belongs to the Poaceae family and recently became a target species for biofuel production. Its rusticity, together with the high biomass production (up to 40 tons/ha) and low input requirements, make it one of the best options as non-food energy crop, especially in the Mediterranean area where it is native (Hardion et al., 2014). Several research studies characterized the major agronomic features of this crop, focusing mainly on the sustainability of its cultivation in different conditions or environments (Dragoni et al., 2015; Mann et al., 2012; Mantineo et al., 2009; Nassi o Di Nasso et al., 2010), especially in marginal lands (Nassi o Di Nasso et al., 2013). Thanks to its high resistance to abiotic stress, *A. donax* became also an important species in remediation of soil contaminated by heavy metals (Papazoglou et al., 2005). With the recent development of transcriptomics resources for this non-model species (Fu et al., 2016; Sablok et al., 2014; Barrero et al., 2015) a large number of gene sequences is now available for functional characterization. To date, however, no validated set of reference genes specific for expression studies in *A. donax* is available.

To fill this important gap in the functional genomics toolbox of *A. donax*, in this study 8 candidate reference genes were selected and ranked using plant material subjected to three different stresses: osmotic, heavy metal and heat shock. Among these candidates, six have been previously used as controls while two other genes were newly selected candidates, which have been obtained from transcriptome analyses of giant reed and sorghum, respectively. Finally, to validate the former results in a real case, we compared the consistence of expression patterns/levels of *A. donax DREB2A* (Dehydration-Responsive Element Binding Protein 2) (Sakuma, 2006), a well-known drought-related gene, using the two best and the two worst reference genes.

4.2 Results

Gene selection and amplification specificity.

In this study, we selected 8 candidate housekeeping genes form different sources. Four of them, namely *TLF* (*translation factor*), *Act2* (*Actin2*), *Tub α* (*Alpha tubulin*) and *EF-1 α* (*Elongation factor 1-alpha*), have been already used in qRT-PCR experiments in foxtail millet (Kumar et al., 2013)(Kumar; 2013), the species with a sequenced genome which is most closely

related to *Arundo donax* (Sablok et al., 2014). For all these genes, *A. donax* homologs used for primer design shared a nucleotide similarity greater than 91% to foxtail millet references. Foxtail millet published primers were slightly adapted to giant reed with exception of TLF for which completely new primers were designed (Table 4.1, Supplementary File 4.1). By mining the sorghum transcriptome (Shakoor et al., 2014), we extracted two more candidates: the commonly used *GAPDH* (*glyceraldehyde-3-phosphate dehydrogenase 2*) and *RPN6* (*26S proteasome non-ATPase regulatory subunit 11*). *A. donax* homologs share respectively 91% and 92% of sequence similarity to their closest sorghum homologs (Supplementary File 4.1). The other gene from sorghum had been used in a previous study (Fu et al., 2016) as reference species to extrapolate an Actin gene (AC1) from Arundo (89% similarity). Finally, one new gene *pDUF221* has also been considered as candidate gene for relative expression quantification through dissection of water-stressed giant reed transcriptome (Fu et al., 2016). *RPN6* and *pDUF221* genes have been selected after analysis of their coefficient of variation in different organs and stress conditions, resulting respectively 9.38% and 8.01%. Each primer pair gave the expected length and specificity of the amplification when checked with agarose gel 2% and melt curve (Supplementary File 4.1). Moreover, for all primer pairs, amplification efficiency resulted to be between 92.85% and 104.03% and correlation coefficient (R^2) between 0.987 and 0.998 (Table 4.1) (Supplementary File 4.6).

Gene symbol	Gene Name	Arundo Accession	Primer Pair (5'-3')	Product size (bp)	E (%)	R^2	Tm (°C)	Ref
AC1	Actin	Unigene036290	F: TCTTGGCTTGCATTCTTGGG R: TGGATTGCGAAGGCTGAGTAC	93	100,72	0,998	81,5	(Fu et al., 2016)
Act2	Actin 2	Unigene057037	F: CGCATACGTGGCACTTGACT R: GGGCATCTGAACCTCTCTGC	126	92,85	0,987	83,5	(Kumar et al., 2013)
EF-1α	Elongation factor 1-alpha	Unigene076509	F: TGACTGTGCTGTGCTCATCA R: GTTGCAGCAGCAGATCATCT	133	97,13	0,996	83	(Kumar et al., 2013)
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase	Unigene069707	F: TGACAAGGAGAAGGCTGCTG R: GAGCAAGGCAGTTTGTGGTG	167	103,83	0,997	82,5	(Shakoor et al., 2014)
pDUF221	Probable membrane protein DUF221-related Calcium-dependent channel	Unigene070087	F: GACAAAGGAGTCAGCCGTCA R: AACGTGCTTCGGACTTGGAT	91	99,94	0,998	81	(Fu et al., 2016)
RPN6	26S proteasome non-ATPase regulatory subunit 11	Unigene067565	F: CACACGACTAGCAGCTTTCAAG R: TTCAAACGTCGGGAAGGTTG	78	104,03	0,993	80	(Shakoor et al., 2014)
TLF	Translation	Unigene076539	F: GACTTCATGGGTGGTGCTGA	110	100,32	0,998	80	(Kumar et

	Factor		R: TGTTTGTTGGGGGACTTGCT					al., 2013)
TUB α	Tubulin alpha	Unigene068813	F: TACCAGCCACCCTCAGTTGT R: AGTCGAACTTGTGGTCAATGC	121	96,24	0,996	85	(Kumar et al., 2013)
DREB2A*	Dehydration-Responsive Element Binding Protein 2	Unigene057213	F: TCCAGCAGGTAGATCATCTCC R: AGCAGGTTCGGTAATAGGCA	98	98.59	0.999	78	(Fu et al., 2016)

Table 4.1. **Candidate reference genes and their primer sequences used in this study, listed by alphabetical order**. Underlined primers/bases are newly design specific for giant reed. E (%) is primer efficiency calculated with standard curves and formula E= 10^ (-1/slope) *100; R^2 is the correlation coefficient; Tm is the melt curve temperature. Asterisc (*) indicates the validation gene.

Gene expression profile of candidate reference genes.

Gene expression profile was evaluated for the 8 *A. donax* candidate reference genes by quantitative real-time PCR. To assess the transcriptional variation of each gene, cycle threshold (Cq) value was considered among all the samples (2 *A. donax* tissues for 3 stress conditions, each with 5 time points plus 1 pre-stress control) (Supplementary File 4.2). The highest Cq value was detected for the *AC1* gene (lowest expression: 28.58 cycles), *while* the lowest Cq value was measured for *GAPDH* (highest expression: 16.16 cycles). Mean expression values per gene varied from 25.27 of *pDUF221* to 18.84 of *GAPDH*. To provide a more informative stability index, we calculated also the difference between 75th and 25th percentile (ΔP), which is inversely proportional to the spread of the data (Mar et al., 2009). Based on this criterion, *RPN6* (ΔP= 0.82) is the most stable gene, followed by *Act2* (1.07), *GAPDH* (1.01), *TUB α* (1.27), *pDUF221* (1.31), *TLF* (1.32), *EF-1α* (1.48) and *AC1* (P= 1.87) (boxes in Figure 4.1). This stability ranking among genes is also confirmed by the coefficient of variation (CV) that ranges from 2.5% of *RPN6* to 6.24% of *AC1* (Supplementary File 4.3). Comparison of the expression profile of each gene from 0 (control) to 24 h after stress application (Supplementary File 4.7) showed that *GAPDH* was always the most expressed gene in every condition/tissue. The lowest expression levels were mainly associated to *pDUF221*, except for heat shock stress. Only for this stress, in shoot *RPN6* has higher Cq at 3h and 6h and, in root, *RPN6* and *AC1* have higher Cq values for the whole time course (Supplementary File 4.7).

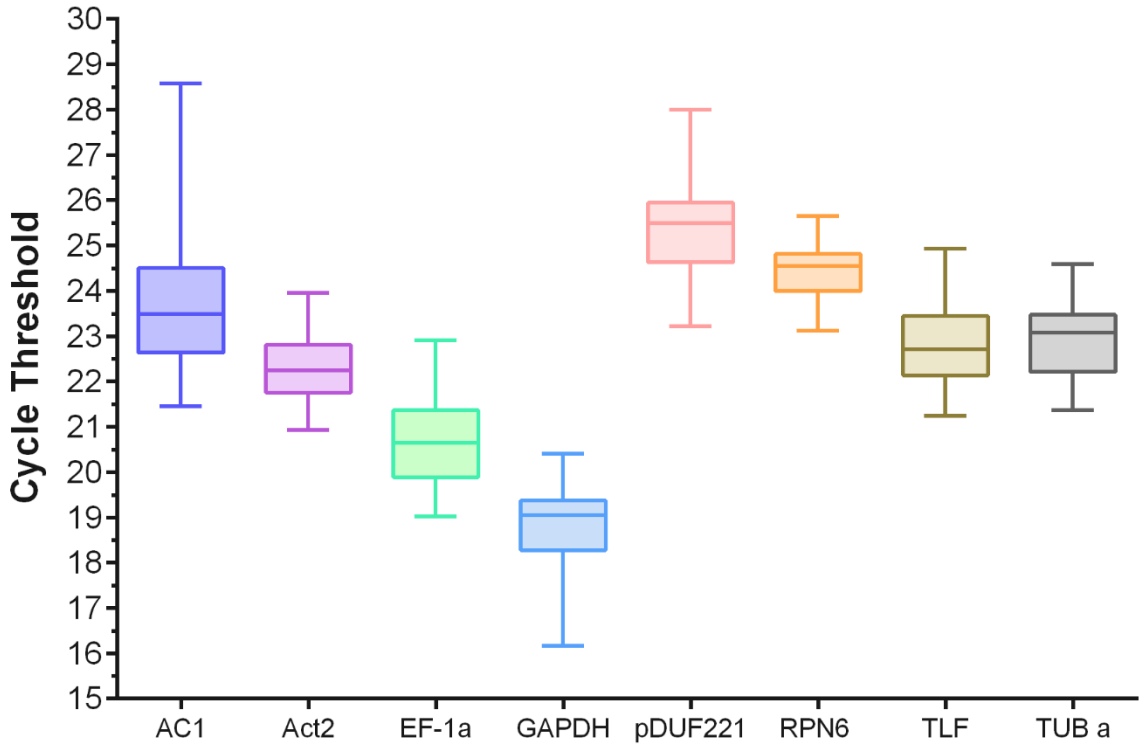


Figure 4.1. **Cq values for 8 candidate genes**. Boxes extend from the 25th to 75th percentiles, whiskers represent the maximum and minimum values, the line across the box represent the median value.

Overall stress analysis.

All treatments, divided according to tissue, were included in an overall analysis to understand which candidate gene is the most stable across samples. For shoot tissue, the results are heterogeneous, with a prevalence of *EF-1α* and *RPN6* in the top 3 genes for all the algorithms (Table 4.2). In this case, geNorm indicates *TLF* and *TUB α* as the most stable genes, while they are ranked as 6th and 4th by NormFinder and as last and 5th by BestKeeper. The least stable genes are instead consistently *AC1* and *pDUF221* with the only exception of BestKeeper that identifies, as said, *TLF* as worst ranking gene. Also, in root *EF-1α* and *RPN6* genes are always in the top 3 positions (with just the intrusion of *Act2* ranked first by BestKeeper), while *pDUF221* and *AC1* are relegated strictly to the last two ranks by all algorithms. Taking in to account both shoot and root organs, *RPN6* is the candidate gene with the best scores in the three algorithms followed by *EF-1α* (first, second and 6th) (Supplementary File 4.4). Overall intergroup and intragroup analyses were also carried out for all samples with NormFinder. This analysis graphically shows the variation of each gene with respect to four subgroups: control, osmotic, heavy metal and heat shock stresses (Supplementary File 4.8A, B). Analysis of the best combination of two genes is consistent with single stability value in root, where *EF-1α+RPN6* (combination value of 0.105) is the best pair of

genes. More surprisingly, in shoot the best pair of genes suggested is *RPN6* (second, S = 0.188) and *TLF* (6th, S= 0.320) with a combination value of 0.091. To notice that also geNorm algorithm includes *TLF* in the best pair of primers together with *TUB α*. Taking together shoot and roots, NormFinder suggests the combination between *RPN6* and *GADPH* (the latter classified at the 6th position as single gene) (Supplementary File 4.8C) These differences with respect to the results of the gene by gene analyses are due to the fact that NormFinder tends to enhance intergroup stability by balancing over/under expression of the two genes to be the closest as possible to 0 in each subgroup.

Rank	Shoot						Root					
	GeNorm		NormFinder		BestKeeper		GeNorm		NormFinder		BestKeeper	
	Gene	Value	Gene	Value	Gene	Value	Gene	Value	Gene	Value	Gene	Value
1	TLF TUB α	0.421	EF-1α	0.132	RPN6	1,00±0,25	EF-1α RPN6	0.353	RPN6	0.111	Act2	2,05±0,45
2	EF-1α	0.480	RPN6	0.188	GAPDH	2,00±0,39	TLF	0.409	EF-1α	0.129	EF-1α	2,40±0,48
3	RPN6	0.559	TUB α	0.232	EF-1α	2,15±0,46	TUB α	0.477	TLF	0.234	RPN6	2,02±0,49
4	GAPDH	0.626	Act2	0.234	Act2	2,70±0,61	GAPDH	0.571	TUB α	0.270	TLF	2,65±0,60
5	Act2	0.655	GAPDH	0.245	TUB α	2,69±0,62	Act2	0.636	Act2	0.306	GAPDH	3,44±0,63
6	AC1	0.700	TLF	0.320	pDUF221	2,69±0,69	pDUF221	0.840	GAPDH	0.421	TUB α	2,76±0,64
7	pDUF221	0.792	AC1	0.470	AC1	3,08±0,72	AC1	1.012	pDUF221	0.783	pDUF221	3,13±0,78
8			pDUF221	0.568	TLF	3,32±0,77			AC1	0.819	AC1	5,82±1,41

Table 4.2. **Overall stress results for different algorithms divided in shoot and root**. Column "value" corresponds to M in geNorm, S in NormFinder and CV±SD in BestKeeper.

Water/Osmotic stress analysis.

Water stress was induced by adding 15% of PEG to the hydroponic solution which decreases the osmotic pressure of the media, consequently inducing osmotic and water limitation stress in the plants (Nicholas and Money, 1989) without toxic effects in the short term (Lawlor, 1970). Reference gene analyses under simulated drought condition in *A. donax* shoot show a high and shared rank for *RPN6* (first in NormFinder and BestKeeper, second in geNorm) followed by *EF-1α* gene (first in geNorm, 4th in NormFinder and third in BestKeeper) (table 4.3). In roots, *RPN6* is confirmed as the most stable gene followed by *GAPDH* that has the best value in geNorm (shared with *RPN6*) and BestKeeper. Using both organ's samples, *RPN6* and *GADPH* are the best two genes (Supplementary File 4.4), resulting in the top 3 ranks provided by all programs. The lowest stability values result from the *AC1* and *pDUF221* genes for all algorithms and tissues with exception of BestKeeper that ranks *pDUF221* 4th in root and *TLF* and *EF-1α* as last in the shoot + root analysis.

Rank	Shoot						Root					
	GeNorm		NormFinder		BestKeeper		GeNorm		NormFinder		BestKeeper	
	Gene	Value	Gene	Value	Gene	Value	Gene	Value	Gene	Value	Gene	Value
1	TLF EF-1α	0.293	RPN6	0.119	RPN6	1,16±0,29	RPN6 GAPDH	0.266	RPN6	0.072	GAPDH	1,73±0,32
2	RPN6	0.345	TUB α	0.141	GAPDH	1,51±0,30	EF-1α	0.308	EF-1α	0.113	RPN6	1,37±0,33
3	TUB α	0.377	TLF	0.208	EF-1α	1,87±0,40	TLF	0.392	GAPDH	0.203	EF-1α	2,22±0,44
4	GAPDH	0.436	EF-1α	0.310	TUB α	1,73±0,41	Act2	0.457	TLF	0.257	pDUF221	1,87±0,47
5	Act2	0.485	GAPDH	0.359	TLF	1,75±0,42	TUB α	0.504	Act2	0.261	Act2	2,53±0,56
6	AC1	0.527	Act2	0.365	Act2	2,02±0,47	pDUF221	0.605	TUB α	0.264	TLF	2,58±0,57
7	pDUF221	0.656	AC1	0.603	AC1	2,50±0,59	AC1	0.705	AC1	0.665	TUB α	3,45±0,78
8			pDUF221	0.726	pDUF221	2,61±0,68			pDUF221	0.676	AC1	4,75±1,09

Table 4.3. **Osmotic stress results for different algorithms divided in shoot and root**. Column "value" corresponds to M in geNorm, S in NormFinder and CV±SD in BestKeeper.

Heavy metal stress analysis.

A. donax is considered a suitable plant for phytoremediation of contaminated soil. Among the heavy metals that can affect soil quality and reduce plant productivity, cadmium (Cd) is one of the most toxic (Benavides et al., 2005). The giant reed seems, however, to cope with it without physiological adaptation, which is an important feature for phytoremediation (Papazoglou et al., 2007, 2005). When subjected to cadmium toxicity stress, reference gene performances are more heterogeneous respect to osmotic stress: in shoot, *RPN6* and *TUB α* rank in the top3 for all the programs, but in root *TLF* is among the best 3 followed by *EF-1α* (first, second and 4th) (Table 4.4). On the other hand, the least stable gene in shoot is *pDUF221*, in root is *AC1* and in the full set of heavy metal stressed samples are *GADPH* and *AC1* (Table 4.4). To notice that BestKeeper indicates as most stable candidate *pDUF221*, which is instead ranked 4th by geNorm and last by NormFinder. Interestingly, analysis of shoot and root together put *pDUF221* always in the top 3 positions, indicating that this gene is the most suitable across different organs for heavy metal stress treatments, but less valuable if organs are taken separately (Supplementary File 4.4). Other general candidate genes to be used in studies encompassing both root and shoot could alternatively be the common *Act2* (ranked 4th, second and first) and *TLF* (ranked third, first and third).

Rank	Shoot						Root					
	GeNorm		NormFinder		BestKeeper		GeNorm		NormFinder		BestKeeper	
	Gene	Value	Gene	Value	Gene	Value	Gene	Value	Gene	Value	Gene	Value
1	RPN6 GAPDH	0.303	TUB α	0.150	RPN6	1,18±0,29	EF-1α RPN6	0.331	TLF	0.117	pDUF221	0,90±0,23
2	TUB α	0.320	RPN6	0.260	GAPDH	2,12±0,41	TLF	0.360	EF-1α	0.139	TLF	1,48±0,33
3	AC1	0.358	AC1	0.311	TUB α	2,26±0,51	Act2	0.388	Act2	0.160	RPN6	1,51±0,36
4	Act2	0.392	TLF	0.324	pDUF221	2,10±0,55	pDUF221	0.448	RPN6	0.162	EF-1α	2,01±0,40
5	EF-1α	0.450	Act2	0.348	Act2	2,49±0,55	TUB α	0.492	GAPDH	0.272	Act2	1,93±0,43
6	TLF	0.489	EF-1α	0.391	EF-1α	2,59±0,55	AC1	0.559	TUB α	0.273	GAPDH	2,76±0,49
7	pDUF221	0.518	GAPDH	0.449	AC1	2,60±0,58	GAPDH	0.627	AC1	0.483	TUB α	2,74±0,62
8			pDUF221	0.737	TLF	3,02±0,68			pDUF221	0.789	AC1	3,35±0,78

Table 4.4. **Heavy metal stress results for different algorithms divided in shoot and root**. Column "value" corresponds to M in geNorm, S in NormFinder and CV±SD in BestKeeper.

Heat shock stress analysis.

Heat shock is an important factor that affects plant physiology and growth (Bita and Gerats, 2013). The capability of *A. donax* to survive in warm environments is a particular and interesting trait of this plant, which can be relevant to forecast its productivity as heat spells become more frequent. Gene stability in heat condition (42 °C) sees the *RPN6* gene as the most stable reference gene in shoot (third, first and first for geNorm, NormFinder and BestKeeper respectively), root (first, first and second) and also considering both organs together (always first) (Table 4.5). Other suitable reference genes for this stress condition are *EF-1α* (first, second, second) and *TUB α* (second, third, 4th) in shoot, *EF-1α* (first, second, first) and *Act2* (third, third and 4th) in root and *TLF* (first, second, 4th) in overall heat shock (Supplementary File 4.4). Again, the genes with overall lower ranking across algorithms are *pDUF221* and *AC1* in all the conditions.

Rank	Shoot						Root					
	GeNorm		NormFinder		BestKeeper		GeNorm		NormFinder		BestKeeper	
	Gene	Value	Gene	Value	Gene	Value	Gene	Value	Gene	Value	Gene	Value
1	TLF EF-1α	0.315	RPN6	0.133	RPN6	0,95±0,24	EF-1α RPN6	0.360	RPN6	0.118	EF-1α	1,92±0,39
2	TUB α	0.367	EF-1α	0.190	EF-1α	1,72±0,37	TLF	0.411	EF-1α	0.131	RPN6	1,88±0,46
3	RPN6	0.398	TUB α	0.282	GAPDH	2,18±0,42	Act2	0.488	Act2	0.219	GAPDH	2,54±0,48
4	GAPDH	0.544	TLF	0.312	TUB α	2,11±0,48	GAPDH	0.542	TLF	0.239	Act2	2,28±0,50
5	Act2	0.599	GAPDH	0.439	Act2	2,17±0,48	TUB α	0.598	TUB α	0.292	TLF	2,48±0,57
6	AC1	0.662	Act2	0.445	AC1	2,15±0,51	pDUF221	0.807	GAPDH	0.352	pDUF221	2,69±0,65
7	pDUF221	0.745	AC1	0.484	TLF	2,42±0,56	AC1	1.019	AC1	0.849	TUB α	3,23±0,76
8			pDUF221	0.496	pDUF221	3,56±0,90			pDUF221	0.851	AC1	5,51±1,41

Table 4.5. **Heat Shock results for different algorithms**. Column "value" corresponds to M in geNorm, S in NormFinder and CV±SD in BestKeeper.

geNorm best reference gene number.

geNorm excel add-in is a useful tool for calculation of the best number of genes that should be used together in a relative qRT-PCR experiment. Therefore, it has been used in this study to predict the optimal number of reference genes to use in each stress experiment (osmotic, heavy metal and heat shock). As expected, the number of references depends on the experimental settings: considering shoot and root together geNorm indicates 3 genes as the most suitable for single stresses and 4 genes considering all stresses (Figure 4.2A). If root and shoot are considered separately, all the values for single stress drop below the suggested threshold of 0.15 (Figure 4.3B-C). Nevertheless, in shoot, by grouping the three stresses, the value is just above to the threshold limit when 2 genes are used while it lays on the threshold with 3.

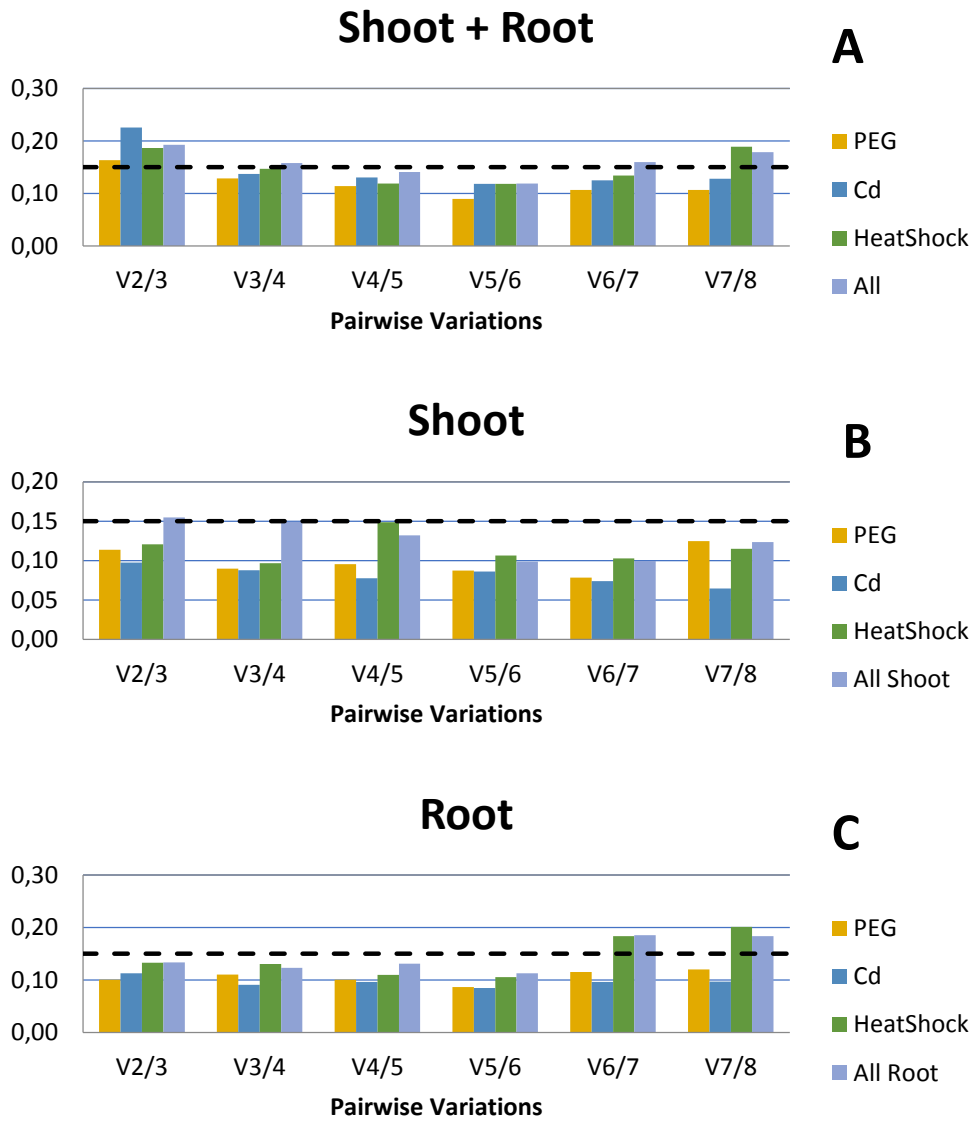


Figure 4.2. **geNorm pairwise variation (Vn/Vn+1)** root and shoot combined (A), only shoot (B) and only root (C) for different treatments (PEG, cadmium, heat shock) and comprising all of them together.

Reference gene validation.

To demonstrate the reliability of the newly analysed reference genes in *A. donax*, a quantitative real time experiment on known *DREB2A* (Dehydration-Responsive Element Binding Protein 2) gene was made. *DREB2A* is one of the key genes in plants triggering the response to both drought and heat shock. We compared the expression profile of *DREB2A* in shoot under osmotic stress against two stable genes (*RPN6* and *EF-1 α*), their combination (*RPN6+EF-1 α*) and the least stable genes (*AC1* and *pDUF221*) based on table 4.3. Fold change of *DREB2A* was calculated with the comparative Ct method (Pfaffl, 2001). The expression pattern among *RPN6*, *EF-1α* and their combination is consistent with a 2-fold increase of DREB2A expression at 6h and 11h and about 4-fold increase at 24h (Fig. 4). On the other hand, the pattern obtained using *AC1* as reference displays an upwards-shifted trend, with a 2-fold increase since the beginning and a 6-fold increase from 6h on. When *pDUF221* is used as reference, the expression pattern decrease at 3h, 6h, and 11h to finally grow again at 24h.

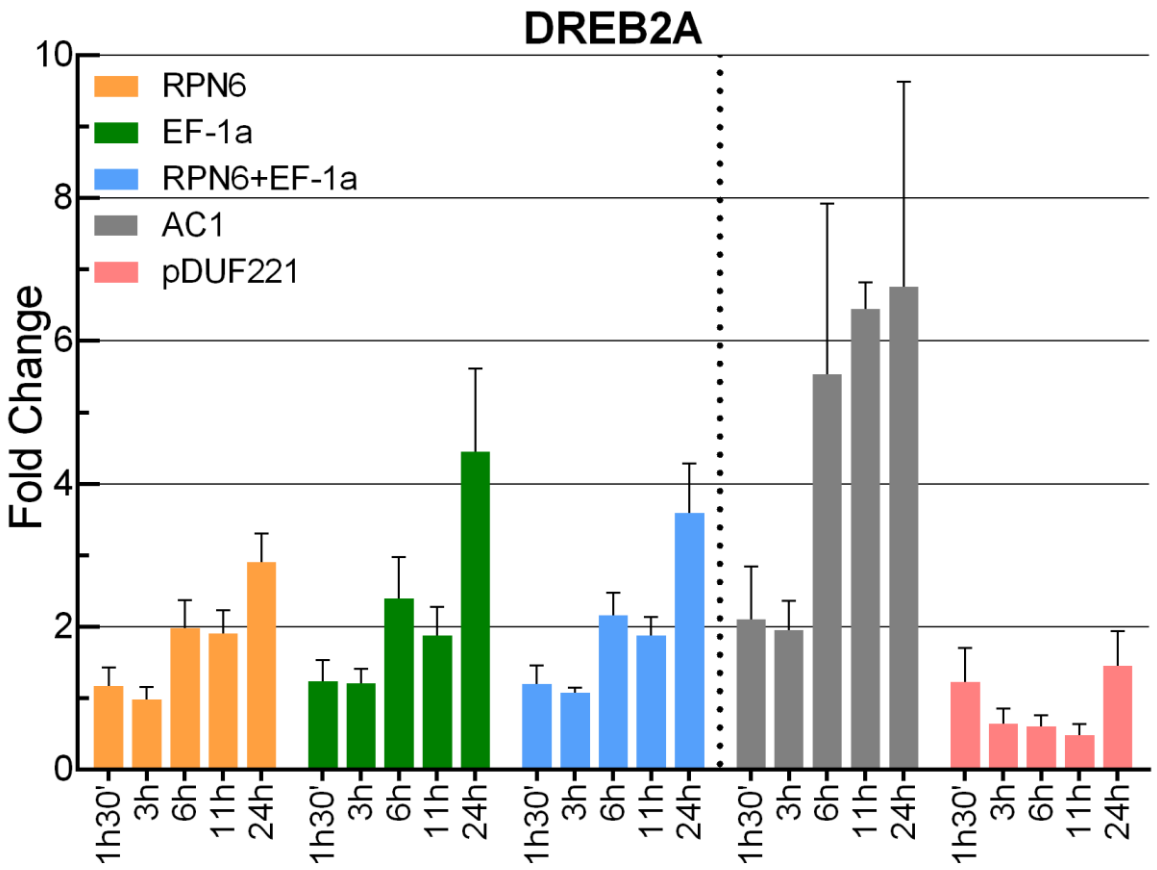


Figure 4.3. **Fold change of DREB2A (Dehydration-Responsive Element-Binding protein 2) in shoot under osmotic stress**. Two most stable genes (RPN6 and EF-1α), their combination (left of the vertical dotted line) and the least stable genes AC1 and pDUF221 (right of dotted line) are considered.

4.3 Discussion

Real Time PCR is a powerful technique for gene expression profiling and functional characterization of genes. Its accuracy, however, is critically dependent from the choice of reference genes whose expression is strictly proportional to the total mRNA amount in the samples to be compared (Guénin et al., 2009). Therefore, ideal reference genes should have the same expression levels (measured in terms of Cq value) among different conditions, tissues, developmental stages and crop varieties (Bustin et al., 2009). In reality, though, gene stability depends greatly on the plant-environment system so that different stress conditions, organs or cultivars can bring undesirable variations in expression of reference genes selected without specific validation, possibly leading also to erroneous quantification (Kozera and Rapacz, 2013). For this reason, a full set of possible reference genes should be developed and systematically tested for each species whenever the experimental conditions change. In this study, 8 candidate reference genes were selected for *Arundo donax* L., an emerging non-food energy crop (Angelini et al., 2009): *AC1* (Actin), *pDUF221* (Probable membrane protein DUF221-related Calcium-dependent channel), *TLF* (Translation Factor), *Act2* (Actin 2), *TUB α* (Tubulin alpha), *EF-1α* (Elongation factor 1-alpha), *RPN6* (26S proteasome non-ATPase regulatory subunit 11) and *GAPDH* (Glyceraldehyde 3-phosphate dehydrogenase). The aim was to evaluate the overall performance of these genes as references for qRT- PCR, but also to provide a detailed indication on which gene is better fit for different organ/stress combinations. The three kinds of abiotic stress applied (osmotic, heavy metal and heat shock) are among the most interesting ones to characterize the functional bases of *A. donax* tolerance to adverse environmental conditions. Three algorithms were applied to the Cq values to measure the stability of the candidate genes: geNorm, NormFinder and BestKeeper. Our results show overall consistence among algorithms on the selection of the reference genes, which indicates the good performance of the methods. We also divided the analysis in stress and organ subsets to identify for each condition the best reference genes. In general, from our study, gene *RPN6* is always among the top 3 genes in both organs under osmotic stress, heat shock and overall analysis that make it a highly suitable reference (Table 4.2). Nevertheless, under heavy metal stress, in root and with complete dataset it presents slightly more variability especially when NormFinder and BestKeeper algorithms are used. On the other hand, geNorm classifies *RPN6* as the best one also in root, making the choice difficult in this context. These results are very interesting, giving the fact that until now, to our knowledge, this gene was considered as possible reference in quantitative RT-PCR only in *Arabidopsis thaliana* (Kim et al., 2014; Lee et al., 2008, 2011) but, to our knowledge,

never in monocot species. Its putative ortholog in *Arabidopsis thaliana* (*AT1G29150*) encodes a protein constituting a lid subunit of the 26S proteasome, which is involved in the ubiquitin-proteasome system (UPS) for degradation of misfolded proteins and stress response (Stone, 2014). Other subunits of this large protease complex have been previously suggested as potential source of new and more stable reference genes in Arabidopsis (Czechowski et al., 2005). Worth of notice, in Arabidopsis *RPN6* is classified as cadmium responsive gene (Sarry et al., 2006), that could explain the variation we observed in response to this stress even though it still remains more stable than some other commonly used reference gene such as *EF-1α* or *AC1* (Table 4.4, Supplementary File 4.4). This fact is possibly due to the high resistance of giant reed to cadmium treatment as demonstrated by (Papazoglou et al., 2007). If shoot and root are considered separately, a useful reference gene across different stresses can be the commonly used *EF-1α* that ranked always in the top three positions. From our analysis, this gene often scores better when the organs are taken separately, especially in osmotic and heat shock stresses (Table 4.3 and Table 4.5), suggesting tissue-specific variability. This makes *EF-1α* a good reference gene if experimental design does not involve the quantitative comparison among different organs. This is particularly interesting when compared with the related *Setaria italica* (L.) P. Beauvois (foxtail millet) gene that was suggested as best internal control under drought and salt condition (Kumar et al., 2013). Here, we confirm the good stability of gene *EF-1α* also in giant reed, but mainly when it is used in single tissues. Taking both tissues together, the *Act2* gene has better performances in each stress condition, especially under heavy metal treatment (Supplementary File 4.4). *GAPDH* is, instead, suitable in osmotic stress, where it results the best among commonly used control genes (Table 4.3). This is supported also by literature where it has been shown that *GAPDH* is one of the best references in plants under drought and salt stress, but not in other conditions (Guo et al., 2014; Lin et al., 2014; Yang et al., 2015).

Finally, we suggest to avoid using two of the candidate genes tested due to their instability: p*DUF221* (a result somehow expected, as it was selected from transcript analysis of osmotic stress only) and *AC1*. As for the number of reference genes to use, geNorm analysis show some interesting traits based on the different employed dataset. In particular, to be notice that the full dataset (all the treatments + both organs) drop below the 0.15 threshold only with 4 reference genes, while in roots it estimates the need of only 2 genes (Figure 4.2A, 2C). This is significant because it implies a differentiation between organs that become difficult to fill in the case of an organ-comparative study. Another point to be notice is the trend of all stresses in shoot (Figure 4.2B) which sees a small change in the values from 2 to 3 references. Based on these results, and

considering that 0.15 is an arbitrary value (Vandesompele et al., 2002), we can suggest that the inclusion of the third gene gives no significant improvements.

DREB2A (Dehydration-Responsive Element Binding Protein 2) is a well-known gene, associated with drought, salt and heat response, that encodes a transcription factor. The *DREB2A* protein interacts with a cis-acting dehydration-responsive element sequence and activates a downstream cascade of drought and heat-responsive genes, thus providing a better tolerance of plants to these stresses (Brulle et al., 2014; Matsukura et al., 2010; Sakuma, 2006). To demonstrate the reliability of the reference genes in this study we have examined the behaviour of *DREB2A* under different reference candidates: the best two (*RPN6* and *EF-1α*) singularly and together and the worse two (*AC1* and *pDUF221*) (Table 4.3). As expected in shoot under osmotic stress, the use of references such as *RPN6*, *EF-1α* and *RPN6+ EF-1α* gave a stable expression profile of *DREB2A*, while the choice of genes with higher variability resulted in both a different pattern and fold change in its expression (Figure 4.4).

In summary, this study provides a wide view of the reference genes that can be used or avoided in *Arundo donax* under specific abiotic stresses, making an important step forward towards reliable and accurate gene expression quantification in this plant. Moreover, thanks to analysis of related species transcriptomes, a new stable gene (*RPN6*) has been successfully used for relative quantification, showing that a deeper comparative analysis of plant transcriptomes can unveil additional candidates for a more precise and reliable quantitative real time PCR.

4.4 Material and Methods

Plant materials and stress treatments.

Cohorts of *A. donax* cuttings (Sesto Fiorentino, Florence, Italy 43°49'01.8"N 11°11'57.0"E) were used in this study. The plant growing condition and procedure for stress treatment were the same as those used previously (Fu et al. 2016). For stress treatments, plants at the 5-leaf stage were transferred from hydroponic solution to fresh one supplemented with 15% PEG 6000 (osmotic stress), 500 μM $CdSO_4$ (heavy metal stress) or prewarmed at 42°C (heat stress). The entire shoots and root system were independently collected before and after stress treatments for 1h 30min, 3h, 6h, 11h and 24h, immediately frozen in liquid nitrogen, and then stored at -80°C till use. Three biological replicates were applied for all the treatments at every sampling time point.

Gene selection and primer design.

Among the sequenced genomes deposited in Phytozome, *Sorghum bicolor* and foxtail millet (*Setaria italica* L.) are the two phylogenetically species most closely related to *A. donax* (Sablok et al., 2014). *A.donax* homologs of four common housekeeping genes from foxtail millet *TLF* (transcription factor), *Act 2* (Actin 2), *Tub α* (Tubulin alpha), and *EF-1 α* (Elongation Factor 1 alpha)(Kumar et al., 2013) were selected by BLASTN searches against the giant reed reference and water-stress transcriptomes (Sablok et al.; 2014; Fu Y. et al; 2016). Analogously, 2 additional genes were identified in the *A. donax* transcriptomes based on sorghum database (*GAPDH* and *RPN6* (Shakoor et al., 2014)). Finally, the last 2 genes were selected directly from the giant reed transcriptome (*AC1* and *pDUF221*; (Fu et al., 2016)) based on their low coefficient of variation (CV) across organs/water stress conditions. Primers were designed with Primer3Plus software (http://primer3plus.com/cgi-bin/dev/primer3plus.cgi) using the following parameters: length 18-25 (optimum 20), product size 75-200bp, melting temperature 59-64°C (optimum 60°C); GC content 30-70% (optimum 50%). Primer pairs with free energy (dG) of dimer formation lower than -5 kcal/mol according to the PerlPrimer v1.1.21 software (http://perlprimer.sourceforge.net/) were discarded.

Total RNA isolation and cDNA synthesis.

Total RNA was isolated with the Spectrum Plant Total RNA Extraction Kit (Sigma) for shoots and the Rneasy® Plant Mini Kit (Qiagen) for roots, respectively. To assure complete absence of genomic DNA contaminations, extracted total RNA was treated with DNase I (Sigma-Aldrich) and checked on 1% agarose gel for integrity control. Concentration and quality of each sample were measured spectrophotometrically through the OD_{260}/OD_{280} absorption ratio. First strand cDNA was reversed transcribed from 1 µg of total RNA primed with oligo-dT in a total reaction mixture of 20 µL using SuperScript® III Reverse Transcriptase (Life Technologies) according to the manufacture's instruction.

PCR and Quantitative Real Time PCR.

To assess the amplification specificity of each primer pair prior to qPCR analysis, PCR amplification was performed in a total volume of 10 µL containing 1 µl of 6-fold diluted cDNA, 1 µL 10x PCR Buffer, 1 µL dNTPs (1 mM), 1 µL of each primer (final concentration of 200 nM) ,0.1

µL of *Taq* polymerase (Sigma) and 4.9 ul of H2O; The PCR programme was as follows: 8 min at 95°C, 33 cycles of 40s at 94 °C, 30s at 60 °C and 20s at 72 °C, with 5 min final extension at 72 °C. The PCR products were run on 2% agarose gel to check single amplification (Supplementary File 4.5A). The qPCR reaction was conducted by mixing 1 µL of 10-fold diluted cDNA (5 ng of starting RNA), 200 nM of each primer and 6.25 µL of Platinum® SYBR® Green qPCR SuperMix-UDG (Invitrogen) in a 12.5 µL total volume. The programme for qRT-PCR in Bio-Rad C1000 Thermal Cycler was set as: 2 min at 50 °C, 2 min at 95 °C, 40 cycles of 15s at 95 °C and 30s at 60 °C (59 °C for validation of *DREB2A* gene). The melting curve was obtained from amplified template for every gene by constantly rising the temperature from 65 °C to 90 °C (Supplementary File 4.5B). A standard curve of qPCR reaction was generated from five points (four points for *pDUF221*gene and validation gene *DREB2A*) of a 6-fold dilution series (10-fold dilution for *pDUF221* and *TLF*). The slope (S) of the standard curve was used to calculate the amplification efficiency (E) of each primer pair as follows: $E = 10^{(-1/S)}$ (Supplementary File 4.6). Three technical replicates were used for each sample and every plate contained one No Template Control (NTC) well for each primer pair used. In order to compare different plates, in Bio-Rad CFX Manager software the baseline threshold was set at 329.82 and one control sample was used in every plate to check for Cq congruency.

geNorm, NormFinder and BestKeeper analyses.

The tests were conducted in different groups of data that comprise overall samples and single stress samples (osmotic, heavy metal and heat shock) divided in organs (shoot and roots). This is necessary because the reference gene may vary depending on the experimental settings (Bustin et al., 2009). Three different Excel-based algorithms have been applied for data analysis. geNorm v3.5 (Vandesompele et al., 2002) and NormFinder v0.953 (Andersen et al., 2004) require relative input data, so the Cq values were converted with the formula $2^{-\Delta Ct}$ where ΔCt is the difference of each Cq value minus the lowest Cq value (highest expression level). BestKeeper (Pfaffl et al., 2004) instead, uses raw Cq values. geNorm calculates stability value (M) based on the average pairwise comparison with a stepwise exclusion of the highest M value (least stable gene). Further, geNorm calculates the number of genes needed for a reliable normalization considering the pairwise variation (V_n/V_{n+1}) between sequential normalization factors, NF_n and NF_{n+1}. This number is optimal when the addition of one more reference does not give a significant contribution to the normalization factor (NF_{n+1}) or, as suggested, the value drops below 0.15. NormFinder uses an ANOVA-based algorithm to estimate intra- and inter- group variation for a given set of experiments, providing a rank where the most stable gene is the one with lowest stability (S) value.

Moreover, NormFinder gives the best gene pair combination that minimizes the expression differences among subgroups, if subgroups are set. BestKeeper, differently from the other algorithms, does not provide a direct rank list but calculates standard deviation (SD [± Cq]) and coefficient of variation (CV [%Cq]) for each gene. We sorted the CV values to rank the genes from most stable (lowest CV value) to least stable (highest CV value).

4.5 References

Adamec, L., 1984. The effect of plasmolysis and deplasmolysis on the permeability of plant membranes. Biol. Plant. 26, 128–131. doi:10.1007/BF02902278

Ahmad, R., Liow, P.-S., Spencer, D.F., Jasieniuk, M., 2008. Molecular evidence for a single genetic clone of invasive Arundo donax in the United States. Aquat. Bot. 88, 113–120. doi:10.1016/j.aquabot.2007.08.015

Ahrar, M., Doneva, D., Koleva, D., Romano, A., Rodeghiero, M., Tsonev, T., Biasioli, F., Stefanova, M., Peeva, V., Wohlfahrt, G., Loreto, F., Varotto, C., Velikova, V., 2015. Isoprene emission in the monocot Arundineae tribe in relation to functional and structural organization of the photosynthetic apparatus. Environ. Exp. Bot. 119, 87–95. doi:10.1016/j.envexpbot.2015.04.010

Alwine, J.C., Kemp, D.J., Stark, G.R., 1977. Method for detection of specific RNAs in agarose gels by transfer to diazobenzyloxymethyl-paper and hybridization with DNA probes. Proc. Natl. Acad. Sci. U. S. A. 74, 5350–4. doi:10.1073/pnas.74.12.5350

Andersen, C.L., Ledet-Jensen, J., Orntoft, T., 2004. Normalization of real-time quantitative RT-PCR data: a mode-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. Cancer Res. 64, 5245–5250.

Angelini, L.G., Ceccarini, L., Nassi o Di Nasso, N., Bonari, E., 2009. Comparison of Arundo donax L. and Miscanthus x giganteus in a long-term field experiment in Central Italy: Analysis of productive characteristics and energy balance. Biomass and Bioenergy 33, 635–643. doi:10.1016/j.biombioe.2008.10.005

Angers, S., Li, T., Yi, X., MacCoss, M.J., Moon, R.T., Zheng, N., 2006. Molecular architecture and assembly of the DDB1-CUL4A ubiquitin ligase machinery. Nature 443, 590–593. doi:10.1038/nature05175

Baillie, R., Drayton, M., Pembleton, L., Kaur, S., Culvenor, R., Smith, K., Spangenberg, G., Forster, J., Cogan, N., 2017. Generation and Characterisation of a Reference Transcriptome for Phalaris (Phalaris aquatica L.). Agronomy 7, 14. doi:10.3390/agronomy7010014

Bell, G.P., 1997. Ecology and Management of Arundo donax, and Approaches to Riparian Habitat Restoration in Southern California., in: Brock, J. (Ed.), Plant Invasions Studies from North America and Europe. Backhuys, pp. 103–113.

Benavides, M.P., Gallego, S.M., Tomaro, M.L., 2005. Cadmium toxicity in plants. Brazilian J. Plant Physiol. 17, 21–34. doi:10.1590/S1677-04202005000100004

Biedermann, S., Hellmann, H., 2011. WD40 and CUL4-based E3 ligases: lubricating all aspects of life. Trends Plant Sci. 16, 38–46. doi:10.1016/j.tplants.2010.09.007

Bita, C.E., Gerats, T., 2013. Plant tolerance to high temperature in a changing environment: scientific fundamentals and production of heat stress-tolerant crops. Front. Plant Sci. 4, 273. doi:10.3389/fpls.2013.00273

Brulle, F., Bernard, F., Vandenbulcke, F., Cuny, D., Dumez, S., 2014. Identification of suitable qPCR reference genes in leaves of Brassica oleracea under abiotic stresses. Ecotoxicology 23, 459–471. doi:10.1007/s10646-014-1209-7

Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L., Vandesompele, J., Wittwer, C.T., 2009. The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. Clin. Chem. 55, 611–622. doi:10.1373/clinchem.2008.112797

Byrt, C.S., Grof, C.P.L., Furbank, R.T., 2011. C4 plants as biofuel feedstocks: optimising biomass production and feedstock quality from a lignocellulosic perspective. J. Integr. Plant Biol. 53, 120–35. doi:10.1111/j.1744-7909.2010.01023.x

Cardoso-Silva, C.B., Costa, E.A., Mancini, M.C., Balsalobre, T.W.A., Costa Canesin, L.E., Pinto, L.R., Carneiro, M.S., Garcia, A.A.F., De Souza, A.P., Vicentini, R., 2014. De novo assembly and transcriptome analysis of contrasting sugarcane varieties. PLoS One 9. doi:10.1371/journal.pone.0088462

Clough, S.J., Bent, A.F., 1999. Floral dip : a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana 16, 735–743.

Czechowski, T., Stitt, M., Altmann, T., Udvardi, M.K., 2005. Genome-wide identification and testing of superior reference genes for transcript normalization. Society 139, 5–17. doi:10.1104/pp.105.063743.1

da Fonseca, R.R., Albrechtsen, A., Themudo, G.E., Ramos-Madrigal, J., Sibbesen, J.A., Maretty, L., Zepeda-Mendoza, M.L., Campos, P.F., Heller, R., Pereira, R.J., 2016. Next-generation biology: Sequencing and data analysis approaches for non-model organisms. Mar. Genomics 30, 1–11. doi:10.1016/j.margen.2016.04.012

Danin, A., 2004. Arundo (Gramineae) in the Mediterranean reconsidered 34204, 361–369. doi:10.3372/wi.34.34204

Danin, A., Raus, T., Scholz, H., 2002. Contribution to the Flora of Greece: A New Species of Arundo (Poaceae). Willdenowia 32, 191–194.

De Spiegelaere, W., Dern-Wieloch, J., Weigel, R., Schumacher, V., Schorle, H., Nettersheim, D., Bergmann, M., Brehm, R., Kliesch, S., Vandekerckhove, L., Fink, C., 2015. Reference gene validation for RT-qPCR, a note on different available software packages. PLoS One 10, 1–13. doi:10.1371/journal.pone.0122515

Divi, U.K., Zhou, X.R., Wang, P., Butlin, J., Zhang, D.M., Liu, Q., Vanhercke, T., Petrie, J.R., Talbot, M., White, R.G., Taylor, J.M., Larkin, P., Singh, S.P., 2016. Deep sequencing of the fruit transcriptome and lipid accumulation in a non-seed tissue of Chinese tallow, a potential biofuel crop. Plant Cell Physiol. 57, 125–137. doi:10.1093/pcp/pcv181

Doebley, J., Gaut, B., Smith, B., 2006. The molecular genetics of crop domestication. Cell 1309–1321.

Dragoni, F., Ragaglini, G., Corneli, E., o di Nasso, N.N., Tozzini, C., Cattani, S., Bonari, E., 2015. Giant reed (Arundo donax L.) for biogas production: Land use saving and nitrogen utilisation efficiency compared with arable crops. Ital. J. Agron. 10, 192–201. doi:10.4081/ija.2015.664

Dugas, D. V, Monaco, M.K., Olsen, A., Klein, R.R., Kumari, S., Ware, D., Klein, P.E., 2011. Functional annotation of the transcriptome of Sorghum bicolor in response to osmotic stress and abscisic acid. BMC Genomics 12, 514. doi:10.1186/1471-2164-12-514

Fu, Y., Poli, M., Sablok, G., Wang, B., Liang, Y., La Porta, N., Velikova, V., Loreto, F., Li, M., Varotto, C., 2016. Dissection of early transcriptional responses to water stress in Arundo donax L. by unigene-based RNA-seq. Biotechnol. Biofuels 9, 54. doi:10.1186/s13068-016-0471-8

Fuentes, J.D., Lerdau, M., Atkinson, R., Baldocchi, D., Bottenheim, J.W., Ciccioli, P., Lamb, B., Geron, C., Gu, L., Guenther, A., Sharkey, T.D., Stockwell, W., 2000. Biogenic Hydrocarbons in the Atmospheric Boundary Layer: A Review. Bull. Am. Meteorol. Soc. 81, 1537–1575. doi:10.1175/1520-0477(2000)081<1537:BHITAB>2.3.CO;2

Ginzinger, D.G., 2002. Gene quantification using real-time quantitative PCR: An emerging technology hits the mainstream. Exp. Hematol. 30, 503–512. doi:10.1016/S0301-472X(02)00806-8

Goodwin, S., McPherson, J.D., McCombie, W.R., 2016. Coming of age: ten years of next-generation sequencing technologies. Nat Rev Genet 17, 333–351. doi:10.1038/nrg.2016.49

Guénin, S., Mauriat, M., Pelloux, J., Van Wuytswinkel, O., Bellini, C., Gutierrez, L., 2009. Normalization of qRT-PCR data: The necessity of adopting a systematic, experimental conditions-specific, validation of references. J. Exp. Bot. 60, 487–493. doi:10.1093/jxb/ern305

Guleria, P., Masand, S., Yadav, S.K., 2014. Overexpression of SrUGT85C2 from Stevia reduced growth and yield of transgenic Arabidopsis by influencing plastidial MEP pathway. Gene 539, 250–257. doi:10.1016/j.gene.2014.01.071

Guo, J., Ling, H., Wu, Q., Xu, L., Que, Y., 2014. The choice of reference genes for assessing gene expression in sugarcane under salinity and drought stresses. Sci. Rep. 4, 7042. doi:10.1038/srep07042

Haddadchi, A., Gross, C.L., Fatemi, M., 2013. The expansion of sterile Arundo donax (Poaceae) in southeastern Australia is accompanied by genotypic variation. Aquat. Bot. 104, 153–161. doi:10.1016/j.aquabot.2012.07.006

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

Hardion, L., Verlaque, R., Callmander, M.M.W., Vila, B., 2012b. Arundo micrantha Lam.(Poaceae), the correct name for Arundo mauritanica Desf. and Arundo mediterranea Danin. Candollea 67, 131–135.

Hardion, L., Verlaque, R., Rosato, M., Rosselló, J. a., Vila, B., 2015. Impact of polyploidy on fertility variation of Mediterranean Arundo L. (Poaceae). C. R. Biol. doi:10.1016/j.crvi.2015.03.013

Hardion, L., Verlaque, R., Saltonstall, K., Leriche, A., Vila, B., 2014. Origin of the invasive Arundo donax (Poaceae): a trans-Asian expedition in herbaria. Ann. Bot. mcu143-. doi:10.1093/aob/mcu143

Harley, P., Vasconcellos, P., Vierling, L., Pinheiro, C.C.D.S., Greenberg, J., Guenther, A., Klinger, L., De Almeida, S.S., Neill, D., Baker, T., Phillips, O., Malhi, Y., 2004. Variation in potential for isoprene emissions among Neotropical forest sites. Glob. Chang. Biol. 10, 630–650. doi:10.1111/j.1529-8817.2003.00760.x

He, Y.J., McCall, C.M., Hu, J., Zeng, Y., Xiong, Y., 2006. DDB1 functions as a linker to recruit receptor WD40 proteins to CUL4-ROC1 ubiquitin ligases. Genes Dev. 20, 2949–2954. doi:10.1101/gad.1483206

Hollingsworth, M.L., Bailey, J.P., 2000. Evidence for massive clonal growth in the invasive weed Fallopia japonica (Japanese Knotweed). Bot. J. Linn. Soc. 133, 463–472. doi:10.1006/bojl.2000.0359

Huang, J., Wang, M.-M., Bao, Y.-M., Sun, S.-J., Pan, L.-J., Zhang, H.-S., 2008. SRWD: a novel WD40 protein subfamily regulated by salt stress in rice (OryzasativaL.). Gene 424, 71–9. doi:10.1016/j.gene.2008.07.027

Janda, L., Tichý, P., Spížek, J., Petříček, M., 1996. A deduced Thermomonospora curvata protein containing serine/threonine protein kinase and WD-repeat domains. J. Bacteriol. 178, 1487–1489.

Ji, H., Pardo, J.M., Batelli, G., Van Oosten, M.J., Bressan, R.A., Li, X., 2013. The salt overly sensitive (SOS) pathway: Established and emerging roles. Mol. Plant 6, 275–286. doi:10.1093/mp/sst017

Kagale, S., Nixon, J., Khedikar, Y., Pasha, A., Provart, N.J., Clarke, W.E., Bollina, V., Robinson, S.J., Coutu, C., Hegedus, D.D., Sharpe, A.G., Parkin, I.A.P., 2016. The developmental transcriptome atlas of the biofuel crop Camelina sativa. Plant J. 879–894. doi:10.1111/tpj.13302

Kausar, S., Mahmood, Q., Raja, I.A., Khan, A., Sultan, S., Gilani, M.A., Shujaat, S., 2012. Potential of Arundo donax to treat chromium contamination. Ecol. Eng. 42, 256–259. doi:10.1016/j.ecoleng.2012.02.019

Khudamrongsawat, J., Tayyar, R., Holt, J.S., 2004. Genetic diversity of giant reed (Arundo donax) in the Santa Ana River, California. Weed Sci. doi:10.1614/WS-03-120R1

Kim, J.H., Nguyen, N.H., Jeong, C.Y., Nguyen, N.T., Hong, S.W., Lee, H., 2013. Loss of the R2R3 MYB, AtMyb73, causes hyper-induction of the SOS1 and SOS3 genes in response to high salinity in Arabidopsis. J. Plant Physiol. 170, 1461–1465. doi:10.1016/j.jplph.2013.05.011

Kim, S., Lee, J., Seo, K., Ryu, B., Sung, Y., Chung, T., Deng, X.W., Lee, J., 2014. Characterization of a Novel DWD Protein that Participates in Heat Stress Response in Arabidopsis. Mol. Cells 37, 833–840. doi:10.14348/molcells.2014.0224

Kong, D., Li, M., Dong, Z., Ji, H., Li, X., 2014. Identification of TaWD40D, a wheat WD40 repeat-containing protein that is associated with plant tolerance to abiotic stresses. Plant Cell Rep. 34, 395–410. doi:10.1007/s00299-014-1717-1

Kozera, B., Rapacz, M., 2013. Reference genes in real-time PCR. J. Appl. Genet. 54, 391–406. doi:10.1007/s13353-013-0173-x

Kumar, K., Muthamilarasan, M., Prasad, M., 2013. Reference genes for quantitative real-time PCR analysis in the model plant foxtail millet (Setaria italica L.) subjected to abiotic stress conditions. Plant Cell, Tissue Organ Cult. 115, 13–22. doi:10.1007/s11240-013-0335-x

Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33, msw054. doi:10.1093/molbev/msw054

Lal, R., 2009. Soil quality impacts of residue removal for bioethanol production. Soil Tillage Res. 102, 233–241. doi:10.1016/j.still.2008.07.003

Lawlor, D.W., 1970. Absorption of Polyethylene Glycols by Plants and Their Effects on Plant Growth. New Phytol. doi:10.1111/j.1469-8137.1970.tb02446.x

Lee, J.-H., Terzaghi, W., Gusmaroli, G., Charron, J.-B.F., Yoon, H.-J., Chen, H., He, Y.J., Xiong, Y., Deng, X.W., 2008. Characterization of Arabidopsis and rice DWD proteins and their roles as substrate receptors for CUL4-RING E3 ubiquitin ligases. Plant Cell 20, 152–67. doi:10.1105/tpc.107.055418

Lee, J.-H., Yoon, H.-J., Terzaghi, W., Martinez, C., Dai, M., Li, J., Byun, M.-O., Deng, X.W., 2010. DWA1 and DWA2, two Arabidopsis DWD protein components of CUL4-based E3 ligases, act together as negative regulators in ABA signal transduction. Plant Cell 22, 1716–32. doi:10.1105/tpc.109.073783

Lee, J.H., Terzaghi, W., Deng, X.W., 2011. DWA3, an Arabidopsis DWD protein, acts as a negative regulator in ABA signal transduction. Plant Sci. 180, 352–357. doi:10.1016/j.plantsci.2010.10.008

Leong, D.T., Gupta, A., Bai, H.F., Wan, G., Yoong, L.F., Too, H.P., Chew, F.T., Hutmacher, D.W., 2007. Absolute quantification of gene expression in biomaterials research using real-time PCR. Biomaterials 28, 203–210. doi:10.1016/j.biomaterials.2006.09.011

Lewandowski, I., Scurlock, J.M.O.O., Lindvall, E., Christou, M., 2003. The development and current status of perennial rhizomatous grasses as energy crops in the US and Europe. Biomass and Bioenergy 25, 335–361. doi:10.1016/S0961-9534(03)00030-8

Li, J., Ye, W.H., 2006. Genetic diversity of alligator weed ecotypes is not the reason for their different responses to biological control. Aquat. Bot. 85, 155–158. doi:10.1016/j.aquabot.2006.02.006

Li, W., Wang, B., Wang, J., 2006. Lack of genetic variation of an invasive clonal plant Eichhornia crassipes in China revealed by RAPD and ISSR markers. Aquat. Bot. 84, 176–180. doi:10.1016/j.aquabot.2005.09.008

Lin, Y., Zhang, C., Lan, H., Gao, S., Liu, H., Liu, J., Cao, M., Pan, G., Rong, T., Zhang, S., 2014. Validation of potential reference genes for qPCR in maize across abiotic stresses, hormone treatments, and tissue types. PLoS One 9. doi:10.1371/journal.pone.0095445

Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25, 402–8. doi:10.1006/meth.2001.1262

Luo, Q., Zhao, Z., Li, D.K., Zhang, Y., Xie, L.F., Peng, M.F., Yuan, S., Yang, Y., 2016. Overexpression of NaKR3 enhances salt tolerance in Arabidopsis. Genet. Mol. Res. 15, 1–11. doi:10.4238/gmr.15016378

Mann, J.J., Barney, J.N., Kyser, G.B., DiTomaso, J.M., 2012. Miscanthus × giganteus and Arundo donax shoot and rhizome tolerance of extreme moisture stress. GCB Bioenergy n/a-n/a. doi:10.1111/gcbb.12039

Mantineo, M., D'Agosta, G.M., Copani, V., Patanè, C., Cosentino, S.L., 2009. Biomass yield and energy balance of three perennial crops for energy use in the semi-arid Mediterranean environment. F. Crop. Res. 114, 204–213. doi:10.1016/j.fcr.2009.07.020

Mar, J.C., Kimura, Y., Schroder, K., Irvine, K.M., Hayashizaki, Y., Suzuki, H., Hume, D., Quackenbush, J., 2009. Data-driven normalization strategies for high-throughput quantitative RT-PCR. BMC Bioinformatics 10, 110. doi:10.1186/1471-2105-10-110

Marguerat, S., Bähler, J., 2010. RNA-seq: from technology to biology. Cell. Mol. Life Sci. 67, 569–79. doi:10.1007/s00018-009-0180-6

Mariani, C., Cabrini, R., Danin, a., Piffanelli, P., Fricano, a., Gomarasca, S., Dicandilo, M., Grassi, F., Soave, C., 2010. Origin, diffusion and reproduction of the giant reed (Arundo donax L.): a promising weedy energy crop. Ann. Appl. Biol. 157, 191–202. doi:10.1111/j.1744-7348.2010.00419.x

Matsukura, S., Mizoi, J., Yoshida, T., Todaka, D., Ito, Y., Maruyama, K., Shinozaki, K., Yamaguchi-Shinozaki, K., 2010. Comprehensive analysis of rice DREB2-type genes that encode transcription factors involved in the expression of abiotic stress-responsive genes. Mol. Genet. Genomics 283, 185–196. doi:10.1007/s00438-009-0506-y

Miller, J.C., Chezem, W.R., Clay, N.K., 2015. Ternary WD40 Repeat-Containing Protein Complexes: Evolution, Composition and Roles in Plant Immunity. Front. Plant Sci. 6, 1108. doi:10.3389/fpls.2015.01108

Minogue, P., Wright, S., 2016. Biology , Control and Invasive Potential of Giant Reed 2–6.

Mirza, N., Mahmood, Q., Pervez, A., Ahmad, R., Farooq, R., Shah, M.M., Azim, M.R., 2010. Phytoremediation potential of Arundo donax in arsenic-contaminated synthetic wastewater. Bioresour. Technol. 101, 5815–9. doi:10.1016/j.biortech.2010.03.012

Mishra, A.K., Muthamilarasan, M., Khan, Y., Parida, S.K., Prasad, M., 2014. Genome-wide investigation and expression analyses of WD40 protein family in the model plant foxtail millet (Setaria italica L.). PLoS One 9, e86852. doi:10.1371/journal.pone.0086852

Mishra, A.K., Puranik, S., Bahadur, R.P., Prasad, M., 2012a. The DNA-binding activity of an AP2 protein is involved in transcriptional regulation of a stress-responsive gene, SiWD40, in foxtail millet. Genomics 100, 252–63. doi:10.1016/j.ygeno.2012.06.012

Mishra, A.K., Puranik, S., Prasad, M., 2012b. Structure and regulatory networks of WD40 protein in plants. J. Plant Biochem. Biotechnol. 21, 32–39. doi:10.1007/s13562-012-0134-1

Monson, R.K., 2002. Volatile organic compound emissions from terrestrial ecosystems: A primary biological control over atmospheric chemistry. Isr. J. Chem. 42, 29–42. doi:10.1560/0jjc-xqaa-jx0g-fxjg

Monson, R.K., Trahan, N., Rosenstiel, T.N., Veres, P., Moore, D., Wilkinson, M., Norby, R.J., Volder, A., Tjoelker, M.G., Briske, D.D., Karnosky, D.F., Fall, R., 2007. Isoprene emission from terrestrial ecosystems in response to global change: minding the gap between models and observations. Philos. Trans. A. Math. Phys. Eng. Sci. 365, 1677–95. doi:10.1098/rsta.2007.2038

Naik, S.N., Goud, V. V., Rout, P.K., Dalai, A.K., 2010. Production of first and second generation biofuels: A comprehensive review. Renew. Sustain. Energy Rev. 14, 578–597. doi:10.1016/j.rser.2009.10.003

Nassi o Di Nasso, N., Angelini, L.G., Bonari, E., 2010. Influence of fertilisation and harvest time on fuel quality of giant reed (Arundo donax L.) in central Italy. Eur. J. Agron. 32, 219–227. doi:10.1016/j.eja.2009.12.001

Nassi o Di Nasso, N., Roncucci, N., Bonari, E., 2013. Seasonal Dynamics of Aboveground and Belowground Biomass and Nutrient Accumulation and Remobilization in Giant Reed (Arundo donax L.): A Three-Year Study on Marginal Land. BioEnergy Res. doi:10.1007/s12155-012-9289-9

Neer, E.J., Schmidt, C.J., Nambudripad, R., Smith, T.F., 1994. The ancient regulatory-protein family of WD-repeat proteins. Nature. doi:10.1038/371297a0

Nicholas, P., Money, N.P., 1989. Osmotic Pressure of Aqueous Polyethylene Glycols : Relationship between Molecular Weight and Vapor Pressure Deficit. Plant Physiol. 91, 766–9.

Osakabe, Y., Osakabe, K., Shinozaki, K., Tran, L.-S.P., 2014. Response of plants to water stress. Front. Plant Sci. 5, 86. doi:10.3389/fpls.2014.00086

Ouyang, Y., Huang, X., Lu, Z., Yao, J., 2012. Genomic survey, expression profile and co-expression network analysis of OsWD40 family in rice. BMC Genomics 13, 100. doi:10.1186/1471-2164-13-100

Papazoglou, E.G., Karantounias, G.A., Vemmos, S.N., Bouranis, D.L., 2005. Photosynthesis and growth responses of giant reed (Arundo donax L.) to the heavy metals Cd and Ni. Environ. Int. 31, 243–9. doi:10.1016/j.envint.2004.09.022

Papazoglou, E.G., Serelis, K.G., Bouranis, D.L., 2007. Impact of high cadmium and nickel soil concentration on selected physiological parameters of Arundo donax L. Eur. J. Soil Biol. 43, 207–215. doi:10.1016/j.ejsobi.2007.02.003

Peck, S.C., Teisberg, T.J., Peck, S.C., Teisberg, T.J., 2016. CETA : A Model for Carbon Emissions Trajectory Assessment Published by : International Association for Energy Economics Stable URL : http://www.jstor.org/stable/41322454 Accessed : 15-04-2016 05 : 36 UTC CETA : A Model for Carbon Emissions Trajectory Assessment * 13, 55–77.

Perdue, R.E., 1958. Arundo donax—Source of musical reeds and industrial cellulose. Econ. Bot. 12, 368–404. doi:10.1007/BF02860024

Pfaffl, M.W., 2001. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. 29, e45. doi:10.1093/nar/29.9.e45

Pfaffl, M.W., Tichopad, A., Prgomet, C., Neuvians, T.P., 2004. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper--Excel-based tool using pair-wise correlations. Biotechnol. Lett. 26, 509–515. doi:10.1023/B:BILE.0000019559.84305.47

Pilu, R., Bucci, A., Badone, F.C., Landoni, M., 2012. Giant reed (Arundo donax L.): A weed plant or a promising energy crop? AFRICAN J. Biotechnol. 11, 9163–9174. doi:10.5897/AJB11.4182

Pilu, R., Cassani, E., Landoni, M., Badone, F.C., Passera, A., Cantaluppi, E., Corno, L., Adani, F., 2013a. Genetic characterization of an Italian Giant Reed (Arundo donax L.) clones collection: exploiting clonal selection. Euphytica. doi:10.1007/s10681-013-1022-z

Pilu, R., Manca, A., Landoni, M., Agrarie, S., 2013b. Arundo donax as an energy crop: pros and cons of the utilization of this perennial plant. Maydica.

Porter, W.C., Barsanti, K.C., Baughman, E.C., Rosenstiel, T.N., 2012. Considering the air quality impacts of bioenergy crop production: a case study involving Arundo donax. Environ. Sci. Technol. 46, 9777–84. doi:10.1021/es3013084

Qin, F., Sakuma, Y., Tran, L.-S.P., Maruyama, K., Kidokoro, S., Fujita, Y., Fujita, M., Umezawa, T., Sawano, Y.,

Miyazono, K.-I., Tanokura, M., Shinozaki, K., Yamaguchi-Shinozaki, K., 2008. Arabidopsis DREB2A-interacting proteins function as RING E3 ligases and negatively regulate plant drought stress-responsive gene expression. Plant Cell 20, 1693–1707. doi:10.1105/tpc.107.057380

Quinn, L.D., Holt, J.S., 2007. Ecological correlates of invasion by Arundo donax in three southern California riparian habitats. Biol. Invasions 10, 591–601. doi:10.1007/s10530-007-9155-4

Rossa, B., Tuffers, a V, Naidoo, G., von Willert, D.J., 1998. Arundo donax L. (Poaceae) - a C3 species with unusually high photosynthetic capacity. Bot. Acta 111, 216–221. doi:10.1111/j.1438-8677.1998.tb00698.x

Sablok, G., Fu, Y., Bobbio, V., Laura, M., Rotino, G.L., Bagnaresi, P., Allavena, A., Velikova, V., Viola, R., Loreto, F., Li, M., Varotto, C., 2014. Fuelling genetic and metabolic exploration of C 3 bioenergy crops through the first reference transcriptome of Arundo donax L. Plant Biotechnol. J. n/a-n/a. doi:10.1111/pbi.12159

Sagehashi, M., Liu, C., Fujii, T., Fujita, H., Sakai, Y., Hu, H.-Y., Sakoda, A., 2011. Cadmium Removal by the Hydroponic Culture of Giant Reed (Arundo donax) and Its Concentration in the Plant. J. Water Environ. Technol. 9, 121–127. doi:10.2965/jwet.2011.121

Sakuma, Y., 2006. Functional Analysis of an Arabidopsis Transcription Factor, DREB2A, Involved in Drought-Responsive Gene Expression. Plant Cell 18, 1292–1309. doi:10.1105/tpc.105.035881.1

Sánchez, E., Scordia, D., Lino, G., Arias, C., Cosentino, S.L., Nogués, S., 2015. Salinity and Water Stress Effects on Biomass Production in Different Arundo donax L. Clones. BioEnergy Res. doi:10.1007/s12155-015-9652-8

Sarry, J.E., Kuhn, L., Ducruix, C., Lafaye, A., Junot, C., Hugouvieux, V., Jourdain, A., Bastien, O., Fievet, J.B., Vailhen, D., Amekraz, B., Moulin, C., Ezan, E., Garin, J., Bourguignon, J., 2006. The early responses of Arabidopsis thaliana cells to cadmium exposure explored by protein and metabolite profiling analyses. Proteomics 6, 2180–2198. doi:10.1002/pmic.200500543

Schroeder, D.F., Gahrtz, M., Maxwell, B.B., Cook, R.K., Kan, J.M., Alonso, J.M., Ecker, J.R., Chory, J., 2016. De-Etiolated 1 and Damaged DNA Binding Protein 1 Interact to Regulate Arabidopsis Photomorphogenesis. Curr. Biol. 12, 1462–1472. doi:10.1016/S0960-9822(02)01106-5

Seo, K.-I., Lee, J.-H., Nezames, C.D., Zhong, S., Song, E., Byun, M.-O., Deng, X.W., 2014. ABD1 Is an Arabidopsis DCAF Substrate Receptor for CUL4-DDB1-Based E3 Ligases That Acts as a Negative Regulator of Abscisic Acid Signaling. Plant Cell 26, 695–711. doi:10.1105/tpc.113.119974

Shakoor, N., Nair, R., Crasta, O., Morris, G., Feltus, A., Kresovich, S., 2014. A Sorghum bicolor expression atlas reveals dynamic genotype-specific expression profiles for vegetative tissues of grain, sweet and bioenergy sorghums. BMC Plant Biol. 14, 35. doi:10.1186/1471-2229-14-35

Sharkey, T.D., Wiberley, A.E., Donohue, A.R., 2008. Isoprene emission from plants: why and how. Ann. Bot. 101, 5–18. doi:10.1093/aob/mcm240

Silver, N., Best, S., Jiang, J., Thein, S.L., 2006. Selection of housekeeping genes for gene expression studies in human reticulocytes using real-time PCR. BMC Mol. Biol. 7, 33. doi:10.1186/1471-2199-7-33

Stirnimann, C.U., Petsalaki, E., Russell, R.B., Müller, C.W., 2010. WD40 proteins propel cellular networks. Trends Biochem. Sci. 35, 565–574. doi:10.1016/j.tibs.2010.04.003

Stone, S.L., 2014. The role of ubiquitin and the 26S proteasome in plant abiotic stress signaling. Front. Plant Sci. 5, 135. doi:10.3389/fpls.2014.00135

Swaminathan, K., Chae, W.B., Mitros, T., Varala, K., Xie, L., Barling, A., Glowacka, K., Hall, M., Jezowski, S., Ming, R., Hudson, M., Juvik, J.A., Rokhsar, D.S., Moose, S.P., 2012. A framework genetic map for Miscanthus sinensis from RNAseq-based markers shows recent tetraploidy. BMC Genomics 13, 142. doi:10.1186/1471-2164-13-142

Tarin, D., Pepper, A.E., Goolsby, J.A., Moran, P.J., Arquieta, A.C., Kirk, A.E., Manhart, J.R., 2013. Microsatellites Uncover Multiple Introductions of Clonal Giant Reed (Arundo donax). Invasive Plant Sci. Manag. 6, 328–338. doi:10.1614/IPSM-D-12-00085.1

Tester, M., Langridge, P., 2010. Breeding Technologies to Increase Crop Production in a Changing World. Science (80-.). 327, 818–822. doi:10.1126/science.1183700

Tuteja, N., Sopory, S.K., 2008. Chemical signaling under abiotic stress environment in plants. Plant Signal. Behav. 3, 525–536. doi:10.4161/psb.3.8.6186

Unamba, C.I.N., Nag, A., Sharma, R.K., 2015. Next Generation Sequencing Technologies: The Doorway to the Unexplored Genomics of Non-Model Plants. Front. Plant Sci. 6, 1074. doi:10.3389/fpls.2015.01074

Valasek, M.A., Repa, J.J., 2005. The power of real-time PCR. AJP Adv. Physiol. Educ. 29, 151–159.

doi:10.1152/advan.00019.2005.

van Nocker, S., Ludwig, P., 2003. The WD-repeat protein superfamily in Arabidopsis: conservation and divergence in structure and function. BMC Genomics 4, 50. doi:10.1186/1471-2164-4-50

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., Speleman, F., 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol. 3, RESEARCH0034. doi:10.1186/gb-2002-3-7-research0034

Virtue, J., Reynolds, T., Malone, J., 2010. Managing the weed risk of cultivated Arundo donax L., in: Seventeenth Australasian Weeds Conference.

Wang, C., Dong, X., Han, L., Su, X.D., Zhang, Z., Li, J., Song, J., 2016. Identification of WD40 repeats by secondary structure-aided profile-profile alignment. J. Theor. Biol. 398, 122–129. doi:10.1016/j.jtbi.2016.03.025

Wang, Y., Jiang, F., Zhuo, Z., Wu, X.H., Wu, Y.D., 2013. A Method for WD40 Repeat Detection and Secondary Structure Prediction. PLoS One 8, 1–13. doi:10.1371/journal.pone.0065705

Xie, F., Xiao, P., Chen, D., Xu, L., Zhang, B., 2012. miRDeepFinder: A miRNA analysis tool for deep sequencing of plant small RNAs. Plant Mol. Biol. 80, 75–84. doi:10.1007/s11103-012-9885-2

Yang, Z., Chen, Y., Hu, B., Tan, Z., Huang, B., 2015. Identification and validation of reference genes for quantification of target gene expression with quantitative real-time PCR for tall fescue under four abiotic stresses. PLoS One 10, 1–15. doi:10.1371/journal.pone.0119569

Yim, A.K.-Y., Wong, J.W.-H., Ku, Y.-S., Qin, H., Chan, T.-F., Lam, H.-M., 2015. Using RNA-Seq Data to Evaluate Reference Genes Suitable for Gene Expression Studies in Soybean. PLoS One 10, e0136343. doi:10.1371/journal.pone.0136343

Yoshida, T., Mogami, J., Yamaguchi-Shinozaki, K., 2014. ABA-dependent and ABA-independent signaling in response to osmotic stress in plants. Curr. Opin. Plant Biol. 21, 133–139. doi:10.1016/j.pbi.2014.07.009

Zhang, Y., Schroeder, D.F., 2010. Effect of overexpression of Arabidopsis Damaged DNA-binding protein 1A on De-etiolated 1. Planta 231, 337–348. doi:10.1007/s00425-009-1056-6

Chapter 5.

General conclusion and discussion

Global warming is becoming a predominant issue for worldwide countries that have to face the consequences of weather anomalies such as prolonged drought period, heavy storms or floods just to cite some. Therefore, a constant effort to reduce the anthropogenic GHG emissions is ongoing through the implementation of multiple actions that insist on the CO_2 sources. An important sector scientists are giving attention to is the production of renewable energy with low or no CO_2 emission in the atmosphere. The production of fuel from plant biomasses (so called biofuel) is the most promising approach to reduce the GHG emission deriving from road transportation. In Europe, several species have been considered as suitable energy crop and *A. donax* appears as the best choice for the Mediterranean area. With this work, we explored the response of the plant to drought stress trhough the analysis of the whole transcriptome that not only enhance our general knowledge about the molecular mechanisms of stress response in this species but also gives the opportunity to identify and characterize in depth candidate genes whose function is still unknown. This was the case of *AdDWD1* gene which was literally unknown in Arundo and in its closely related species (rice, maize or sorghum) and only little was known in the model plant *Arabidopsis*. In this study, *AdDWD1* was caracterized both in *Arundo* or close related species (rice, maize or sorghum) and in the model plant *Arabidopsis* as an heterologous system. The selection of this gene was based on the preliminary general information about the importance of the WD40 family in plants and specifically on the role of DWD subfamily group on stress response. The results reported here showed a high involvement of *AdDWD1* during simalted drought and salt stresses making it a good candidate for further studies also in other species. Moreover, due to the importance of quantitative real-time PCR in molecular biology we validated a set of reliable genes to be used as reference in relative quantification approach. They will likely constitute a valuable resource for further studies on abiotic stress in *A. donax*.

In conclusion, this work sheds light for the first time into the genetic response of *A. donax* to osmotic stress, developing valuable tools and information for the improvement of the species for biofuel production. Our results could be the foundation for further studies addressig the unique genetics basis of adapative and yield-related traits of giant reed, towards the full exploitation of this semi-domesticated plant as a crop.

Chapter 6.

Acknowledgment

I whish to thanks all the people that have supported me during these years, starting from **Yuan** and the small **Gaia**. A huge thanks also to my supervisors **Claudio** and **Silvio** that allowed me to start and finish this project and especially support me along the way. Many other people deserve many thanks: my mother, **Marta**, to be always availbale to help in any way, my brother **Luca** and sister **Elisa** (**Martino**, **Valentina**, **Filippo**) that also helped much, my dad, **Paolo**, that would be happy for what I have done here. In the lab, thanks to **Mingai** and **Enrico** (**Eliana, Davide, Arianna**) for the material and motivational support. Also, all the collegues that I met and from which I learnd (and hopefully taught) something: the VIP **Mastaneh**, **Yu**, **Luisa**, **Ilaria**, **Huan**, **Jia** and than "Master" **Josè** (**Hannah** and **Jonah**), the "great mind" **Shili** (**Hui** and **Suci**), "il filosofo" **Luca**, **Ginco**, **Bo** and all the volleyball team (especially **Chida**). Big thanks also to my "old" friends **Alessandro** and **Roberto**, **Valentina**, **Martina**, **Stefano**, **Grazia** because they are still my friends. As last, I whish to thanks also **Emma** and **Mauro** for their help and the great barbeques! Sorry if I forgot someone!

The project has been funded with support from **Fondazione Edmund Mach**.

Chapter 7.

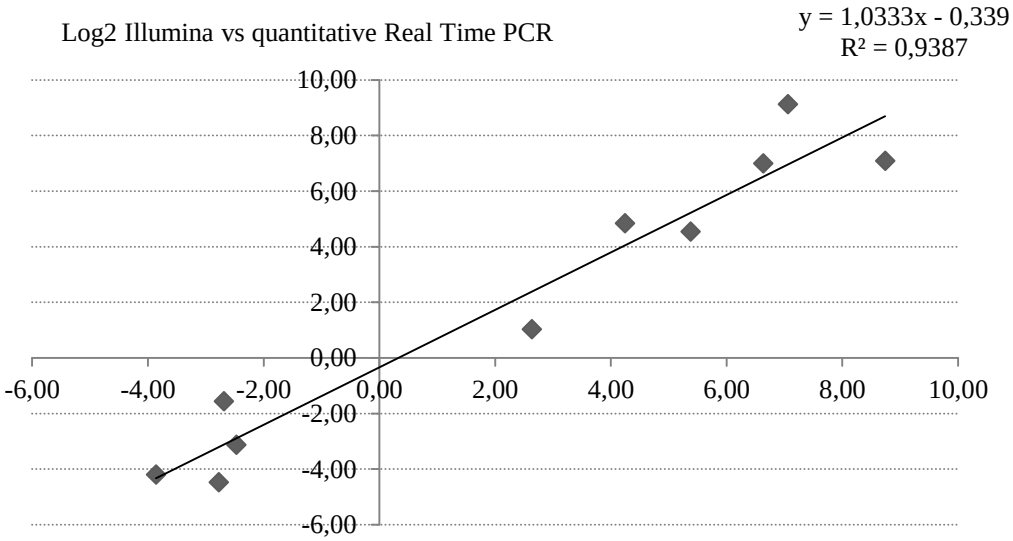
Supplementary material.

7.1 Arundo donax L. transcriptome under osmotic stress.

Library name	Number of Sequencing Reads*	Treatment/Condition	Organ
Ad_Control_1	47792849	Untreated control	Shoot
Ad_Control_2	35432010	Untreated control	Shoot
Ad_Control_3	32186149	Untreated control	Shoot
Ad_PEG10%_1	38952041	10% PEG	Shoot
Ad_PEG10%_2	34331086	10% PEG	Shoot
Ad_PEG10%_3	30234460	10% PEG	Shoot
Ad_PEG20%_1	32648859	20% PEG	Shoot
Ad_PEG20%_2	34100100	20% PEG	Shoot
Ad_PEG20%_3	38857109	20% PEG	Shoot
Ad_Control_1	41912793	Untreated control	Root
Ad_Control_2	35027592	Untreated control	Root
Ad_Control_3	35234162	Untreated control	Root
Ad_PEG10%_1	42457897	10% PEG	Root
Ad_PEG10%_2	42854988	10% PEG	Root
Ad_PEG10%_3	43240510	10% PEG	Root
Ad_PEG20%_1	40387040	20% PEG	Root
Ad_PEG20%_2	42169350	20% PEG	Root
Ad_PEG20%_3	46354173	20% PEG	Root

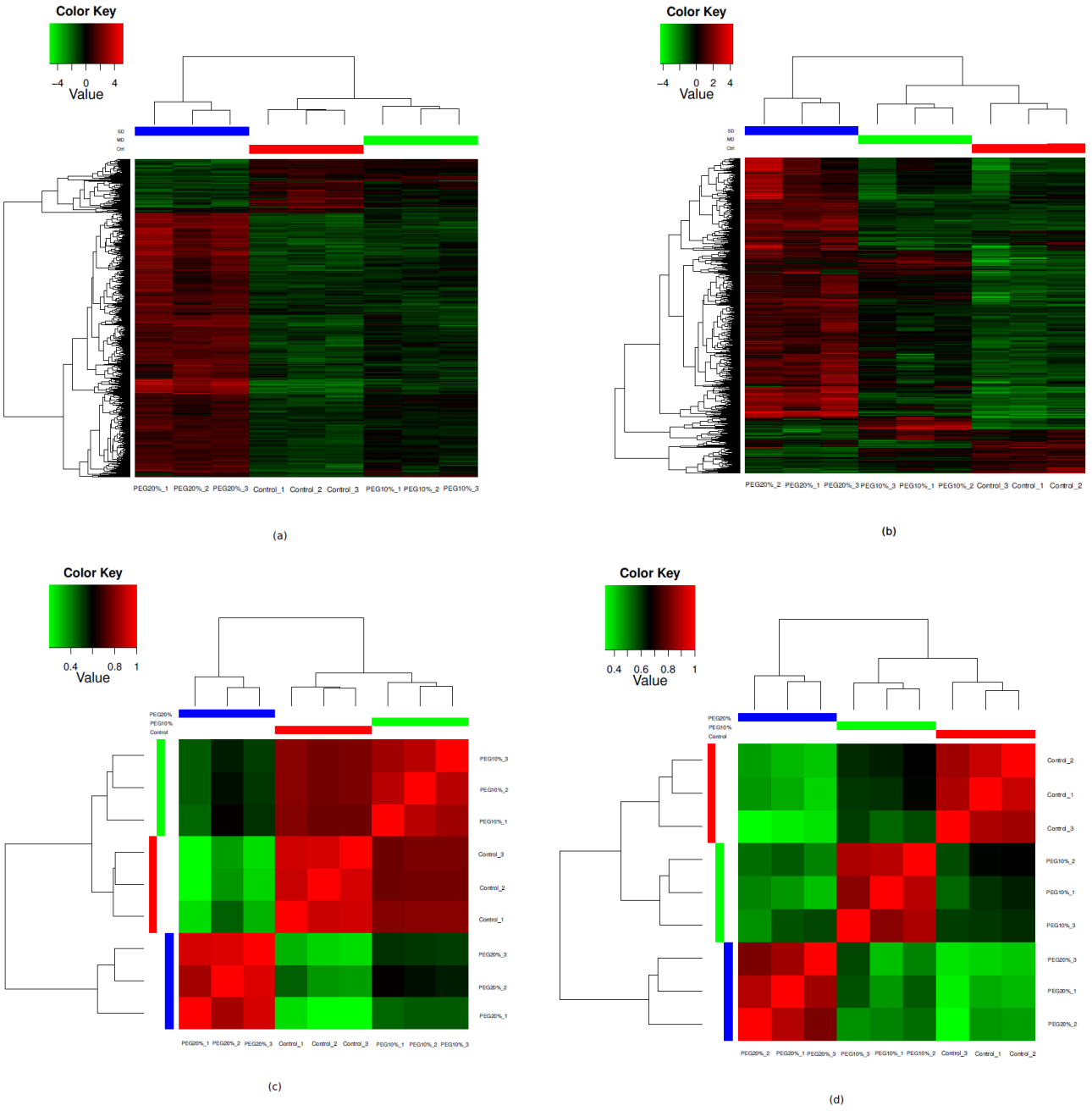
Note: * 2x100 PE

Supplementary File 2.1. **Experimental design and number of reads per library**.



Organ/Regulation	Gene ID	Primer Forward	Primer Reverse
Shoot_up	Arundo_WS_Unigene028474	AACCTCCCTCATTTCCTCGAAC	TTTTGCGCTGATACGTGTGG
	Arundo_WS_Unigene065161	AACCATCGTTGCTTCGCTTG	TGGCTAACCAACAACAACCG
Shoot_down	Arundo_WS_Unigene045711	ACTGACAGGCTGAGTTCGAC	AGCACAGCTGGCATTCAGAT
	Arundo_WS_Unigene074110	GTCTCTTCTTTGTGCCATCTGC	AGCAGCGATCTCCTTTATCTGG
Root_up	Arundo_WS_Unigene011476	GCAGCACGACTACTACTTTCAC	ATGGCATGCAACAGCAATCC
	Arundo_WS_Unigene053771	AATCGTGGCGTTGTATCGTG	TTGGCGACCAAATTCATGCC
	Arundo_WS_Unigene020231	GATGAAATGGGGATGGGTGAAC	TGCTTTTGCCAGCATTGTCC
	Arundo_WS_Unigene023719	ACAAGGGAACAGGTTCAGGTC	CGCCAACGAAATGCCAAAAG
Root_down	Arundo_WS_Unigene042884	AGGGAAAGCTTGACACGATG	ATGGTACCATTGCGTCTTGG
	Arundo_WS_Unigene077521	GCTCGTCGAACTTCTTGGTG	ACCTTGCTAGTGACCGTATCAG
All	AdoACT	TCTTGGCTTGCATTCTTGGG	TGGATTGCGAAGGCTGAGTAC

Supplementary File 2.2. **Validation of 10 A. donax DEGs by Real Time qrt-PCR**.



Supplementary File 2.3. **Visualization of differentially expressed genes**. Heat maps visualization of the differentially expressed genes between three conditions for shoots (a) and roots (b) using Euclidean distances between TMM normalized expression values. Expression levels for genes in each cDNA library were measured as fragments per kilobase per million reads (FPKM), and color-coded from green (lowly expressed) to red (highly expressed). Hierarchical clustering level at individual libraries is represented by the dendrogram for shoots (c) and roots (d), and color-coded from green (weak correlation) to red (strong correlation).

Additional file 4_
Table S3.xlsx

Supplementary File 2.4. **Stress-related A. donax DEGs**. See online version at http://www.biotechnologyforbiofuels.com/content/9/1/54.

Additional file 5_
Table S4.xlsx

Supplementary File 2.5. **Functional annotation, fold change values and statistical significance of DEGs**. See online version at *http://www.biotechnologyforbiofuels.com/content/9/1/54*.

Additional file 6_
Tables S5.xlsx

Supplementary File 2.6. **Slimmed GO terms overrepresented under different conditions**. See online version at http://www.biotechnologyforbiofuels.com/content/9/1/54.

Additional file 7_
Table S6.xlsx

Supplementary File 2.7. **Transcription factors responsive to water stress in A. donax.** See online version at http://www.biotechnologyforbiofuels.com/content/9/1/54.

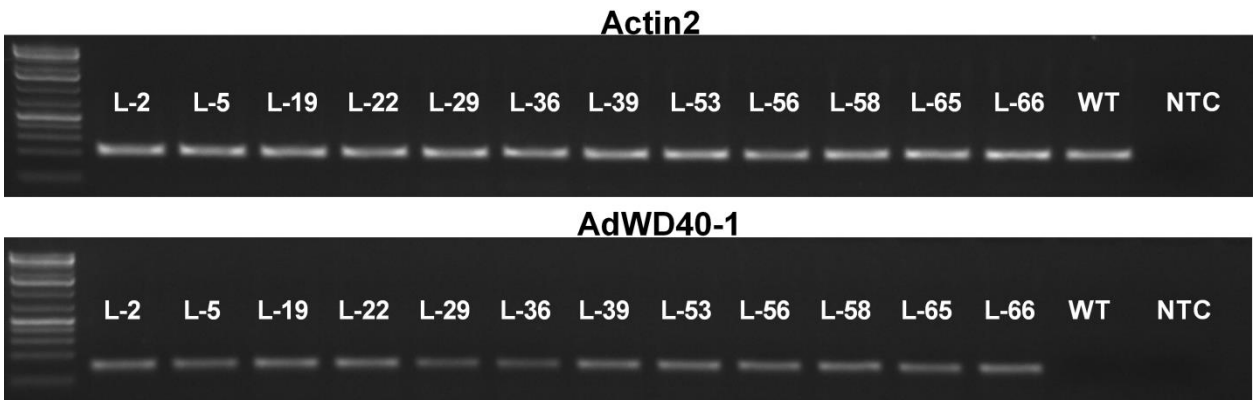
Additional file 8_
Table S7.xlsx

Supplementary File 2.8. **Comparison of co-regulated expression modules responsive to water stress in A. donax and rice**. See online version at http://www.biotechnologyforbiofuels.com/content/9/1/54.

Additional file 9_
Table S8.xlsx

Supplementary File 2.9. **Drought responsive orthologs identified among A. donax, foxtail, sorghum and rice.** See online version at http://www.biotechnologyforbiofuels.com/content/9/1/54.

7.2 Characterization of a stress-related Arundo gene.



Supplementary File 3.1. **T2 lines of 35S::AdWD40-1 Arabidopsis plants**. WT, wild type Arabidopsis, ecotype Columbia; NTC, no template control.

A

Repeats	Score	Start	End	Strand_d	Loop_da	Strand_a	Loop_ab	Strand_b	Loop_bc	Strand_c	Loop_cd	H_bonds	Hotspots_on_the_top_face
WD1	40.64	154	194	AVKSTI	AHFQ	LRNLLW	ATSKH	DVYSVQ	NY	SVMHWS	PLLKRG	NA	Q163 R165
WD2	51.01	195	251	KEVLDV	ASQLVPAEE AHGARPLSR	VQVSTM	ALKDN	LIVAGG	FQGE	LICKYV	DKPGVA	NA	R218 Q220
WD3	54.75	252	296	FCTNLT	GNNNSI	TNAVDI	YQTPN GAT	RVTAAN	NDC	VVRTFD	TERF	NA	I263 N265 N284
WD4	69.48	297	337	SLLSQF	TFPWS	VNNTSV	SPDGK	LLAVLG	DSS	DCLIAD	PNSG	NA	N309
WD5	138.48	338	380	KEVATL	RGHLDY	SFSSAW	HPGGH	VLATGN	QDA	TCRLWD	VRNPS	penta d	Y349 F351 Q367
WD6	87.56	381	423	RSFAVL	GGRIGA	VRGLRF	SADGR	FLAAAE	AAD	FVHVYD	AAAGY	NA	R394
WD7	60.16	424	463	GAEQEV	DLFGE	ISGAAF	SPDGET	LFVSVA	DRTYG	GLLEFR		NA	E434

B

Query sequence (Arundo_Drought_Unigene032775, Len=476):

```
(BLUE - beta-strands, RED - alpha-helices, predicted by PSIPRED)
                        10        20        30        40        50        60        70        80
                ....*....|....*....|....*....|....*....|....*....|....*....|....*....|....*....|
Query         1 LISLQRFPRSLPCPALFPTSSAQLATSPVGACQPMAHDLRDDLELVAADGDYYYFDDFDADDGFRTAGASEEGGATDQNK 80
WD40-rpts

Query        81 QMDDTSALDYKEGKDMQGIPWERLNYSRDQYREMRLKQYKNYQSLARSRSGLEKECKQVQRKDTFFDFHFNTRAVKSTIA 160
WD40-rpts                                                                                (__________

Query       161 HFQLRNLLWATSKHDVYSVQNYSVMHWSPLLKRGKEVLDVASQLVPAEEAHGARPLSRVQVSTMALKDNLIVAGGFQGEL 240
WD40-rpts       ______WD 1_________________)

Query       241 ICKYVDKPGVAFCTNLTGNNNSITNAVDIYQTPNGATRVTAANNDCVVRTFDTERFSLLSQFTFPWSVNNTSVSPDGKLL 320
WD40-rpts                (____________________WD 2____________________) (____________________WD 3_____

Query       321 AVLGDSSDCLIADPNSGKEVATLRGHLDYSFSSAWHPGGHVLATGNQDATCRLWDVRNPSRSFAVLGGRIGAVRGLRFSA 400
WD40-rpts       _____________)    (_________________WD 4_________________)     (_________________WD

Query       401 DGRFLAAAEAADFVHVYDAAAGYGAEQEVDLFGEISGAAFSPDGETLFVSVADRTYGGLLEFRRRRSYGYLDAYFY 476
WD40-rpts       5_________________)    (_________________WD 6_________________)
```

6 WD40-repeats found in QUERY=Arundo_Drought_Unigene032775 (Len=476):

(Red - Z-score>4.0)

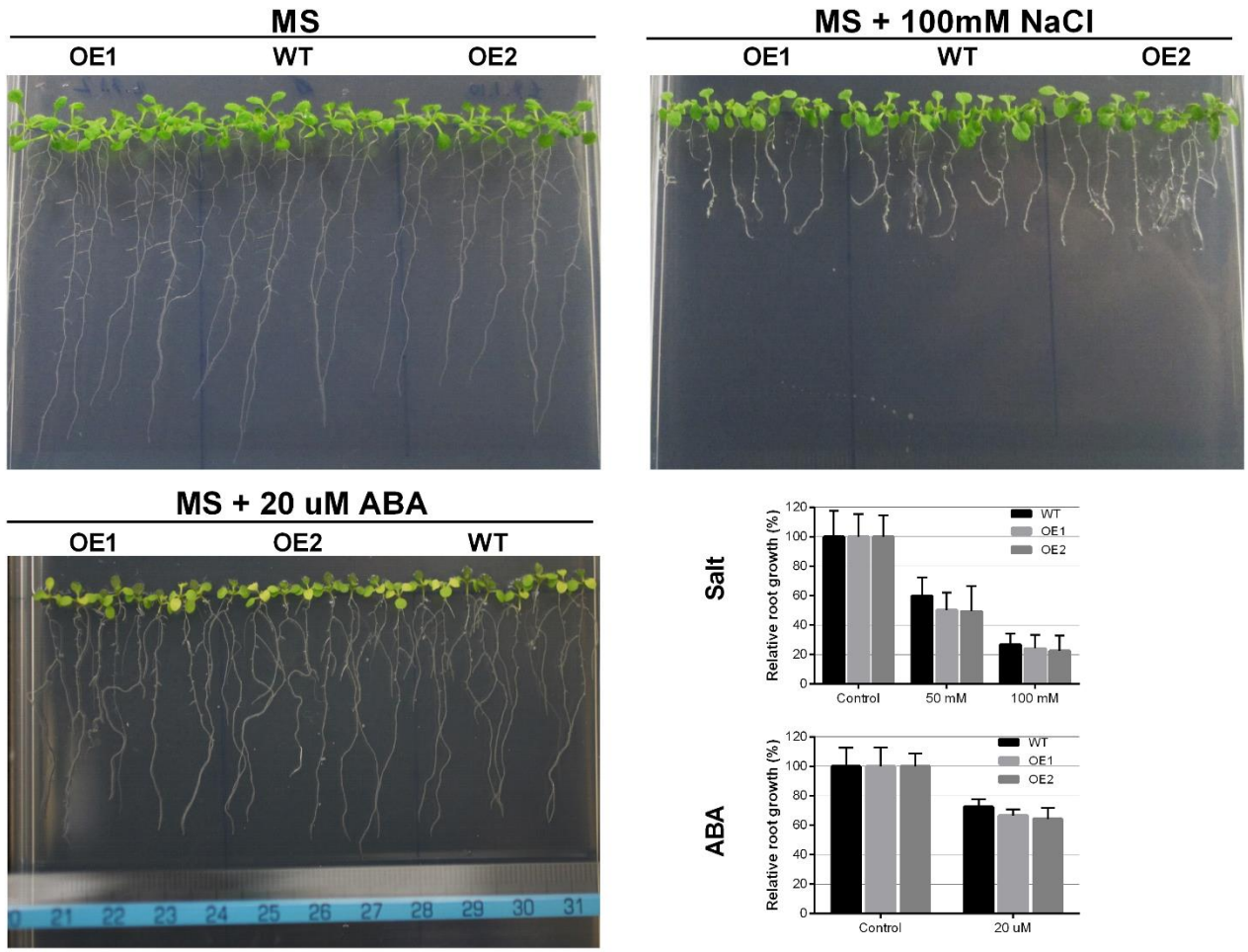
Rpts	Segment	Boundary	Z-score	Score	P-value	Identities	SS-score
WD 1:	Seg 2	149-188	4.26	290.52	0.00755	10.0%	37.43
WD 2:	Seg 11	250-292	6.63	388.61	0.00036	20.0%	55.70
WD 3:	Seg 15	293-333	7.85	439.00	7.5e-05	27.5%	57.81
WD 4:	Seg 19	337-375	10.22	536.95	3.6e-06	35.0%	53.26
WD 5:	Seg 23	380-418	9.05	488.68	1.6e-05	25.0%	56.05
WD 6:	Seg 27	422-461	7.45	422.27	0.00013	22.5%	54.70

Multiple sequence alignment of 6 WD40-repeats:

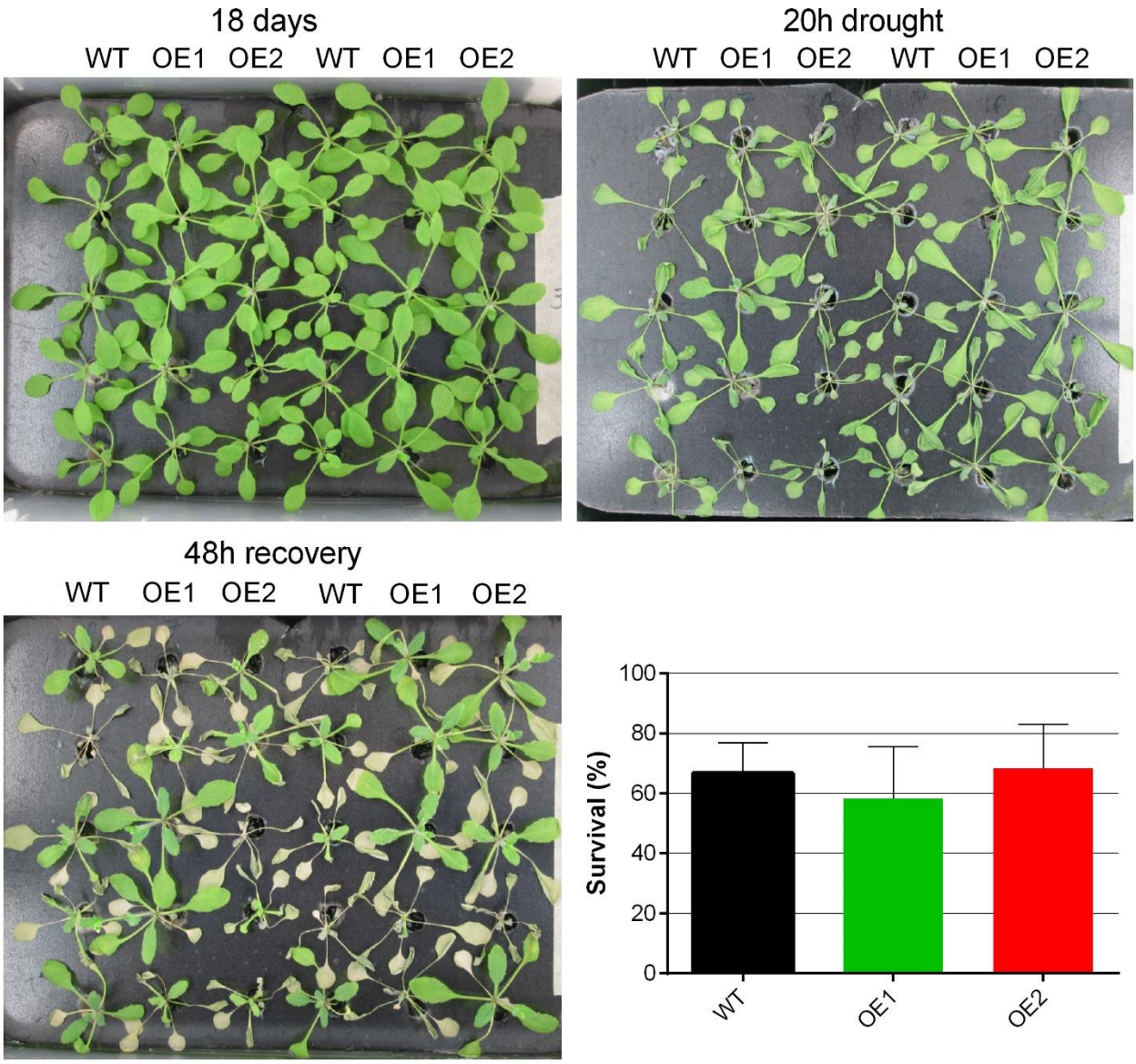
```
(BLUE - beta-strands, RED - alpha-helices, predicted by PSIPRED)
   WD 1  149-188    HFN-TRAVKSTIAHFQLR---NLLWATSK-HDVYSVQNYSV-MHWS
   WD 2  250-292    -VA-FCTNLTGNNNSITNAVDIYQTPNGATRVTAANNDCVV-RTFD
   WD 3  293-333    TERFSLLSQFTFPWSVNN---TSVSPDGK-LLAVLGDSSDC-LIAD
   WD 4  337-375    -GK-EVATLRGHLDYSFS---SAWHPGGH-VLATGNQDATC-RLWD
   WD 5  380-418    -SR-SFAVLGGRIGAVRG---LRFSADGR-FLAAAEAADFV-HVYD
   WD 6  422-461    -GY-GAEQEVDLFGEISG---AAFSPDGE-TLFVSVADRTYGGLLE

   WD40 Template    TGE-CLRTLSGHTSGVTS---VSFSPDGK-RLVSGSEDGTI-KVWD
```

Supplementary File 3.2. **WD40 repeats predicted from online software** WDSP (A) (http://wu.scbb.pkusz.edu.cn/wdsp/index.jsp) and WRDD (B) (http://protein.cau.edu.cn/wdrr).



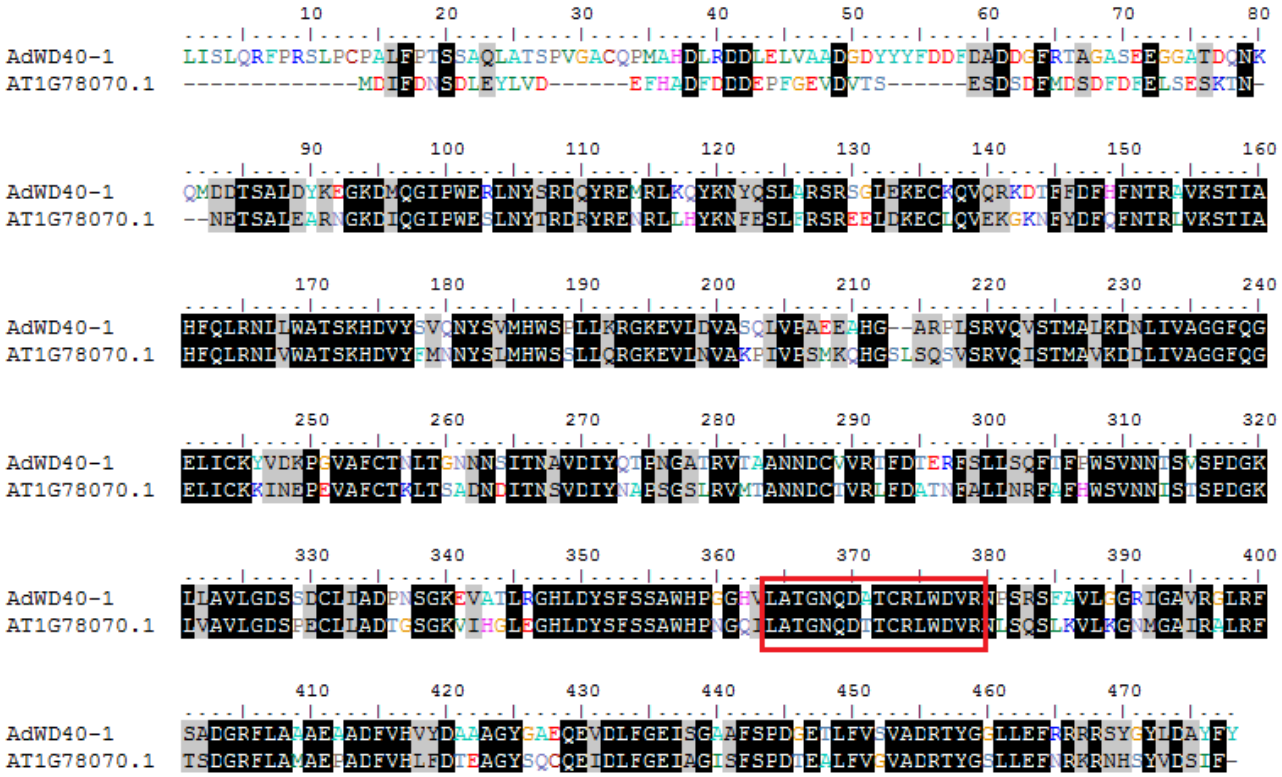
Supplementary File 3.3. **Root growth of WT and two overexpressing lines OE1 and OE2**. Plants have been germinated for 4 days in standard half strength MS medium and transfer on vertical plates containing MS (A), MS + 100 mM NaCl (B) or MS + 20 µM ABA (C). Graphs illustrate the growth length compare to controls.



Supplementary File 3.4. **Air exposed Arabidopsis WT and OE lines behave similar**. Plants were grown 18 days, exposed to air for 20 hours and then recovered for 48 hours. Experiment consist of 6 independent replicas.

	Primer name	Primer sequence 5'-3'	Amplicon lenght	Reference
Arundo donax cloning	Ad_Uni032775_2F	CACCAGCTCGCAACTAGTCCAGTCGG	1322	
	Ad_Uni032775_2R	CGGCTCATCACTCGATCACC		
Arabidopsis thaliana promoter cloning	At_WD40_prom_F	CACCGAGATTTGGTTCCTTTAATTTGGACAC	1853	
	At_WD40_prom_R	CTTTAAACCTTTGGCAATAATCAAAATCG		
qRT-PCR	Ad_AC1_F*	TCTTGGCTTGCATTCTTGGG	93	Fu et al.; 2016
	Ad_AC1_R*	TGGATTGCGAAGGCTGAGTAC		
	Ad_GAPDH_F*	TGACAAGGAGAAGGCTGCTG	167	
	Ad_GAPDH_R*	GAGCAAGGCAGTTTGTGGT		
	Ad_Uni032775_rt_F	TAGCGCATTTTCAGCTGAGG	139	
	Ad_Uni032775_rt_R	ACCAATTGGCTGGCAACATC		
	At_Actin2_F	CAAGGCCGAGTATGATGAGG	228	Seo et al.; 2014
	At_Actin2_R	GAAACGCAGACGTAAGTAAAAAC		
	ABI1_RT_F	AGAGTGTGCCTTTGTATGGTTTTA	206	Kong et al.; 2014
	ABI1_RT_R	CATCCTCTCTCTACAATAGTTCGCT		
	AtSOS1_RT_F	GCAAACACTTTGATATTTATCCTCAG	89	Luo et al.; 2016
	AtSOS1_RT_R	CATGAATTCCCTTGGTAGGC		
	AtSOS3_RT_F	CATTCACGGTAGAAGAAGTGGA	230	Kong et al.; 2014
	AtSOS3_RT_R	GCTTGGATGGAAGACACCTAA		
	COR15A_RT_For	GATACATTGGGTAAAGAAGCTGAGA	199	Lim et al.; 2015
	COR15A_RT_Rev	ACATGAAGAGAGAGGATATGGATCA		
	DREB2A_RT_For	CTACAAAGCCTCAACTACGGAATAC	175	Lim et al.; 2015
	DREB2A_RT_Rev	AAACTCGGATAGAGAATCAACAGTC		
	P5CS_RT_For	GAAGGATTACTTACAACGAGATGGA	155	Lim et al.; 2015
	P5CS_RT_Rev	CTCTCCTCAAGTCTCAACCAAATAC		
	RD29A_RT_2F	GGAAGTGAAAGGAGGAGGAGGAA	134	Kong et al.; 2014
	RD29A_RT_2R	CACCACCAAACCAGCCAGATG		

Supplementary File 3.5. **List of primers**. Asterisks indicate the genes used as reference in real-time qPCR.



Supplementary File 3.6. Pairwise alignment of AdWD40-1 and its Arabidopsis homolog At1g78070. Shading made with BLOSUM62 matrix. DWD box is marked in red.

7.3 Arundo donax qRT-PCR reference gene selection for abiotic stress.

Gene	Plant resource	Original accession num.	Arundo donax best hit	Score (bits)	E Value	Identities	Arundo primers F/R (5'-3')
AC1 Actin	Sorghum bicolor	Sb01g010030	Unigene036290	749	0.0	587/655 (89%)	TCTTGGCTTGCATTC TTGGG TGGATTGCGAAGGCT GAGTAC
Act2 Actin2	Setaria italica	Si026509m	Unigene057037	698	0.0	496/544 (91%), 347/399 (86%)	CGCATACGTGGCACT TGACT GGGCATCTGAACCTC TCTGC
EF-1α Elongation factor 1-alpha	Setaria italica	Si022039m	Unigene076509	1049	0.0	643/681 (94%)	TGACTGTGCTGTGCT CATCA GTTGCAGCAGCAGAT CATCT
GAPDH Glyceraldehyde-3-phosphate dehydrogenase 2	Sorghum bicolor	Sb07g002220	Unigene069707	1289	0.0	923/1014 (91%)	TGACAAGGAGAAGG CTGCTG GAGCAAGGCAGTTTG TGGTG
RPN6 26S proteasome non-ATPase regulatory subunit 11	Sorghum bicolor	Sb06g017780	Unigene067565	1855	0.0	1212/1304 (92%)	CACACGACTAGCAGC TTTCAAG TTCAAACGTCGGGAA GGTTG
pDUF221 Probable membrane protein DUF221-related Calcium-dependent channel	Arundo donax		Unigene070087				GACAAAGGAGTCAG CCGTCA AACGTGCTTCGGACT TGGAT
TLF	Setaria italica	Si000298m	Unigene076539	3640	0.0	2358/2532 (93%)	GACTTCATGGGTGGT GCTGA

Translation factor							TGTTTGTTGGGGGACTTGCT
TUB α Alpha tubulin	Setaria italica	Si029822m	Unigene068813	829	0.0	490/514 (95%), 440/498 (88%)	TACCAGCCACCCTCAGTTGT AGTCGAACTTGTGGTCAATGC
DREB2A * Dehydration-Responsive Element Binding Protein 2	Setaria italica	Si022619m	Unigene057213	416	e-115	327/366 (89%)	TCCAGCAGGTAGATCATCTCC AGCAGGTTCGGTAATAGGCA

Supplementary File 4.1. **Locus details, identities and original primers**. The modified single bases or new primers are underlined. Asterisc (*) indicates the validation gene.

	pDUF221	TLF	AC1	Act2	TUB α	EF-1α	RPN6	GAPDH
Control_shoot_1	26,84	24,18	23,00	22,49	22,81	22,02	25,04	19,53
Control_shoot_2	26,85	23,55	22,87	22,96	23,14	21,62	25,20	20,01
Control_shoot_3	27,99	24,20	23,27	23,71	23,76	22,14	25,65	20,08
Control_root_1	25,18	22,55	21,45	21,26	21,62	19,76	23,39	18,60
Control_root_2	25,20	21,44	21,85	20,93	21,37	19,02	23,20	18,23
Control_root_3	25,72	21,79	21,59	21,31	21,43	19,27	23,21	18,30
Drought_shoot_1h30'_1	26,19	22,68	22,40	22,27	21,96	20,51	24,00	19,09
Drought_shoot_1h30'_2	26,38	23,46	22,86	22,76	23,19	21,19	24,42	19,20
Drought_shoot_1h30'_3	26,41	23,76	23,58	23,33	23,53	21,53	24,72	19,79
Drought_shoot_3h_1	25,77	24,01	23,55	23,07	23,48	21,56	24,31	19,55
Drought_shoot_3h_2	25,58	24,06	23,44	23,59	23,60	21,86	24,79	20,11
Drought_shoot_3h_3	25,83	23,15	22,58	22,90	23,10	20,71	24,27	19,12
Drought_shoot_6h_1	25,25	23,96	23,42	23,56	22,93	21,51	24,97	19,85
Drought_shoot_6h_2	24,73	23,29	24,03	23,30	22,93	21,65	24,77	19,55
Drought_shoot_6h_3	25,14	24,15	24,56	23,95	23,57	22,09	24,83	20,06
Drought_shoot_11h_1	25,02	24,05	24,39	23,75	23,92	21,38	24,99	19,47
Drought_shoot_11h_2	24,58	24,22	24,01	23,22	23,35	21,31	24,82	19,35
Drought_shoot_11h_3	24,94	24,45	24,99	23,66	23,90	21,92	25,01	19,71
Drought_shoot_24h_1	26,14	24,93	24,90	22,93	24,59	22,90	25,28	19,30
Drought_shoot_24h_2	26,11	24,65	23,68	22,36	23,56	22,23	25,13	19,31
Drought_shoot_24h_3	25,62	24,26	23,36	21,95	23,21	21,76	24,83	19,08
Drought_root_1h30'_1	25,68	22,65	22,15	21,96	22,62	19,65	24,03	19,11
Drought_root_1h30'_2	25,56	22,18	21,90	21,52	22,04	19,35	23,80	18,69
Drought_root_1h30'_3	25,58	21,96	21,66	21,74	21,81	19,62	23,53	18,57

Drought_root_3h_1	24,98	22,64	24,15	22,24	23,45	19,48	23,87	18,84
Drought_root_3h_2	25,59	22,78	23,38	22,24	23,21	19,42	23,93	18,66
Drought_root_3h_3	26,05	23,41	23,74	23,01	23,75	20,75	24,55	19,24
Drought_root_6h_1	24,80	23,06	23,41	22,93	23,14	20,24	23,72	19,07
Drought_root_6h_2	25,16	22,94	24,94	22,96	23,89	20,50	24,14	18,59
Drought_root_6h_3	24,68	23,40	25,14	22,82	24,09	20,77	24,32	19,38
Drought_root_11h_1	23,83	21,47	24,27	22,36	23,36	20,04	24,03	18,49
Drought_root_11h_2	24,17	21,58	24,18	22,25	23,14	19,86	23,68	18,76
Drought_root_11h_3	24,39	22,31	23,91	22,34	23,40	20,28	24,18	19,42
Drought_root_24h_1	24,59	21,38	22,39	21,55	22,12	19,42	23,53	18,15
Drought_root_24h_2	25,10	21,77	22,13	21,86	22,21	19,19	23,26	18,25
Drought_root_24h_3	24,83	21,82	22,31	23,38	22,12	19,51	23,48	18,27
Cadmium_shoot_1h30'_1	26,26	22,10	22,11	21,71	21,72	21,02	24,87	19,14
Cadmium_shoot_1h30'_2	26,32	21,71	21,85	21,43	21,97	20,21	24,78	19,21
Cadmium_shoot_1h30'_3	26,75	21,94	21,76	21,75	22,12	20,34	24,79	19,55
Cadmium_shoot_3h_1	26,96	23,16	23,20	22,82	23,25	21,44	25,41	20,40
Cadmium_shoot_3h_2	26,54	22,86	23,28	22,61	22,77	21,73	25,29	19,80
Cadmium_shoot_3h_3	26,61	22,57	22,97	22,42	22,35	21,58	25,08	19,48
Cadmium_shoot_6h_1	25,44	21,95	22,94	22,72	22,37	20,65	24,79	19,78
Cadmium_shoot_6h_2	25,52	21,24	21,66	21,96	21,56	19,56	24,16	18,84
Cadmium_shoot_6h_3	25,53	22,31	23,20	22,98	22,48	21,58	25,20	19,87
Cadmium_shoot_11h_1	25,55	22,08	21,59	21,98	21,56	20,69	24,33	18,11
Cadmium_shoot_11h_2	25,51	22,43	22,24	22,24	21,90	21,29	24,71	19,36
Cadmium_shoot_11h_3	25,51	22,01	22,17	21,91	21,75	21,10	24,90	19,13
Cadmium_shoot_24h_1	26,21	22,16	21,82	21,12	21,92	21,19	24,78	19,01
Cadmium_shoot_24h_2	26,23	22,30	21,70	21,04	22,21	21,15	24,59	19,05
Cadmium_shoot_24h_3	26,85	23,66	22,21	22,08	22,85	22,39	25,27	19,65
Cadmium_root_1h30'_1	26,05	21,92	22,44	22,25	22,37	19,32	23,66	17,29
Cadmium_root_1h30'_2	25,58	22,13	22,94	22,53	22,32	19,62	23,61	17,62
Cadmium_root_1h30'_3	25,50	21,96	23,57	22,25	23,19	20,20	24,11	17,83
Cadmium_root_3h_1	25,51	22,60	24,15	22,68	23,19	20,92	24,42	18,31
Cadmium_root_3h_2	25,40	22,26	22,84	21,74	22,12	19,52	23,52	17,79
Cadmium_root_3h_3	25,63	22,76	24,33	22,17	23,48	19,88	23,99	18,31

Cadmium_root_6h_1	25,36	22,57	23,58	22,40	22,81	20,09	24,36	17,94
Cadmium_root_6h_2	25,50	22,06	22,80	21,75	21,96	19,28	23,71	17,14
Cadmium_root_6h_3	25,60	22,15	23,53	22,74	22,80	19,50	24,12	18,03
Cadmium_root_11h_1	25,87	22,17	23,68	22,36	23,15	19,92	23,45	17,06
Cadmium_root_11h_2	25,77	23,26	25,07	23,08	24,28	20,45	24,36	18,12
Cadmium_root_11h_3	25,62	22,05	23,44	21,92	22,94	19,05	23,12	16,16
Cadmium_root_24h_1	26,09	21,83	23,20	22,18	22,90	19,29	23,56	17,24
Cadmium_root_24h_2	26,17	22,68	24,40	22,03	23,22	19,96	23,47	17,49
Cadmium_root_24h_3	25,97	22,21	23,94	21,87	23,11	19,59	23,19	17,25
HeatShock_shoot_1h30'_1	25,79	23,96	24,55	21,93	24,18	21,99	25,12	18,53
HeatShock_shoot_1h30'_2	25,42	23,40	23,58	21,12	23,32	21,27	24,26	17,82
HeatShock_shoot_1h30'_3	25,65	23,66	24,24	22,08	24,04	21,83	24,92	18,26
HeatShock_shoot_3h_1	24,17	22,88	24,69	22,20	23,11	21,34	24,69	18,64
HeatShock_shoot_3h_2	24,29	23,22	24,24	22,13	23,05	21,20	24,76	18,78
HeatShock_shoot_3h_3	24,54	23,23	24,58	22,36	23,23	21,43	24,95	18,88
HeatShock_shoot_6h_1	24,09	21,85	23,67	22,61	22,15	20,88	24,33	19,09
HeatShock_shoot_6h_2	24,28	22,21	23,93	22,91	22,20	20,91	24,68	19,28
HeatShock_shoot_6h_3	24,68	22,89	23,86	22,99	22,38	21,17	24,68	19,65
HeatShock_shoot_11h_1	24,77	22,51	23,21	22,35	22,49	20,73	24,66	19,20
HeatShock_shoot_11h_2	24,79	22,27	22,99	22,09	22,11	20,65	24,49	19,19
HeatShock_shoot_11h_3	24,78	22,61	23,15	22,46	22,52	21,05	24,66	19,34
HeatShock_shoot_24h_1	25,78	23,27	23,14	21,41	23,01	21,45	24,82	19,05
HeatShock_shoot_24h_2	26,08	23,79	23,43	21,63	23,21	21,82	24,80	19,31
HeatShock_shoot_24h_3	26,25	23,57	23,24	21,56	22,76	21,51	24,64	19,06
HeatShock_root_1h30'_1	24,86	22,48	25,88	23,12	23,30	19,91	24,72	17,83
HeatShock_root_1h30'_2	24,62	22,48	26,22	21,85	24,37	20,17	24,57	17,84
HeatShock_root_1h30'_3	24,52	22,74	26,07	21,47	23,58	20,40	24,54	17,90
HeatShock_root_3h_1	24,07	23,71	26,05	22,12	23,47	20,90	24,67	18,80
HeatShock_root_3h_2	23,57	23,56	26,81	21,96	23,84	20,40	24,67	18,90
HeatShock_root_3h_3	23,55	23,74	27,08	21,66	24,28	20,49	24,53	18,80
HeatShock_root_6h_1	23,22	24,20	28,58	23,02	24,55	21,40	25,25	20,10
HeatShock_root_6h_2	23,32	23,40	26,73	22,18	24,31	20,68	24,79	19,50
HeatShock_root_6h_3	23,80	23,07	26,82	22,30	24,50	20,50	24,78	18,95

HeatShock_root_11h_1	24,16	24,06	26,58	22,46	24,17	20,78	25,30	19,48
HeatShock_root_11h_2	23,26	23,31	26,43	22,13	23,53	20,09	24,87	19,14
HeatShock_root_11h_3	24,19	23,41	24,84	21,69	23,06	20,31	25,08	18,93
HeatShock_root_24h_1	23,32	23,20	25,63	21,33	23,09	20,33	24,23	19,10
HeatShock_root_24h_2	23,48	23,55	25,53	21,19	23,30	20,46	24,53	19,30
HeatShock_root_24h_3	23,32	23,08	25,23	21,33	23,06	20,29	24,20	19,15

Supplementary File 4.2. **Raw Cq values for each sample and gene**.

	pDUF221	**TLF**	**AC1**	**Act2**	**TUB α**	**EF-1α**	**RPN6**	**GAPDH**
Average	25,276	22,822	23,703	22,278	22,939	20,616	24,398	18,840
SD	0,962625	0,836003	1,478618	0,697158	0,815829	0,898593	0,61424	0,796856
CV	3,81%	3,66%	6,24%	3,13%	3,56%	4,36%	2,52%	4,23%
Median	25,50	22,71	23,49	22,24	23,08	20,65	24,55	19,05
25% Percentile	24,63	22,14	22,64	21,75	22,21	19,89	23,99	18,27
75% Percentile	25,95	23,45	24,51	22,82	23,48	21,37	24,81	19,38
Percentile Difference	1,31	1,32	1,87	1,07	1,27	1,48	0,82	1,10

Supplementary File 4.3. **Cq details. SD = standard deviation, CV (coefficient of variation)**.

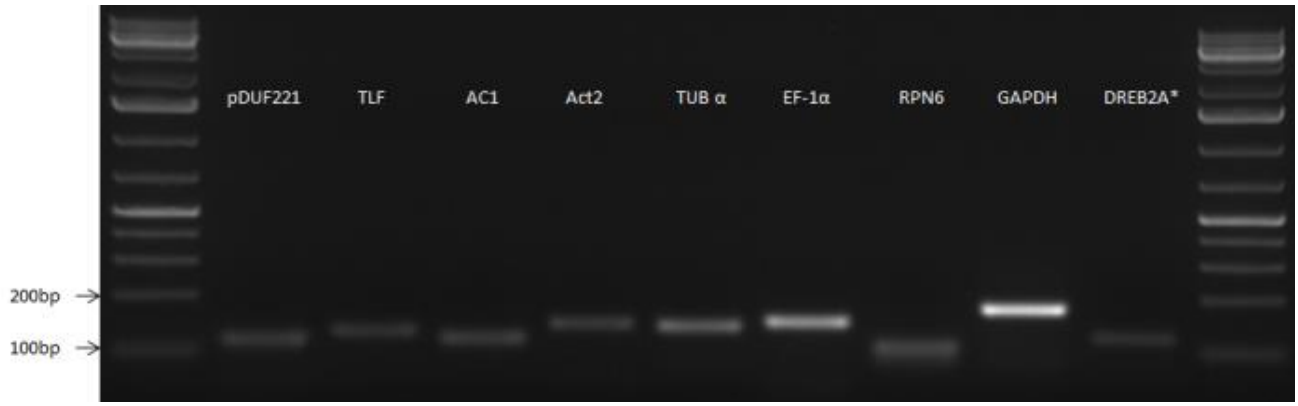
Rank	**All**						**Osmotic**					
	GeNorm		**NormFinder**		**BestKeeper**		**GeNorm**		**NormFinder**		**BestKeeper**	
	Gene	**Value**	**Gene**	**Value**	**Gene**	**Value**	**Gene**	**Value**	**Gene**	**Value**	**Gene**	**Value**
1	EF-1α RPN6	0.493	RPN6	0.134	RPN6	2,03±0,49	TLF EF-1α	0.407	RPN6	0.062	GAPDH	2,48±0,47
2	GAPDH	0.493	EF-1α	0.145	Act2	2,43±0,54	RPN6	0.474	GAPDH	0.209	RPN6	2,33±0,57
3	TLF	0.581	TLF	0.210	GAPDH	3,34±0,63	GADPH	0.513	TUB α	0.291	TUB α	2,77±0,64
4	Act2	0.646	TUB α	0.229	TUB α	2,80±0,64	Act2	0.563	Act2	0.311	pDUF221	2,54±0,65
5	TUB α	0.711	Act2	0.231	TLF	3,17±0,73	TUB α	0.593	TLF	0.345	Act2	2,94±0,66
6	pDUF221	0.752	GAPDH	0.246	EF-1α	3,65±0,75	AC1	0.666	EF-1α	0.348	AC1	3,65±0,85
7	AC1	0.885	AC1	0.595	pDUF221	3,08±0,78	pDUF221	0.758	AC1	0.579	TLF	3,83±0,89
8			pDUF221	0.646	AC1	4,74±1,13			pDUF221	0.648	EF-1α	4,61±0,95

Rank	**Heavy metal**						**Heat shock**					
	GeNorm		**NormFinder**		**BestKeeper**		**GeNorm**		**NormFinder**		**BestKeeper**	
	Gene	**Value**	**Gene**	**Value**	**Gene**	**Value**	**Gene**	**Value**	**Gene**	**Value**	**Gene**	**Value**
1	EF-1α RPN6	0.404	TLF	0.234	Act2	2,19±0,49	TLF RPN6	0.488	RPN6	0.167	RPN6	1,39±0,34
2	pDUF221	0.586	Act2	0.275	pDUF221	1,94±0,50	EF-1α	0.527	TLF	0.196	GAPDH	2,44±0,46
3	TLF	0.615	pDUF221	0.306	TLF	2,30±0,52	GADPH	0.599	Act2	0.332	Act2	2,36±0,52
4	Act2	0.657	RPN6	0.307	TUB α	2,67±0,60	Act2	0.624	EF-1α	0.335	TLF	2,46±0,57
5	TUB α	0.694	TUB α	0.398	RPN6	2,64±0,64	TUB α	0.686	TUB α	0.340	EF-1α	2,89±0,60
6	GADPH	0.755	EF-1α	0.417	AC1	3,29±0,75	pDUF221	0.850	GAPDH	0.385	TUB α	2,83±0,66

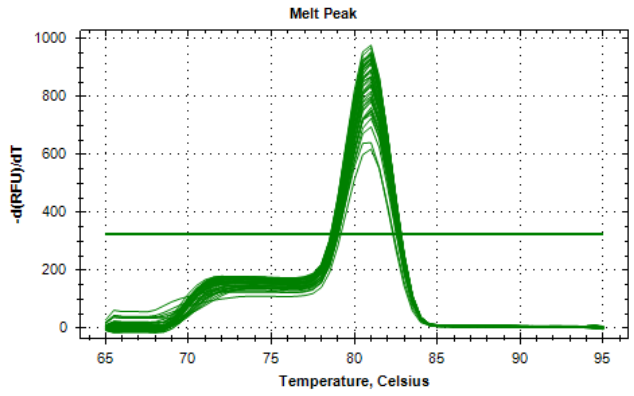
7	AC1	0.832	GAPDH	0.624	EF-1α	4,14±0,85	AC1	1.081	pDUF221	0.936	pDUF221	3,64±0,90
8			AC1	0.677	GAPDH	4,83±0,90			AC1	1.186	AC1	5,84±1,44

Supplementary File 4.4. **GeNorm, NormFinder and BestKeeper ranks** computed with shoot and root together for all the single stress and their combination (referred as all).

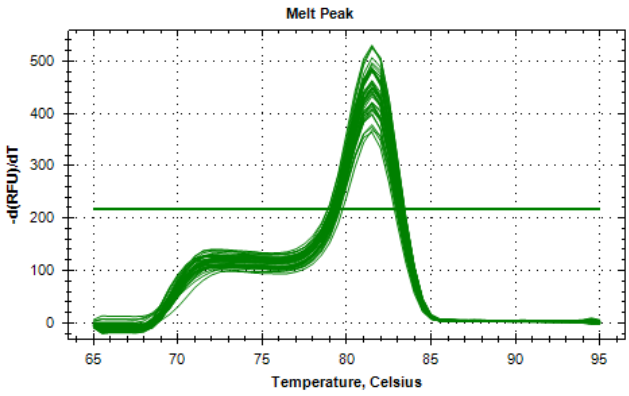
A



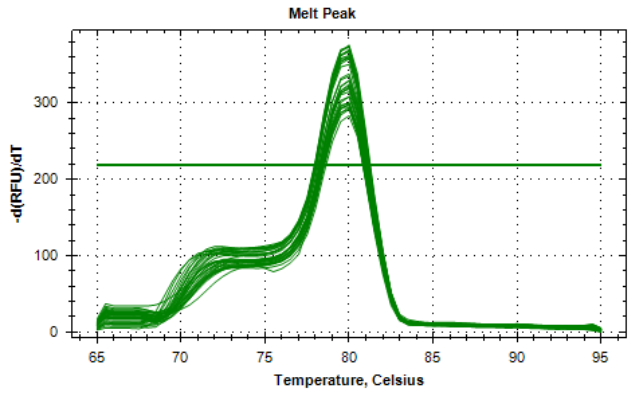
B



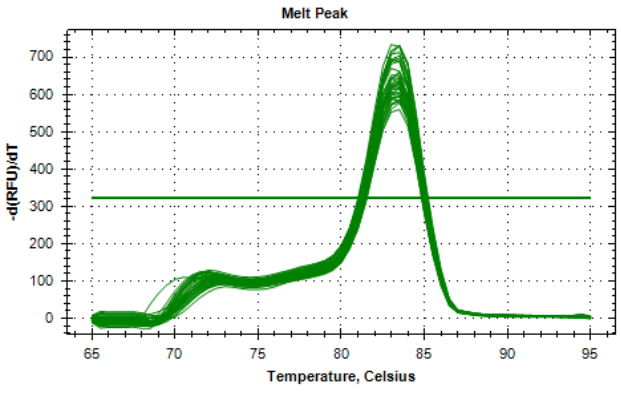
pDUF221



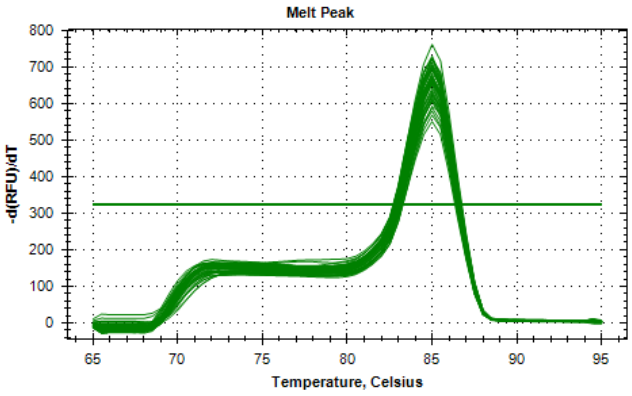
AC1



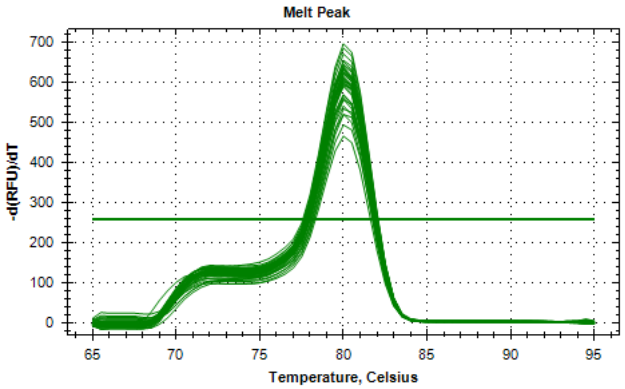
TLF



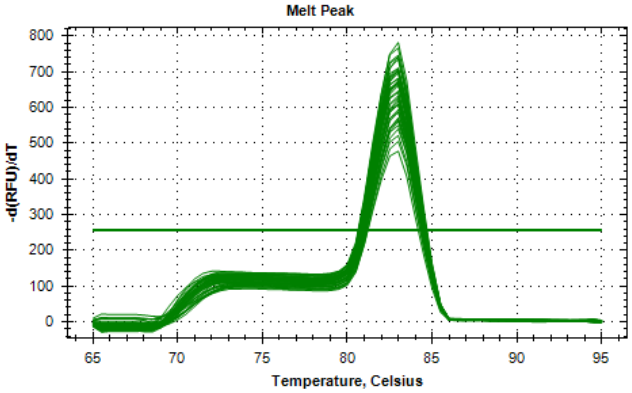
Act2



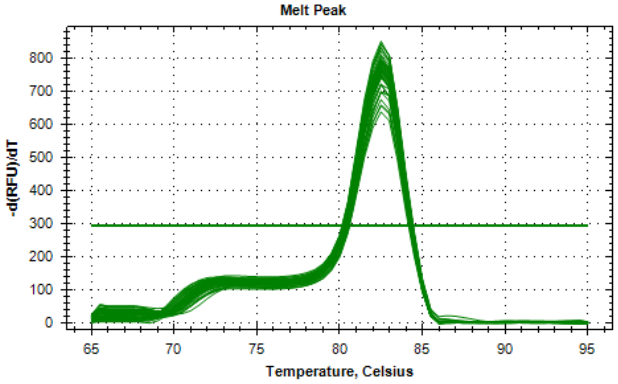
TUB α



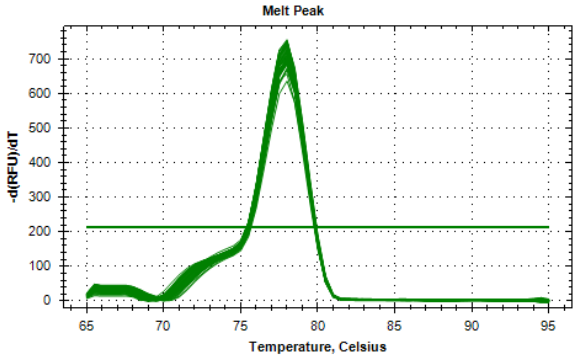
RPN6



EF-1α

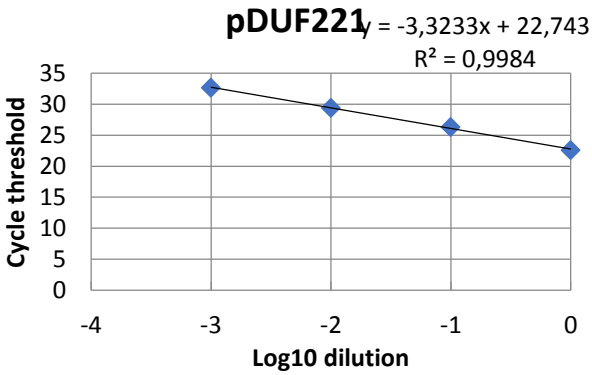


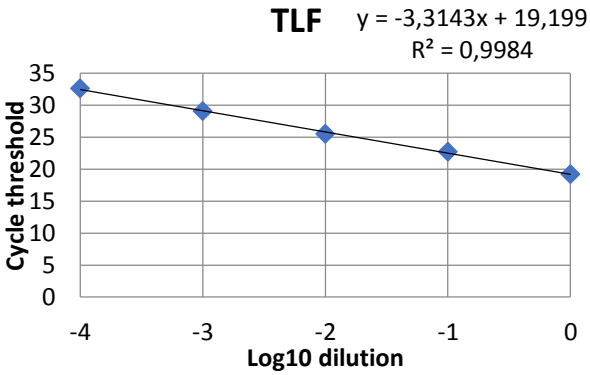
GAPDH

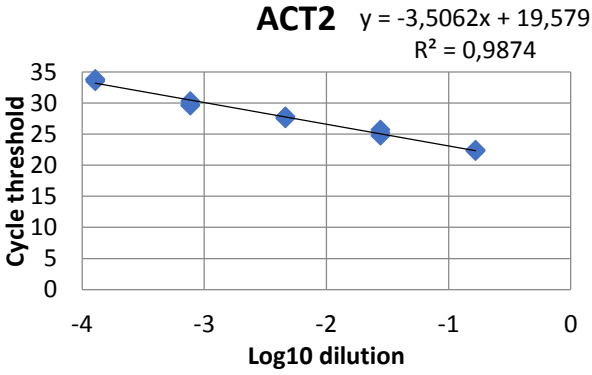


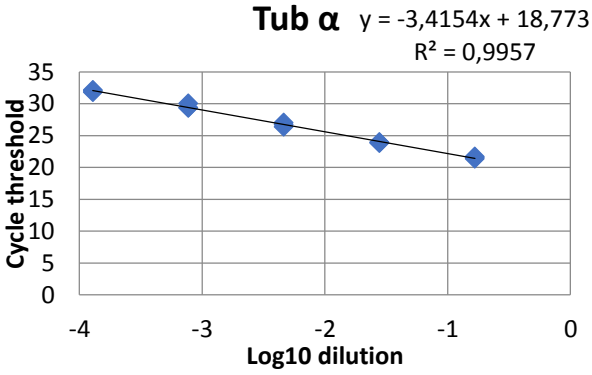
DREB2A*

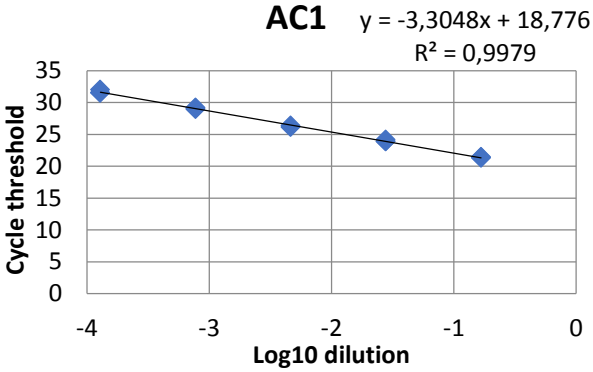
Supplementary File 4.5. **Size of the gene visualized on 2% agarose gel (A) and melting curves of 8 candidate genes (B)**. Validation gene indicated by "*".

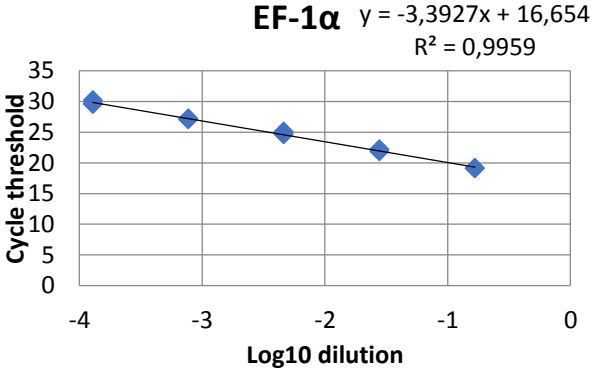

pDUF221
y = -3,3233x + 22,743
R² = 0,9984
Cycle threshold
Log10 dilution

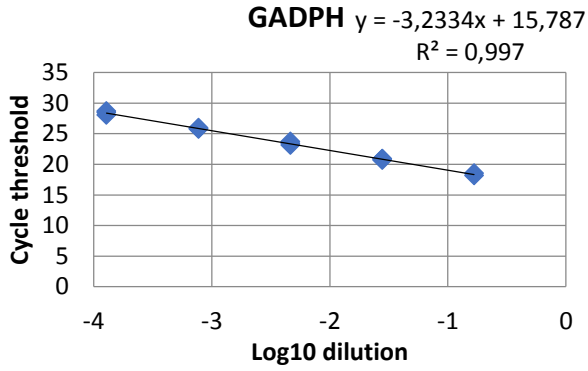

TLF
y = -3,3143x + 19,199
R² = 0,9984
Cycle threshold
Log10 dilution

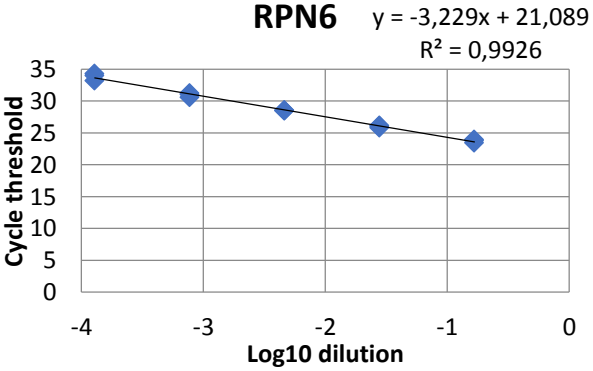

ACT2
y = -3,5062x + 19,579
R² = 0,9874
Cycle threshold
Log10 dilution

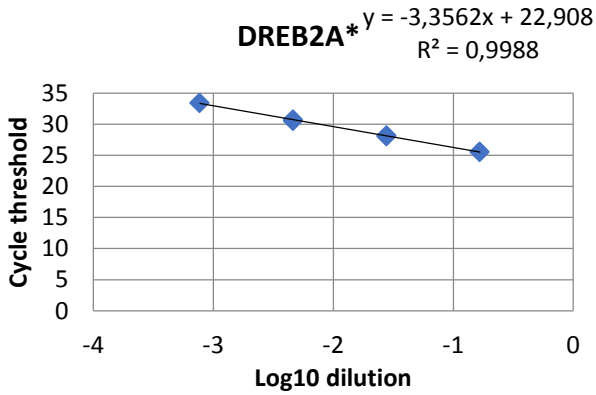

Tub α
y = -3,4154x + 18,773
R² = 0,9957
Cycle threshold
Log10 dilution


AC1
y = -3,3048x + 18,776
R² = 0,9979
Cycle threshold
Log10 dilution

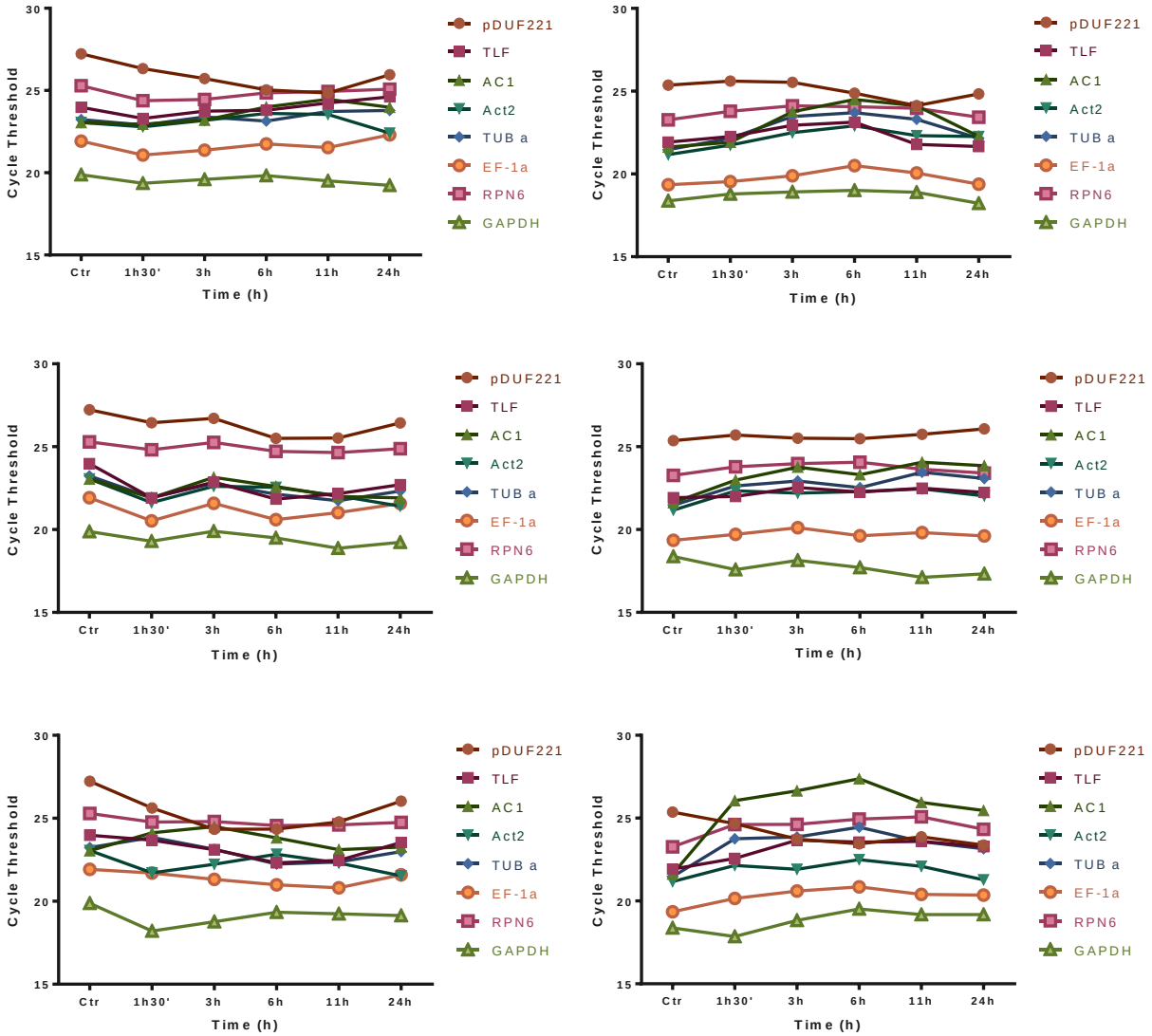

EF-1α
y = -3,3927x + 16,654
R² = 0,9959
Cycle threshold
Log10 dilution


GADPH
y = -3,2334x + 15,787
R² = 0,997
Cycle threshold
Log10 dilution

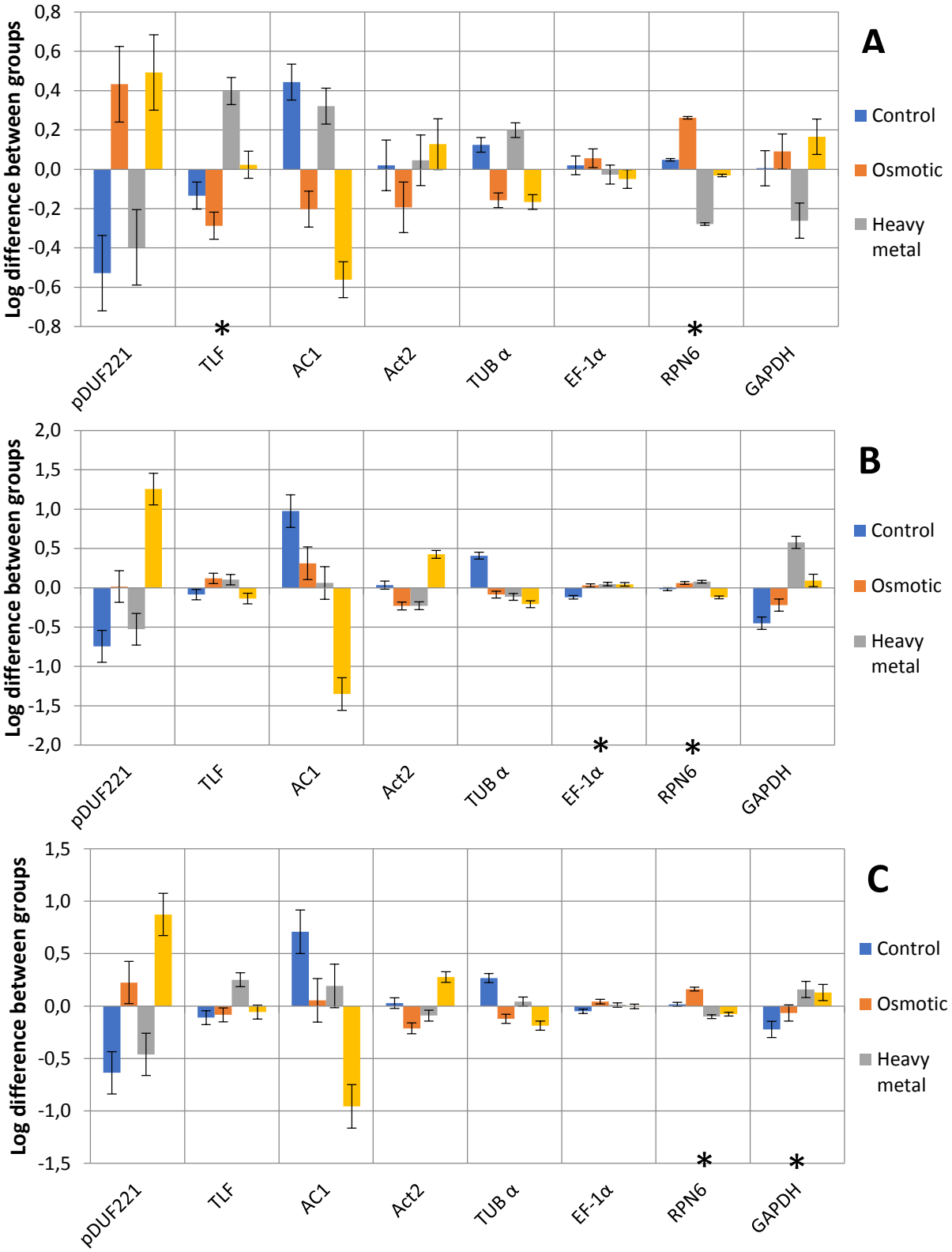

RPN6
y = -3,229x + 21,089
R² = 0,9926
Cycle threshold
Log10 dilution



Supplementary File 4.6. **Standard curves for 8 candidate genes**. Curves are made with a serial 6-fold dilution. Blue squares indicate the three technical replicas for each dilution step. Validation gene indicated by "*".



Supplementary File 4.7. **Expression profile of 8 candidate genes divided by stress and tissues**: Osmotic stress in shoot (A) and root (B); Heavy metal stress in shoot (C) and root (D); Heat shock stress in shoot (E) and root (F).



Supplementary File 4.8. **NormFinder intergroup variation.** Dataset was divided in shoot (A), root (B), shoot + root (C) and input in Normfinder algorithm with four subgroups: control plus three stresses (osmotic, heavy metal and heat shock). Bars rapresent the intergroup variation respect the average and vertical lines the mean of intragroup variation, asterisc (*) indicates the best combination of two genes.