**University of Naples Federico II**
**Department of Agriculture**

**PhD course in Agricultural and Food Sciences**
**XXX Cycle (2014-2017)**

# TOWARDS UNDERSTANDING THE MOLECULAR DYNAMICS OF COLD STRESS RESPONSE IN A TOLERANT POTATO SPECIES, *SOLANUM COMMERSONII* DUN.

**Ph.D. Dissertation**
**Dr. Salvatore Esposito**

**TUTOR:**
Prof. Domenico Carputo

***INDEX***

## CHAPTER I. GENERAL INTRODUCTION

**1.1.1 LOW TEMPERATURE STRESS IN PLANTS: A BRIEF OVERVIEW…….…….....4**
**1.1.2 BREEDING FOR COLD RESISTANCE (OR TOLERANCE) IN THE GENOMIC ERA…….……9**
**1.2 THE CULTIVATED POTATO AND ITS WILD GERMPLASM……..……10**
**1.2.1 *SOLANUM COMMERSONII*: a useful source of resistance to cold……..……14**
**1.3 OBJECTIVES OF THE THESIS……20**
**1.4 REFERENCES……..22**

## CHAPTER II. TRANSCRIPTOMIC PROFILING OF COLD TOLERANT *SOLANUM COMMERSONII* REVEALS NEW INSIGHTS IN THE RESPONSE TO LOW TEMPERATURES

**2.1 INTRODUCTION……..28**

**2.2 MATERIALS AND METHODS……...……30**

**2.3 RESULTS……....33**

**2.4 DISCUSSION……..43**

**2.5 CONCLUSIONS……46**

**2.6 REFERENCES……...……46**

## CHAPTER III. DEEP-SEQUENCING OF *SOLANUM COMMERSONII* SmRNA LIBRARIES REVEALS RIBOREGULATORS INVOLVED IN COLD STRESS RESPONSE

**3.1 INTRODUCTION……..71**

**3.2 MATERIALS AND METHODS……...73**

**3.3 RESULTS……75**

**3.4 DISCUSSION……..83**

**3.5 CONCLUSIONS……..…88**

**3.6 REFERENCES……...89**

**CHAPTER IV. DICER-LIKE AND RNA-DEPENDENT RNA POLYMERASE GENE FAMILIES IDENTIFICATION AND ANNOTATION IN THE CULTIVATED *SOLANUM TUBEROSUM* AND ITS WILD RELATIVE *S. COMMERSONII***

**4.1 INTRODUCTION……113**

**4.2 MATERIALS AND METHODS……..…115**

**4.3 RESULTS……..117**

**4.4 DISCUSSION……133**

**4.5 CONCLUSIONS……...…135**

**4.6 REFERENCES……136**

**CHAPTER V. GENERAL CONCLUSIONS……. 147**

# Chapter I. General introduction

Partially published in: Esposito S., D'Amelia V., Villano C., Contaldi F., Carputo D., Aversano R. (2017) The Wild Side of Potato: Insights into the Genome Sequence of the Stress-Tolerant *S. commersonii* Dun.. In The Potato Genome. Eds. Chakrabarti SK, Conghua X and Tiwari JK. Springer. In press. ISBN 978-3-319-66133-9

### 1.1.1 LOW TEMPERATURE STRESS IN PLANTS: A BRIEF OVERVIEW

Abiotic factors such as heat, cold, drought, salinity and nutrient deficiency have a huge impact on world agriculture. It has been suggested that they reduce average yields by >50% for most major crop plants (Viswanathan and Zhu, 2002; Puyaubert and Baudouin, 2014) (Tab. 1). Current climate prediction models indicate that surface temperatures will rise by 3–5 °C in the next 50–100 years, drastically affecting global agricultural systems (IPCC, 2007). This will be concurrent with an increased frequency of drought, flood and heat waves (Mittler and Blumwald, 2010). In particular, warmer and drier summers in mid-continental regions such as central Europe and central Africa are predicted, along with a reduction in the growing season of many regions, extensive salinization as sea levels rise and a decrease in land suitable for agriculture (Morison et al. 2008). A change in the variability of rainfall and temperature may itself affect yields as well as adversely impact the nutritional quality of products (Porter and Semenov, 2005). Crop plants are therefore likely to encounter a greater range and number of environmental stresses which, when occurring simultaneously, can have severe consequences.

**Tab 1.** Estimated losses in percentage of six major crops (wheat, barley, soybean, maize, potato, sugar beet) due to abiotic and biotic stresses. Record and average yield are also reported.

| Type of crop | Record (max) yield ($kg/ha^{-1}$) | Average yield (kg/ha-1) | Average loss due to abiotic stress (%) | Average loss due to biotic stress (%) |
|---|---|---|---|---|
| Wheat | 14.500 | 1.880 | 82% | 5% |
| Barley | 11.400 | 2.050 | 75% | 7% |
| Soybean | 7.390 | 1.610 | 69% | 9% |
| Maize | 19.300 | 4.600 | 66% | 10% |
| Potato | 91.400 | 28.300 | 54% | 20% |
| Sugar beet | 121.000 | 43.600 | 50% | 14% |

Freezing or extremely low temperature constitutes a key factor influencing plant growth, development and crop productivity. Generally, plants greatly differ in their abilities to cope with this stress. When they are exposed to gradually decreasing temperatures below a certain threshold, they acclimatize (low-temperature acclimation) to the stress, a process called *cold hardening*. In spite of various adaptations to cold, plants may be injured through exposure to cold temperatures in a variety of ways, depending on the temperature range. One type of injury, called chilling injury, occurs below the freezing point of water. Sometimes ice crystals form in the protoplasm of cells, resulting in cell and, possibly, plant death. Crops may be classified into three groups according to

their tolerance to low temperatures. Frost-sensitive crops are intolerant of ice in their tissues and are hence sensitive to chilling injury. Among them, beans, corn, tomato plants can be killed when temperatures fall just below 0°C. Frost-resistant plants can tolerate some ice in their cells and can survive cold temperatures of up to − 40°C. Cold-hardy plants are predominantly temperate woody species. They can survive temperatures of up to − 196°C. Thus, some species have evolved extraordinary mechanisms to resist and enhance their tolerance during exposure to periods of low temperatures. Understanding the molecular basis of the response to freezing temperatures is therefore a fundamental goal of plant researchers. It has been established that some cold-regulated genes could contribute to freezing tolerance. Whereas it is agreed that low-temperature tolerance is a complex trait, neither the mode of gene action governing the expression of the trait nor its pathway have been established yet. It is known that plants may sense low temperatures through changes in the physical properties of membranes, because their fluidity is reduced during cold stress (Medina et al. 2011). This induces the expression of many genes (*COR,* cold-responsive genes), which led to the accumulation of compatible osmolytes such as proline, betaine, polyols and soluble sugars. Thanks to the advent of molecular biology, today some of the events comprised between membrane rigidification and the expression of *CORs* are defined (for reviews, see Sakuma et al. 2002; Thomashow 2010; Park et al. 2015). For example, it is known that membrane rigidification induces an increase of cytosolic $Ca^{2+}$ that implies the activation of $Ca^{2+}$ sensors. Plants possess many groups of $Ca^{2+}$ sensors, including CaM (calmodulin) and CMLs (CaM-like), CDPKs ($Ca^{2+}$-dependent protein kinases), CCaMK ($Ca^{2+}$-and $Ca^{2+}$/CaM-dependent protein kinase), CAMTA (CaM-binding transcription activator), CBLs (calcineurin B-like proteins) and CIPKs (CBL-interacting protein kinases) (Fig. 1).

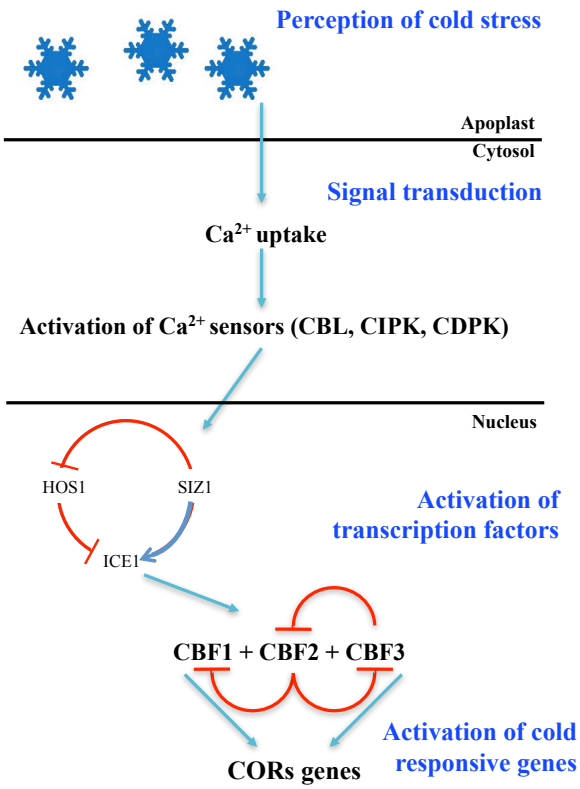


**Fig 1.** Cold-sensing and signalling pathway. Interactions between different transcription factorsare also shown by red and blue arrows.

Genetic analysis demonstrated that CDPKs work as positive regulators (Luan et al. 2002), while calmodulin3 is a negative regulator of gene expression and cold tolerance (Townley et al. 2002). CAMTA3 also has been identified as a positive regulator of *CBF* expression through binding to a regulatory element (CG-1 element, vCGCGb) in their promoter (Doherty et al. 2009). The expression of these proteins led the activation of transcription factors (TFs), especially those belonging the AP2/ERF family, called *CBFs*. Indeed, the *CBF*-dependent cold signaling pathway is the best characterized and the key regulatory pathway involved in cold stress (Thomashow, 2010; Knight and Knight, 2012). In *Arabidopsis*, three duplicated *CBFs* are involved in the regulation of *COR* gene expression (Maruyama et al. 2004; Vogel et al. 2005). These genes are also duplicated in other species such as wheat and barley, and they are located in tandem with highly conserved coding sequence. (Stockinger et al. 2007; Knox et al. 2010). Various studies demonstrated that winter varieties of barley and wheat underwent to an expansion of *CBF* genes compared with their spring counterpart, leading to the idea that freezing tolerance in these species is associated with

higher *CBF* gene copy number (Skinner et al. 2005; Badawi et al. 2007; Zhu et al. 2014). *CBF* gene deletion and mutation are characteristic only of spring varieties (Fig. 2).

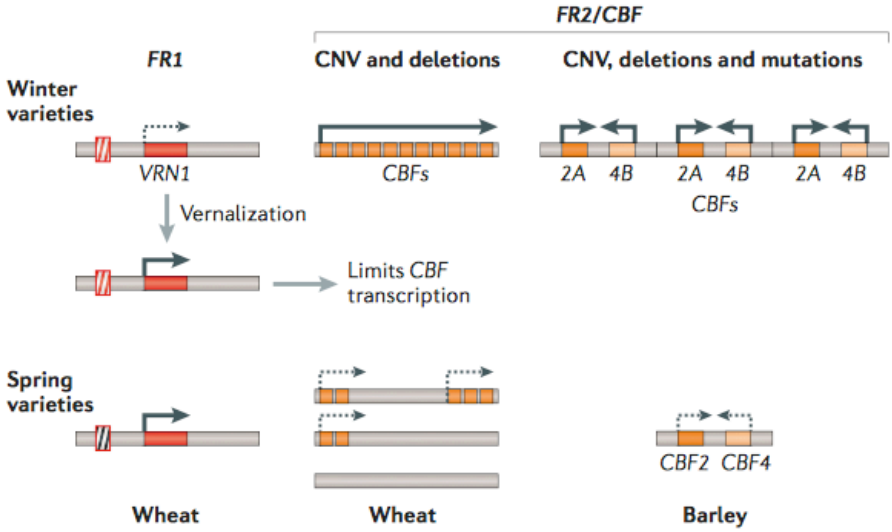


**Fig 2.** Cold tolerance in wheat and barley at *FR2*/*CBF* locus. Freezing tolerance of winter wheat and barley is associated with high *CBF* copy number. *CBF* gene duplications are peculiar of winter tolerant varieties, whereas deletion and mutation are characteristic of spring varieties. The thickness of the right-angled arrows represents transcriptional activity, and dashed arrows indicate low levels of inducible expression. The striped rectangles represent different regulatory sequences (Michael et al. 2015).

The *CBF* (mainly *CBF3*) pathway is controlled by a MYC-type transcription factor *ICE1* (inducer of *CBF* expression1) (Chinnusamy et al. 2007; Zhu et al. 2007; Liu et al. 2010; Lissarre et al. 2010). *ICE1* can bind the MYC recognition *cis*-elements (CANNTG) in the promoter of *CBF3* and induce its regulon during cold acclimation. Approximately 40% of *COR* genes are regulated by *ICE1*, suggesting that it is the master regulator that controls the expression of these genes (Lee et al. 2005). In turn, *ICE1* is regulated by ubiquitylation (Dong et al. 2006) and sumoylation (Miura et al. 2007). However, the precise mechanisms for its activation and inactivation remain still unknown. The final expression of *COR* genes regulated by *ICE1* has been shown to be critical in plants for both chilling tolerance and cold acclimation. Among the *COR* genes, *COR78/RD29A, COR47, COR15A* and *COR6.6* encode dehydrins, which are known as group 2 LEA (LEA II) proteins (Ingram et al. 1996). They are important for membrane stabilization and prevent protein aggregation (Hundertmark et al. 2008). Other dehydrins such as ERD10 (early response to dehydration10) and ERD14 function as chaperones and interact with phospholipid vesicles through

electrostatic forces (Kiyosue et al. 1994; Kovacs et al. 2008). In addition to LEA proteins, HSPs (heat shock proteins) expression is also induced by cold (Kaur et al. 2011). They function in membrane protection, in the refolding of denatured proteins and in preventing protein aggregation. In *Arabidopsis* it has been shown that also pathogen-related proteins (PR) such as PR1, PR2 (*β*-1,3-glucanase) and PR5 (thaumatin-like proteins) are induced by cold treatment, acting as *COR* genes (Seo et al. 2008). The antifreeze activity of *β*-1,3-glucanase, chitinases and thaumatin-like proteins inhibits the recrystallization of intercellular ice in the apoplastic space and prevents intracellular ice formation (Janska et al. 2010). In addition to all these proteins, many other enzymes are involved in the cold response machinery, such as those involved in detoxification and antioxidation, photosynthesis, lignin metabolism, secondary metabolism, cell wall polysaccharide remodeling, starch metabolism, sterol biosynthesis and oligosaccharide synthesis (reviewed in (Janska et al. 2010).

The mechanisms described above are not the unique determinant involved in cold tolerance. Indeed, although some molecular mechanisms on how different species respond to cold stress remain to be elucidated, recent studies have shown that cold also induces aberrant expression of many small non-coding RNA (sncRNA) in several plant species (Sunkar et al. 2007; Khraiwesh et al. 2012). Two main classes of small regulatory RNAs are distinguished: microRNAs (miRNAs) and small interfering RNAs (siRNAs). Zhou et al. (2008) found that 19 miRNA genes of 11 miRNA families in *A. thaliana* are up-regulated by cold stress. Six of them were induced, while the remaining 5 showed either transient or mild regulation. Cold stress also changed the expression of siRNAs in wheat (Yao et al. 2010) and *Populus trichocarpa* (Lu et al. 2008). Both miRNAs and siRNAs are synthetized and processed by several proteins including DCL, RDR and AGO (Fig. 3).

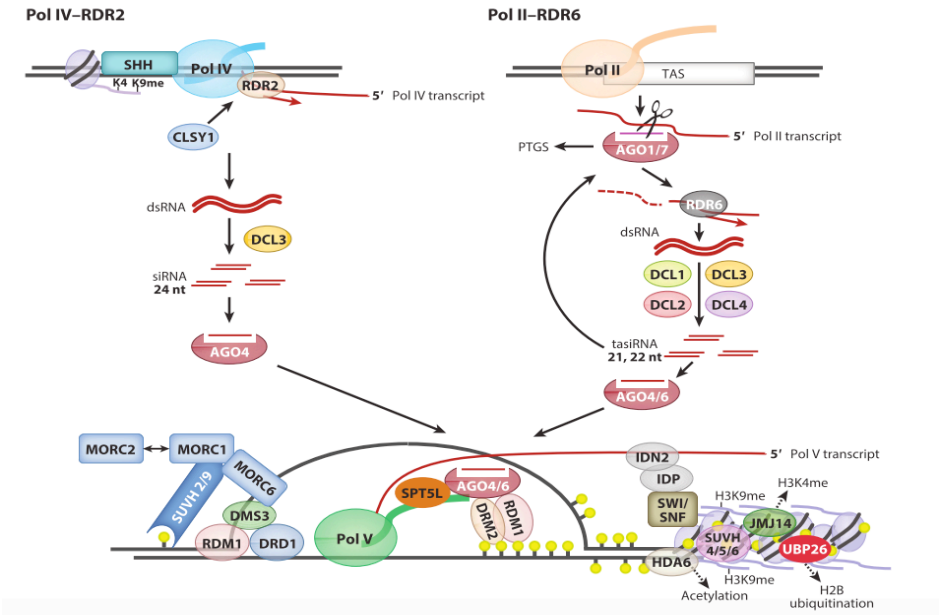

**Fig. 3** Biogenesis of miRNAs and siRNAs. (Upper left) Pol IV–RDR2–dependent small interfering RNA (siRNA) biogenesis. The dsRNAs are processed into 24-nucleotide (nt) siRNAs by DCL3, and the guide strand is incorporated into AGO4 or AGO6, which then enters the Pol V–mediated pathway of *de novo* DNA methylation. (Upper right) Pol II–RDR6–dependent siRNA biogenesis. Pol II transcribes TAS noncoding RNAs, which undergo microRNA-guided slicing by either AGO1 or AGO7. An RNA cleavage product is copied by RDR6 into dsRNA, which is processed into 21–24-nt siRNAs by various DCL activities (Matzke et al. 2015).

All of them guide target degradation at the post-transcriptional level or at the transcriptional level through a pathway termed RNA-directed DNA methylation (RdDM). This latter mechanism seems to be particularly significant, as trans-generational effects in plants are associated with alterations in methylation of genomic DNA. Epigenetic mechanisms may also be involved in the response to stress. Boyko et al. (2010) found that stress-induced trans-generational responses in *Arabidopsis* depended on altered DNA methylation and small RNA-silencing pathways. One of the factors contributing to these changes involves mobile genetic elements. Different stress factors may decrease the methylation level of these sequences (Kalinka et al. 2009), leading to their activation and transposition. Cold was found to down-regulate MET1, resulting in demethylation of mobile genetic elements in *Zea mays* (Steward et al. 2002) and *Antirrhinum majus* (Hashida et al. 2006). Ito et al. (2011) showed that some retrotransposons become active in *Arabidopsis* seedlings subjected to cold stress. The siRNA pathway plays a crucial role in restricting retrotransposition triggered by environmental stress. As changes in methylation at mobile genetic element insertions affect nearby genes, mobility bursts may generate novel, stress-responsive regulatory gene networks. However, it is the miRNA pathway that seems more involved in stress adaptation responses. Although a large number of siRNAs and miRNAs have been identified, only a few dozen small RNAs have been annotated with specific functions.

### 1.1.2 BREEDING FOR COLD TOLERANCE IN THE GENOMIC ERA

Plant breeding is based on the production of genetic variability followed by efficient selection of newly produced genotypes suited to the needs of farmers and consumers and industries. Plant breeding purposes have been enormously successful on a global scale, with examples in the development of $F_1$ hybrids in maize, the introduction of wheat and rice varieties that spawned the

Green Revolution, and the commercialization of transgenic crops. Genetic variability can be produced in several ways, e.g. through crossing parental cultivated genotypes or exploring wild germplasm. Similarly, the recent integration of advances in biotechnology, genomic research, and molecular marker applications with conventional plant breeding practices has created the foundation for molecular plant breeding, an interdisciplinary science that is revolutionizing crop improvement.

To cope with abiotic stressors such as low temperatures, classical plant breeding alone can hardly achieve the expected results. It is now widely accepted that genes, transcripts, proteins, and metabolites that control the architecture and/or the stress tolerance of a crop need to be identified to facilitate breeding efforts. Furthermore, several questions must be elucidated regarding, e.g.: genes and proteins that are up- or down-regulated, the functions of responsive genes, proteins, and metabolites, the characteristics of stress perception, signal transduction, gene activation, protein expression, metabolite production and whole plant response. Many studies demonstrated the success of transgenic approaches in increasing tolerance to low temperatures. For example, the chloroplast *GPAT* (glycerol-3-phosphate acyltransferase) of squash, *Cucurbita maxima*, and *Arabidopsis* is involved in phosphatidyl glycerol fatty acid desaturation and increases the ratio of unsaturated fatty acids in plant cell membranes, leading to enhancement of cold tolerance (Murata et al. 2002). In addition, the citrus *LEA* gene, *CuCOR19* and wheat dehydrin *WCO410* enhanced cold tolerance of transgenic tobacco and strawberry (Hara et al. 2003; Houde et al. 2004). Nowadays, the identification of genes, transcripts, proteins and metabolites useful for breeding involves the use of new advanced tools. Among them, high-throughput genome sequencing efforts have dramatically increased knowledge of and ability to characterize genetic diversity in the germplasm pool. Next Generation Sequencing (NGS) technologies are allowing the mass sequencing of genomes and transcriptomes, which is producing a vast array of information. The analysis of NGS data by bioinformatics allows discovering new genes and regulatory sequences and their positions, and makes large collections of molecular markers available. These advances, coupled with detailed characterization of genes involved in the tolerance and studies on the molecular, cellular and physiological processes underlying plant environmental adaptation, will accelerate the recognition of key loci and will facilitate future breeding efforts for cold tolerance.

## 1.2 THE CULTIVATED POTATO AND ITS WILD GERMPLASM

The cultivated potato, *Solanum tuberosum* L., originated in the Andean region of South America. Since its introduction in Europe, it has been spread globally and nowadays it is the third most important crop in the world, following maize and wheat (Liu et al. 2016). Potato is a rich source of

energy, with a starch content that accounts for 80% of the tuber dry weight and with a high content of quality protein and vitamin C (Scott et al. 2000). In addition, it yields on average more food, energy and protein per unit of land than cereals (Horton, 1988). Potato is a crop of temperate climate and it is sensitive to frost. The young plants grow well at 24 °C, and late growth is favoured at 18 °C. Tuber production is maximum at 20 °C and decreases with rise in temperature. At about 30 °C tuber production is heavily compromised. After tomato, in Italy potato is the most widespread crop, with a production of about 1.3 million tons and an area down by about 72,000 hectares (data FAOSTA 2013). The average yield of potatoes around the world is far below its physiological potential of 120 tons/ha (Papademetriou et al. 2008). Advances in potato breeding have been constrained by its complex biological system including vegetative propagation, autotetraploidy, and high levels of heterozygosity (Mendiburu et al. 1997). However, the potato genome (The Potato Genome Sequencing Consortium et al. 2011) and accompanying gene complement are powerful resources for understanding this complex system and advancing molecular breeding. The potato genome is roughly 844 Mb with 39,031 protein-coding genes and present evidence for at least two genome duplication events indicative of a palaeopolyploid origin. As the first genome sequence of an asterid, the potato genome reveals 2,642 genes specific to this large angiosperm clade. Gene family expansion, tissue-specific expression and recruitment of genes to new pathways contributed to the evolution of tuber development.

*S. tuberosum* has an extremely large secondary and tertiary gene pools consisting of related tuber-bearing species. They have useful genes lacking in cultivated varieties and represent a reservoir of allelic variability to increase heterosis for polygenic traits. Among traits possessed are the high dry matter content, the low content of reducing sugars, resistance to biotic (fungi, insects, nematodes, bacteria and viruses) and abiotic (e.g. cold and drought) stressors, and many others related to quality and productivity. Sources of resistance to the *Leptinotarsa decemlineata* were found in *S. pinnatisectum* and *S. tarijense*. Resistance to *Verticillium spp*. and *Clavibacter spp*. has been reported in *S. lesteri, S. polyadenium* and *S. jemesii.* Particularly interesting are the wild species *S. bulbocastanum*, *S. polyadenium, S. pinnatisectum*, resistant to *Phytophthora infestans*, the main adversity of the potato in terms of economic losses. About 200 species and many intraspecific taxa have been described. These taxa have been classified in series, with different authors recognizing different numbers of series, often with different circumscriptions. Hawkes (1989) suggested a division of the series into two super series, *Stellata* and *Rotata*, emphasizing the outline of the corolla as a major distinctive character. Some of the series contain only one or just a few species, indicating that their relationship to the other species is not clear. Series such as *Piurana* and, especially, *Tuberosa*, are large groups of species that may not be closely related to

each other. Hijmans and Spooner (2001) documented the geographic distribution of wild potato species, with the majority occurring in Argentina, Bolivia, Mexico and Peru, many with only restricted distribution areas (Fig. 4).


**Fig 4** Distribution of wild potatoes (grey shade) (Spooner and Hijmans 2001).

Potato species largely differ in ploidy level, with 12 being the basic (x) chromosome number. Most wild potatoes are diploids (2n=2x=24) and six of these diploid have additional triploid populations with 36 chromosomes (3x). Seven species are exclusively triploid, 22 exclusively tetraploid (48 chromosomes, 4x), one exclusively pentaploid (60 chromosomes, 5x), and 12 exclusively hexaploid (72 chromosomes, 6x). Three species have populations with more than one even ploidy level (*S. acaule* 4x, 6x; *S. leptophyes* 2x, 4x; *S. oplocense* 2x, 4x, 6x). The triploid and pentaploid populations are generally highly sterile. They are less likely to be discovered as most germplasm collecting expeditions collected seeds rather than tubers. It is likely, therefore, that the number of species with additional triploid or pentaploid populations is greater than currently known.

Among wild potato species, *S. commersonii* is particularly attractive for its freezing tolerance and capacity to cold acclimate (i.e., ability to increase cold tolerance after exposure to low, non-freezing temperatures) (Fig. 5).

**Fig 5.** Phenotype of cold treated *S. tuberosum* and *S. commersonii* plants following a cold stress treatment (-2°C for one hour). The picture has been taken after recovery (24 hours at 24°C).

It also possesses several other noteworthy traits, mainly related to resistance to biotic stresses, such as PVX and common scab. Since this species is the object of our research, the next paragraph is aimed at giving an exhaustive overview of its biological and genetics characteristics.

### 1.2.1 *SOLANUM COMMERSONII:* A USEFUL SOURCE OF TOLERANCE TO COLD

*S. commersonii* is a tuber-bearing wild diploid species native to Central and South America (Fig. 6). The French taxonomist Michel-Felix Dunal named this species in honor of Philibert Commerson (1727-73), who collected the type specimen (No. 47) in 1767 at Montevideo, Uruguay. This was probably the first wild potato to be collected on a scientific expedition (Hawkes, 1889).

**Fig 6.** Plant, flower and tuber phenotype of wild *Solanum commersonii* (Dun.).

Analyses of chloroplast genome restriction sites confirmed that *S. commersonii* is phylogenetically distinct from the cultivated potato (Rodríguez and Spooner, 2009; Aversano et al. 2015) and phylogenetic analyses indicated that they diverged ~2.3 million years ago. Consistently, *S. commersonii* and *S. tuberosum* are sexually incompatible (Jackson and Hanneman, 1999). Despite being genetically isolated from the cultivated potato, *S. commersonii* has garnered significant research interest. It possesses several resistance traits not found in the cultivated potato. Among them resistance to root knot nematodes, soft rot and blackleg, Verticillium wilt, Potato Virus X (PVX), Tobacco Etch Virus (TEV), common scab, late blight and *Ralstonia solanacearum* ( Hawkes, 1889; Micheletto et al. 2000; González et al. 2013). Importantly, the genome sequence of this species has been deciphered (Aversano et al. 2015). Its genome is roughly 830 Mb with an N50 of 44,303 bp. Compared with potato, *S. commersonii* shows a striking reduction in heterozygosity

and differences in genome sizes were mainly due to variations in intergenic sequence length. Genome annotation revealed a catalog of 39,290 protein-coding genes, 126 cold-related genes that are lacking in *S. tuberosum,* and 1.703 predicted microRNAs.

As already mentioned above, particularly interesting are its cold tolerance and ability to cold acclimate. Equally important for research purposes is the presence of intra-species variability in terms of cold tolerance. Indeed, available at the Department of Agricultural Sciences are two *S. commersonii* clones with contrasting phenotype: one is tolerant and able to acclimate, the other is susceptible but able to acclimate (Carputo et al. 2013; Fig. 7).

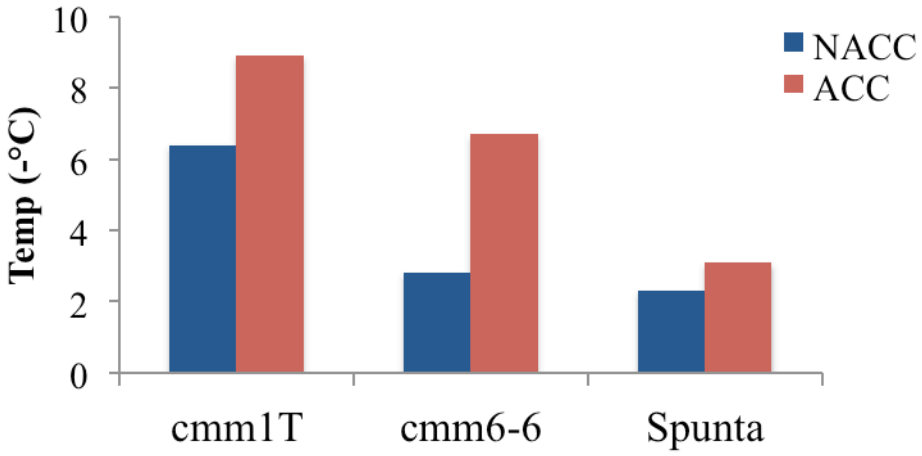


**Fig.7** Killing temperatures (-°C) in non-acclimated (NACC) and acclimated (ACC) conditions of clones cmm1T and cmm6-6 of *S. commersonii* and cv. Spunta of *S. tuberosum* (data from Carputo et al. 2013)

One of the first study regarding its cold tolerance was reported by Chen and Li (1980), who compared the biochemical changes involved in the cold acclimation process in *S. commersonii* vs. *S. tuberosum*. Their research showed that there were similar increases in sugar and starch content during cold acclimation. However, the net synthesis of soluble proteins and the level of total lipids and phospholipids were higher in our wild species. Chen et al. (1992) have also observed changes in endogenous ABA levels to increase transiently in *S. commersonii* after four days of cold acclimation and also detected two separate peaks of free ABA on the second and sixth days of cold acclimation. Evidence for the role of ABA as a signaling molecule was revealed when exogenously applied ABA was observed to improve its freezing tolerance. In *S. commersonii,* freezing tolerance is defined in genetic terms as a complex quantitative trait controlled by several, as yet unknown, combinations of genes and gene families (Stone et al. 1993). In spite of the lack of information on target traits and its sexual isolation, breeders have successfully introgressed genes from *S.*

*commersonii* into the cultivated potato either sexually or through somatic fusion (Cardi et al. 1993; Carputo et al. 1997; González et al. 2013; Zuluaga Cruz et al. 2014). Despite such efforts, very little progress has been made in the release of new varieties. This is at least partially due to the lack of genomics resources available for *S. commersonii*. With the help of its genome sequence, advances in the identification of genes involved in cold stress and the presence of interesting paralogs has been elucidated. For example, Aversano et al. (2015) identified genes involved in freezing and cold acclimation responses comparing the transcript expression profiles of frost stress non-acclimated (NACC) and acclimated (ACC) plants. Among the transcriptomic changes, proteins involved in the cold response machinery, such as antioxidant, secondary metabolism, cell wall remodeling, starch metabolism, and heat shock protein were found differentially expressed. As regard as the cold response machinery, in light of their prominent role in plant cold acclimation, the authors revealed the structural organization of *CBF* genes and they analyzed their gene expression patterns under NACC and ACC conditions. The cross-species comparisons indicated that the *CBFs* underwent to rapid expansion via duplication processes in *S. commersonii*. In total, four *S. commersonii CBFs* (*CBF1*, *CBF*2, *CBF*3, and *CBF*4) and two pseudo-genes of *CBF2* (*cCBF2*) and *CBF3* (*cCBF3*) were identified. The authors hypothesize that the duplication event occurred after the *S. tuberosum-S. commersonii* divergence and may have led to a different functionalization of the *ScCBF3* pseudogene, resulting in enhanced cold response capability in *S. commersonii*. Prior that the *S. commersonii* genome sequence was released, to shed further light on wild potato freezing tolerance, studies on *CBFs* genes were carried out using transgenic plants. For example, Pino et al. (2008) demonstrated that the over-expression of *AtCBF1* in transgenic lines of *S. tuberosum* and *S. commersonii* did not result in a further increase in freezing tolerance of cultivated potato, indicating that probably no additional cold-regulated gene beyond those regulated by *CBF* could increase freezing tolerance. Later, Carvallo et al. (2011) reported that both species have *CBF* regulons (genes regulated by *CBFs*) composed of hundreds of genes and both plants altered gene expression in response to low temperature with similar kinetics. However, there were considerable differences in the sets of genes that comprised the low temperature transcriptomes and *CBF* regulons. The results reported by Carvallo et al. (2011) indicated that the overexpression of *AtCBF3* up-regulated 160 cold-induced genes in *S. commersonii,* and only 54 in *S. tuberosum.* These results suggest that this difference in freezing tolerance is not due to 'macro-scale' differences in gene regulation in response to low temperatures or the size of their *CBF* regulons, but reveal rapid evolution of the *CBF* pathways in the two plant species that may contribute to their differences in freezing tolerance. Hence, to further investigate the genomic differences between the two species some genomic improvement are needed to discover new polymorphisms useful for plant breeding.

In line with this need, to further exploit *S. commersonii* sequence we have just performed a simple sequence repeat (SSR) recognition and analysis to identify genomic structural polymorphisms between *S. commersonii* and *S. tuberosum*. Using MISA (MIcroSAtellite) software and perl scripts, a total of 242,923 and 168,375 SSR loci were identified in *S. commersonii* and *S. tuberosum*, respectively (Esposito et al. 2017). A large proportion of mono-, di-, tri-, tetra-, penta- and hexa-nucleotides repeat motifs were identified in both species (Fig. 7).

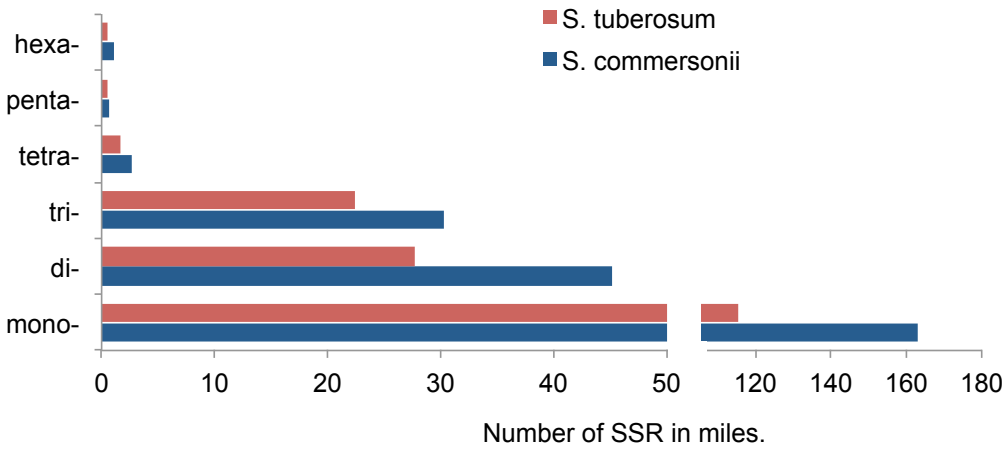


**Fig 8.** SSR loci identified in *S. commersonii* and *S. tuberosum* genomes. mono- = mononucleotides; di- = dinucleotides; tri- = trinucleotides; tetra- = tetranucleotides; penta- = pentanucleotides; hesa- = esanucleotides.

The mononucleotide repeats exhibited a strong bias towards A/T motif (62%) compared with C/G (5%). The AT/AT was the most common motif in both species, whereas the CG/CG was present at very low level, showing a consistent trend with those of many other plants species, such as apple (Guang L et al. 2012) and grape (Cai B et al. 2009). The comparison of SSR distribution and patterns in *S. commersonii* and *S. tuberosum* genomes showed that the former is significantly enriched in terms of SSR loci towards the latter. Comparative transcript profiling also suggested that in *S. tuberosum* post-transcriptional modifications (i.e. alternative splicing) were potentially more abundant. Indeed, more predicted proteins per gene were found in *S. tuberosum* (23,000) than in *S. commersonii* (4,900). The number of their genes, proteins and the specific annotation are represented in Fig. 8.

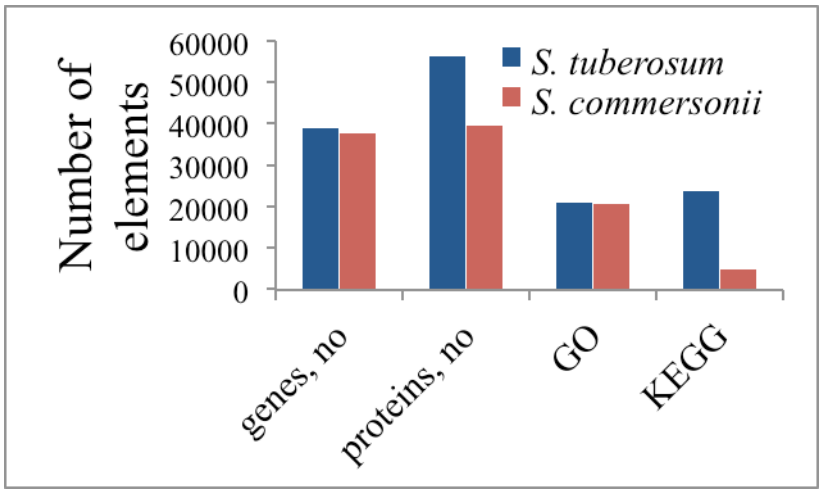


**Fig 9.** Number of predicted genes in *S. commersonii* and *S. tuberosum.* Genes with GO and KEGG pathways are also shown.

To understand also the type and number of RNA molecules in *S. commersonii,* transcriptomic analysis were carried out in five different tissues, namely flowers, leaves, roots, stolons and tubers. Only those genes with an FPKM value > 1 in at least one tissue were designated as expressed. Based on this criterion, 26.408 (71%) of the annotated genes were expressed in at least one tissue, indicating substantially higher representation of the transcriptome. Only a small quantity of genes was not expressed in the analyzed tissues. Diversity of transcriptional activity was highly variable across tissues, with flowers expressing the largest number of genes (≈ 24.000 of all genes), and the tubers expressing the smallest number of genes (less than 20.000) (Fig. 9). This diversity is enhanced also from the number of genes that were tissue-specific. Indeed, a total of 1.090 genes were unique in flowers while 736, 726 and 90 were expressed only in leaves, roots and stolons, respectively. Expansion of the atlas to include stressed tissues may provide a broader representation of the full *S. commersonii* transcriptome. This will help to identify the dynamic transcriptional profiles representing different cell types and developmental processes, provide new regulatory targets and allow the manipulation of specific pathways involved in the control of traits important for breeding.

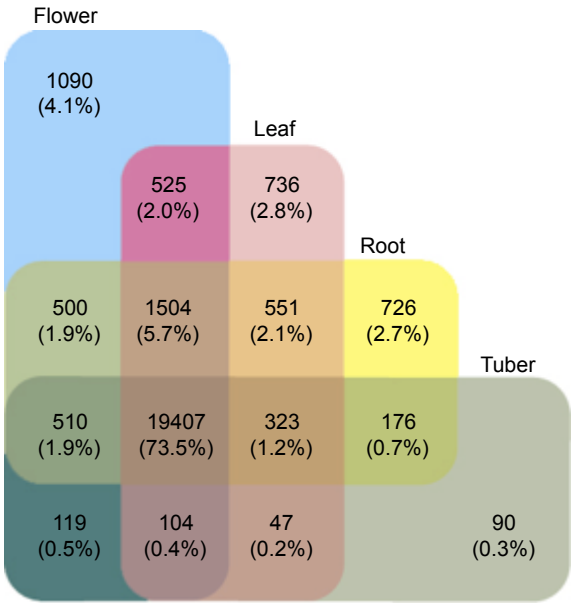


**Fig 10.** Venn diagram of genes expressed in different *S. commersonii* tissues.

Thanks to the availability of *S. commersonii* and *S. tuberosum* genome sequence, the characterization of their repetitive complements was also possible. Therefore, in *S. tuberosum* and *S. commersonii* we performed a de-novo retrotransposon (LTR-RTs) annotation. Of all structurally intact elements (SIE) identified, 2,976 and 4,010 showed the RT–INT order of the Retro-Transcriptase and Integrase domains typical for retroelements of the *Gypsy* superfamily in *S. commersonii* and *S. tuberosum*, whereas in 1,296 and 2,161 candidates, INT–RT domain hit order was observed, indicating the possible presence of retrotransposons of the *Copia* superfamily.

## 1.3 OBJECTIVES OF THE THESIS

It has been estimated that, with climate changes, the potato as well as all crop plants will be exposed to more unexpected events, ranging from abiotic to biotic stress conditions. These will reduce plant survival, production and geographic distribution (Beerling et al. 2000). Among abiotic stresses, cold is one of the suboptimal conditions that is more harmful to the cultivated potato, classified as frost-sensitive. Luckily, some wild potato species are frost hardy and capable of cold acclimation. Thus, they are a potential genetic resource for introgressing freezing tolerance into cultivated varieties. Among the wild potatoes, *S. commersonii* is the one displaying the highest tolerance to low temperatures and whose genome sequence has been deciphered (Aversano et al. 2015). Although the cold response machinery is widely studied in plant, the capacity of a genotype to tolerate low temperatures has not been fully understood yet and relatively little progress has been made in terms of breeding. This is partly due to the fact that the genetic control of cold tolerance is a quantitative and complex trait, with low heritability, presenting additive, dominance and epistatic gene actions. A combination of new and efficient approaches is necessary to accelerate the identification, characterization and effective exploitation of loci affecting tolerance to low temperatures not only in potato but in all crops. These strategies should include the use of genome sequences of target crops and related germplasm, genome-wide association studies, mutation detection, gene discovery and regulation, and –omics databases (Michelbart et al. 2015). With these thoughts in mind, this thesis is aimed at investigating the molecular mechanisms contributing to cold tolerance in the potato species possessing the highest capacity to withstand low temperatures, *S. commersonii*. The treatment is organized in five chapters. The first one is a general introduction underlying the importance of cold tolerance in plants, how they adapt to stress conditions, and some important genomic characteristics of *S. commersonii* vs. *S. tuberosum*. Some new original data have been also included, such as those related to SSR mining and transposable elements structure, distribution and evolution (paragraph 1.2.1). These new data have been recently published in a book chapter (Esposito et al. 2017). The other three chapters are organized in three original papers. Two of them (chapter II and chapter III) will be submitted, the other (chapter IV) has already been sent to Planta. The final chapter is dedicated to a general conclusion. In the next paragraph a few details on the specific objectives of chapter II, III and IV are given.

Chapter II deals with RNAseq analysis of 24 libraries from cold stressed and control plants of *S. commersonii.* They were sequenced using an Illumina technology. Two different clones of *S. commersonii,* contrasting in their cold response, were used. To better understand the genetic control of cold tolerance and considering the small amount of RNA-seq information available for this trait, our objectives were:

1. To generate a landscape of genes differentially expressed during cold stress, with particular attention to those uniquely expressed in the tolerant clone and to those in common between the tolerant and the sensitive clone but displaying opposite expression pattern..
2. To shed new lights on cold signaling and perception, in order to identify the candidates involved in signal transduction.

Chapter III concerns the identification of riboregulators involved in cold stress response. Indeed, recent studies have shown that abiotic stress induces aberrant expression of many small non-coding RNA (sncRNA) in several plant species. Among them, microRNAs (miRNAs) have been already demonstrated to play an essential role during cold stress in tomato. Although miRNAs and their targets have been identified in *S. tuberosum*, their role during cold accumulation in potato remains unknown. Further, no studies on miRNA as well as on secondary siRNA in *S. commersonii* have been reported to date. Therefore, in chapter III the following objectives were settled:

1. Predict and annotate miRNA and siRNA in *S. commersonii* genome. We developed 24 single-end smRNAseq libraries and homemade bioinformatics tools were experimented for the annotation. Towards this goal, the availability of the *S. commersonii* genomic sequence allowed us to identify putative miRNA involved in plant environment-interaction and development.
2. Once annotated, new miRNAs and siRNAs were analyzed to study their involvement in freezing tolerance and cold acclimation. To reach this objective, bioinformatics analyses were performed to report, for the first time, the identification of cold-inducible miRNAs and siRNAs in *S. commersonii*.

Chapter IV approaches the regulation of gene expression through RNA silencing. As known, it predominantly relies on the accurate functioning of Dicer-like (DCL), Argonaute (AGO) and RNA-dependent RNA polymerases (RDR) proteins, whose genes are present in multiple copies in eukaryotic genomes (Margis et al. 2006). DCL, AGO and RDR comprise the core components of RNA-induced silencing complexes, which trigger RNA silencing. These proteins are prominent players in the post-transcriptional control of gene expression, as they control small RNA-mediated gene silencing pathways and function in the epigenetic regulation of the genome under various environmental stresses (Yadav et al. 2015). here we pursued the following objectives:

1. To perform a genome wide analysis of *DCL* and *RDR* genes involved in miRNA and siRNA production in *S. tuberosum* and *S. commersonii*. This is felt important given that the diversity between wild and cultivated species in terms of candidate orthologous gene pairs with important role in RNAi has not received much attention.
2. To determine whether *DCL* and *RDR* genes possess a diverse regulation in different tissues and after cold stress. This object is in line with the fact that there is little information on proteins controlling *DCL* and *RDR* biogenesis in non-model wild species, which are known to often tolerate environmental stresses better than their closely related cultivated species.

**REFERENCES**

Aversano R, Contaldi F, Ercolano MR, et al (2015) The *Solanum commersonii* genome sequence provides insights into adaptation to stress conditions and genome evolution of wild potato relatives. Plant Cell 27:954-968.

Beerling DJ (2000) Global terrestrial productivity in the Mesozoic. Present and Future Climates. Geological Society of London Special Publication 181:17–32.

Boyko A, Blevins T, Yao Y, Golubov A, Bilichak A, Ilnytskyy Y, Hollunder J, Meins F Jr, Kovalchuk I (2010) Transgenerational adaptation of *Arabidopsis* to stress requires DNA methylation and the function of Dicer-like proteins. PLoS ONE 5:9514

Cardi T, D'Ambrosio E, Consoli D, Puite KJ, and Ramulu KS (1993). Production of somatic hybrids between frost-tolerant *Solanum commersonii* and *S. tuberosum*: characterization of hybrid plants. TAG Theor Appl Genet 87:193-200.

Carputo D, Barone A, Cardi T, Sebastiano A, Frusciante L, and Peloquin SJ (1997) Endosperm balance number manipulation for direct in vivo germplasm introgression to potato from a sexually isolated relative (*Solanum commersonii Dun*.). Proc. Natl. Acad. Sci. U.S.A. 94:12013–12017.

Carvallo MA, Pino MT, Jeknic Z, Zou C, Doherty CJ, Shiu SH, Chen THH, and Thomashow M (2011). A comparison of the low temperature transcriptomes and *CBF* regulons of three plant species that differ in freezing tolerance: *Solanum commersonii*, *Solanum tuberosum*, and *Arabidopsis thaliana*. J Exp Bot 62:3807–3819.

Chen HH and PH Li 1980 Biochemical changes in tuber-bearing Solanum species in relation to frost hardiness during cold acclimation. Plant Physio l66:414-421

Chen HH and PH Li 1980 Characteristics of cold acclimation and deacclimation in tuber-bearing *Solanum* species. Plant Physiol 65: 1146-1148

Doherty CJ, Van Buskirk HA, Myers SJ, Thomashow MF (2009) Roles for *Arabidopsis* CAMTA transcription factors in cold-regulated gene expression and freezing tolerance. Plant Cell 21:972–984.

Dong CH, Agarwal M, Zhang Y, Xie Q, Zhu JK (2006) The negative regulator of plant cold responses, *HOS1*, is a RING *E3* ligase that mediates the ubiquitination and degradation of ICE1. Proc. Natl. Acad. Sci. USA 103:8281–8286.

Esposito S, D'Amelia V, Villano C, Contaldi F, Carputo D, Aversano R (2017) The Wild Side of Potato: Insights into the Genome Sequence of the Stress-Tolerant *S. commersonii* Dun. In The Potato Genome. Eds. Chakrabarti SK, Conghua X and Tiwari JK. Springer. In press. ISBN 978-3-319-66133-9

González M, Galván G, Siri, M. I., Borges, A., and Vilaró, F. (2013) Resistencia a la marchitez bacteriana de la papa en *Solanum commersonii*. Agrociencia Urug 7;45–54.

Hara M, Terashima S, Kuboi T (2001) Characterization and cryoprotective activity of cold-responsive dehydrin from *Citrus unshiu*. J Plant Physiol 158:1333–1339.

Hashida SN, Uchiyama T, Martin C, Kishima Y, Sano Y, Mikami T (2006) The temperature-dependent change in methylation of the *Antirrhinum* transposon Tam3 is controlled by the activity of its transposase. Plant Cell 18:104–118

Hawkes, JG (1989) Nomenclatural and taxonomic notes on the infrageneric taxa of the tuber-bearing Solanums (*Solanaceae*). Taxon 38:489-492.

Houde M., Dallaire S., N'Dong D., Sarhan F. Overexpression of the acidic dehydrin (2004) WCOR410 improves freezing tolerance in transgenic strawberry leaves. Plant Biotechnol J 2:381–387.

Ingram J, Bartels D (1996) The molecular basis of dehydration tolerance in plants. Annu Rev Plant Physiol Plant Mol Biol 47:377−403

Ito H, Gaubert H, Bucher E, Mirouze M, Vaillant I, Paszkowski J. (2011) An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress. Nature 472:115–119.

Jackson SA and Hanneman RE (1999) Crossability between cultivated and wild tuber-and non-tuber-bearing *Solanum*. Euphytica 109:51-67.

Janskà, Marsik P, Zelenkov´ a S, Ovesn´ a J (2010) Cold stress and acclimation – what is important for metabolic adjustment? Plant Biol 12:395–405

Kalinka A, Achrem M, and Rogalska S. (2009) Application of BSP method in methylation pattern comparison of reverse transcriptase (rt) gene in wheat-rye hybrids and their parental species. In: Naganowska B, Kachlicki P, Krajewski P [ed.], Genetyka i Genomika w Doskonaleniu Roślin Uprawnych, 53–61. Instytut Genetyki Roślin PAN. Poznań.

Kiyosue T, Yamaguchi-Shinozaki K, Shinozaki K (1994) Characterization of two cDNAs (*ERD10* and *ERD14*) corresponding to genes that respond rapidly to dehydration stress in *Arabidopsis thaliana*. Plant Cell Physiol 35:225−231

Knight MR and Knight H. (2012) Low-temperature perception leading to gene expression and cold tolerance in higher plants. New Phytol 195:737–751.

Knox AK, Dhillon T, Cheng HM, Tondelli A, Pecchioni N, Stockinger EJ. (2010). *CBF* gene copy number variation at *Frost Resistance-2* is associated with levels of freezing tolerance in temperate-climate cereals. Theoretical and Applied Genetics 121:21–35.

Kovacs D, Kalmar E, Torok Z, Tompa P (2008) Chaperone activity of ERD10 and ERD14, two disordered stress-related plant proteins. Plant Physiol 147:381−390

Lee BH, Henderson DA, and Zhu JK (2005) The *Arabidopsis* cold-responsive transcriptome and its regulation by *ICE1*. Plant Cell. **17**: 3155–3175.

Lu SF, Sun YH, Chiang VL (2008). Stress-responsive microRNAs in *Populus*. Plant J 55:131–151.

Luan S, Kudla J, Rodriguez-Concepcion M, Yalovsky S, Gruissem W (2002) Calcium sensors for specific signal response coupling in plants. Plant Cell.;14:S389–400.

Maruyama K, Sakuma Y, Kasuga M, Ito Y, Seki M, Goda H, Shimada Y, Yoshida S, Shinozaki K and Yamaguchi-Shinozaki K (2004) Identification of cold-inducible downstream genes of the *Arabidopsis DREB1A/ CBF3* transcriptional factor using two microarray systems. Plant J 38:982–993.

Mendiburu AO and Peloquin SJ. (1997) The significance of 2N gametes in potato breeding. Theoretical and Applied Genetics 49:53–61

Micheletto S, Boland R, and Huarte M (2000) Argentinian wild diploid *Solanum* species as sources of quantitative late blight resistance - Springer. Theoretical and Applied Genetics.

Mittler R, Blumwald E. (2010) Genetic engineering for modern agriculture: challenges and perspectives. Annual Review of Plant Biology 61:443-462)

Miura K, Rus A, Sharkhuu A, Yokoi S, Karthikeyan AS, Raghothama KG, Baek D, Koo YD, Jin JB, Bressan RA, et al. (2005) The *Arabidopsis* SUMO E3 ligase *SIZ1* controls phosphate deficiency responses. Proc. Natl. Acad. Sci. USA 102:7760–7765.

Murata N, Ishizaki-Nishizawa O, Higashi S, Hayashi H, Tasaka Y, Nishida I. (1992) Genetically engineered alteration in the chilling sensitivity of plants. Nature 356:710–713.

Papademetriou MK (2008) In: RAP Publication (FAO), no. 2008/07; Workshop to commemorate the international year of potato, Bangkok (Thailand), 6 May 2008/FAO, Bangkok (Thailand). Regional Office for Asia and the Pacific, p 84

Park S, Lee CM, Doherty CJ, Gilmour SJ, Kim Y, Thomashow MF (2015) Regulation of the *Arabidopsis CBF* regulon by a complex low-temperature regulatory network. Plant J 82: 193–207
Philosophical Transactions of the Royal Society B: Biological Sciences, 2008 363:639-658

Pino MT, Skinner JS, Park EJ, Jeknić Z, Hayes PM, Thomashow MF, et al. (2008) Use of a stress inducible promoter to drive ectopic *AtCBF* expression improves potato freezing tolerance while minimizing negative effects on tuber yield. Plant Biotechnol J 5:591–604.

Porter JR, Semenov MA (2005) Crop responses to climatic variation, Philosophical Transactions of the Royal Society B: Biological Sciences 360:2021-2035

Rodríguez F and Spooner DM (2009) Nitrate Reductase Phylogeny of Potato (*Solatium* sect. Petota) Genomes with Emphasis on the Origins of the Polyploid Species. Systematic botany. Systematic Botany.34:207-219.

Sakuma Y, Liu Q, Dubouzet JG, Abe H, Shinozaki K, and Yamaguchi-Shinozaki K (2002) 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. 290998–1009.

Scott GJ, Rosegrant MW and Ringler C **(**2000) Roots and tubers for the 21st century – trends, projections, and policy options. International Food Policy Research Institute. http://www.ifpri.org.

Seo PJ M,m , Lee AK, Xiang F, Park CM (2008). Molecular and functional profiling of *Arabidopsis* pathogenesis-related genes: insights into their roles in salt response of seed germination. Plant and Cell Physiology 49:334-344

Steward et al. (2002) Steward N, Ito M, Yamaguchi Y, Koizumi N, Sano H. Periodic DNA methylation in maize nucleosomes and demethylation by environmental stress. Journal of Biological Chemistry 277:37741–37746

Stockinger EJ, Mao, Y, Regier MK, Triezenberg SJ, and Thomashow MF (2001) Transcriptional adaptor and histone acetyltransferase proteins in *Arabidopsis* and their interactions with *CBF1*, a transcriptional activator involved in cold-regulated gene expression. Nucleic Acids Res 29:1524–1533.

Stone JM, Palta JP, Bamberg JB, Weiss LS, and Harbage JF (1993) Inheritance of freezing resistance in tuber-bearing *Solanum* species: evidence for independent genetic control of nonacclimated freezing tolerance and cold acclimation capacity. Proc Natl Acad Sci USA 90: 7869–7873.

Thomashow MF (2010) Molecular basis of plant cold acclimation: insights gained from studying the CBF cold response pathway. Plant Physiol 154:571–577.

Townley HE, Knight MR (2002) Calmodulin as a potential negative regulator of *Arabidopsis COR* gene expression. Plant Physiol 128:1169–1172

Vogel JT, Zarka DG, Van Buskirk HA, Fowler SG and Thomashow MF (2005) Roles of the *CBF2* and *ZAT12* transcription factors in configuring the low temperature transcriptome of *Arabidopsis*. Plant J 41:195–211.

Yadav CB, Muthamilarasan M, Pandey G, Prasad M (2015) Identification, characterization and expression profiling of *Dicer-Like*, *Argonaute* and *RNA-dependent RNA polymerase* gene families in foxtail millet. Plant Mol Biol Rep 33:43–55.

Yao Y, Ni Z, Peng H, Sun F, Xin M, Sunkar R, et al. (2010) Non-coding small RNAs responsive to abiotic stress in wheat (*Triticum aestivum* L.). Funct Integr Genomics 10:187–190

Zhou X, Wang G, Sutoh K, Zhu JK, Zhang W (2008) Identification of cold-inducible microRNAs in plants by transcriptome analysis. Biochim Biophys Acta 1779:780−788.

Zuluaga Cruz AP, Ferreira V, Julia Pianzzola M, Ines Siri M, Coll NS, Valls M, et al. (2014) A novel, sensitive method to evaluate potato germplasm for bacterial wilt resistance using a luminescent *Ralstonia solanacearum* reporter strain. Mol Plant Microbe Interact 27:277–285.

s

# Chapter II. Transcriptomic profiling of cold tolerant *Solanum commersonii* reveals new insights in the response to low temperatures

**Transcriptomic profiling of cold tolerant *Solanum commersonii* reveals new insights in the response to low temperatures**

**Abstract**

Low temperatures lead to major crop losses every year. Although several studies have been conducted focusing on the molecular mechanisms involved in cold tolerance in several crops, including wild and cultivated potatoes, transcriptome-scale molecular understanding is still lacking. Taking advantage from *Solanum commersonii* genome sequence, we carried out RNAseq analysis of two different clones of *S. commersonii.* One is cold tolerant and able to cold acclimate, the other is cold sensitive but able to acclimate. We developed 24 RNAseq libraries from leaf tissues subjected to non-acclimated and acclimated conditions, and produced one billion reads using Illumina high-throughput RNA sequencing. Results showed that more than 450 million reads could be mapped uniquely to *S. commersonii* genome, whereas other 300 million reads mapped to multiple loci. Data indicated that in not acclimated stress condition, a lower number of differentially expressed genes is present in the tolerant clone vs. the sensitive one. By contrast, when acclimated, both clones showed a similar behavior in terms of differentially expressed genes. New lights regarding *CBF* expression and their regulation have been also given. Among them, *CBF3*, *CBF4* and *ZAT12* genes are the major candidate, which might play an important role in conferring cold tolerance in *S. commersonii*. Indeed, under non-acclimated conditions, they were induced in the cold tolerant clone, whereas they showed an opposite trend in the cold sensitive. By contrast, following cold acclimation, both clones induced these genes. The mechanisms described in the present work will be useful for future investigations and for the detailed validation in marker assisted selection projects for cold tolerance in potato.

## 2.1. Introduction

How plants respond to abiotic stresses has been an important goal for plant scientist and breeders since many decades. Indeed, crop yield reduction as a consequence of climatic events is threatening global food security (Bailey-Serres et al. 2012; Boyer et al. 2013). Among abiotic stresses, low temperature is a major factor affecting crop productivity. Nowadays, what is known from the literature is that plants use either stress tolerance or stress avoidance through acclimation and adaptation mechanisms to withstand low temperatures. Several studies showed that biochemical and physiological changes occur under cold stress, resulting in modifications of the plant cell structural, biochemical and photosynthetic proprieties (Stitt et al. 2002; Goulas et al. 2006). Changes in gene expression and protein accumulation also occur, including up-regulation of antioxidant proteins

(e.g. superoxide dismutase (SOD), catalase (CAT), ascorbate peroxidase (APX), synthesis and accumulation of osmoprotectants (polyamines and prolin) and sugars (sucrose, maltose, glucose and fructose) (Lissarre et al. 2010; Janskà et al. 2010). In most of cases, the function of these proteins under low temperatures is yet to be conclusively shown. However, it is generally assumed that there is a strong correlation between their accumulation and the development of freezing tolerance and (Dahal et al. 2012). Despite an exponential increase in the understanding of the molecular dynamics of cold tolerance and acclimation, their comprehension in a manner that allow breeders to improve this trait in economically important crops has provided to be an elusive goal. This is partially due to the complexity of cold tolerance as already shown for rice (da Maia et al. 2017). One of the major advances in recent years has been the discovery of numerous cold-responsive (*COR*) genes, which enable the plant to tolerate the dehydration stress associated with freezing temperatures and the presence of extracellular ice (Thomashow et al. 2001; Shinozaki et al. 2003; Chinnusamy et al. 2007). In turn, these are influenced by the expression of regulatory factors, namely *CBFs* (C-repeat binding factors) (van Buskirk and Thomashow, 2006; Chinnusamy et al. 2007). They have been identified as tandem duplicated in the model species *Arabidopsis thaliana* as well as in other plants species. For example, the Frost Resistance2 (FR2) in cereals define a region containing numerous duplicated *CBF* genes, most of them present as tandem repeats with highly conserved coding sequence (Skinner et al. 2005; Badawi et al. 2007). It has been shown that *CBF* overexpression increased the freezing tolerance of *A. thaliana* (Liu et al. 1998), *Brassica napus* (Jaglo et al. 2001), poplar (Benedict et al. 2006) and potato (Pino et al. 2007), but did not result in increased freezing tolerance in tomato (Zhang et al. 2004) and rice (Dubouzet et al. 2003). The reason for this difference is not clear, but it might be due to differences in the composition of the *CBF* regulons (genes regulated by specific *CBF*). Some additional studies revealed that differential expression of a cold-responsive gene is caused by differences in cold tolerance in plants (Liu et al. 2012; Dong et al. 2014). For example, specific modifications such as differences in physiology and global gene expression occur in cold tolerant *Solanum habrochaites* but not in not in the susceptible *S. lycopersicum* (Liu et al. 2012). These findings were also observed in *Brassica rapa,* where many genes involved in cold stress response were specifically expressed in the tolerant genotype (Dong et al. 2014). Therefore, given the demonstrated role of cold-regulated genes in plant cold tolerance, it is reasonable to think that differences in gene expression are likely to contribute to differences in the ability of plants to withstand cold. This could explain why some plants within species may be killed at low temperatures while others are not. Nowadays, recent advances in the availability of sequenced plants genomes and the ability to sequence entire transcriptomes using next generation sequencing (NGS) has added a valuable interest to study the dynamics of cold stress. Hence, these

thinking motivate comparative transcriptome studies using contrasting genotypes to identify genes with critical roles in freezing tolerance, and to better understand the molecular basis of cold tolerance, as already reported in tomato (Chen et al. 2015) and rice (da Maia et al. 2017).

Potato (*S. tuberosum*) is one of the most important cultivated species for food production, being consumed by over half of the world's population (Fao 2012). Due to a large increase in the world population forecasted for the next decades, an increase in potato yield and production is needed. However, low temperature is one of the suboptimal conditions that are more harmful to potato. In fact, similar to several other plants species, *S. tuberosum* is classified as frost-sensitive species (Chen & Li 1980). It is also unable to acclimate. By contrast, some wild potato species are much frost hardy and capable of cold acclimation. Thus, they are a potential powerful genetic resource for introgressing freezing tolerance traits into cultivated varieties. Among the wild potato species, *S. commersonii* is the one possessing genotypes with the highest tolerance to low temperatures. It can survive to about -5 °C pre-acclimation, and to as low as -11 °C after becoming fully cold acclimated (Chen & Li 1980). Importantly, it is also the first potato relative whose genome sequence has been deciphered (Aversano et al. 2015). In light of the reasons described above, and to better understand the genetic control and the molecular dynamics of *S. commersonii* capacity to face low temperatures, we carried out a comparative transcriptome study (RNA-seq) using two different clones of *S. commersonii* contrasting in their cold response. This allowed us to evaluate differentially expressed genes underlying why one clone can survive freezing temperatures and acclimate, whereas the other is susceptible but can acclimate.

## 2.2. Materials and Methods

### 2.2.1 Plant material, growth conditions and treatments.

We used two clones of *S. commersonii* belonging to two different accessions: cmm1T (clone of PI243503) and cmm6-6 (clone of PI590886). The former is freezing tolerant and able to cold acclimate. The latter is susceptible but has the capacity to cold acclimate (Fig.7 chapter I). They also display phenotypic differences as shown in Fig 1 and Fig 2. Four-weeks old vitroplants were transplanted into 14-mm pots filled with sterile soil and grown for two weeks at 24°C in growth chamber prior submitting them to non-acclimated (NACC) and acclimated (ACC) stress conditions. In particular, in NACC experiment three plants of each genotype were challenged for 30min at -2°C, while three plants were chosen as control and kept at 24°C. In the ACC experiment, 6 plants/genotype were acclimated at 4°C for 2 weeks. Then, three of them were transferred for 30min at -2°C, while the others were used as control and kept at 4°C. In both experiments, an environmentally controlled cold room was used. Young leaf samples were collected from all

replicates at the end of each stress and from control plants. Samples were individually stored at -80°C before RNA extraction.

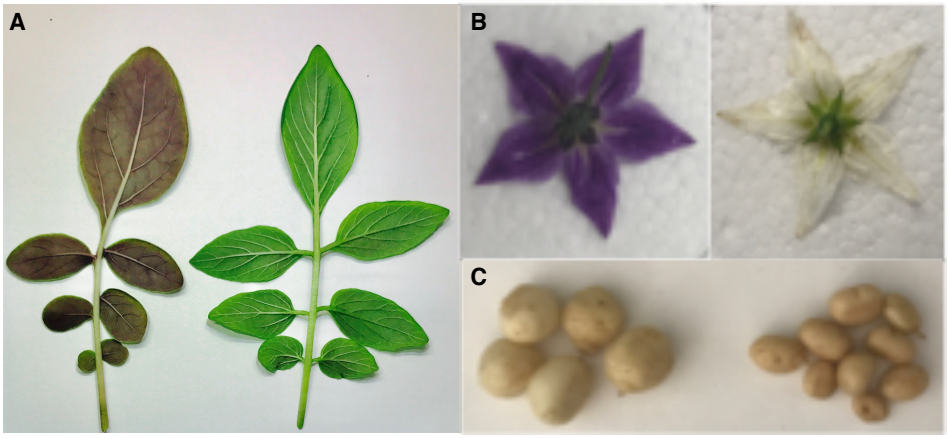


**Fig. 1** Phenotypic variation between cmm1T and cmm6-6 in leaves (A) flowers (B) and tubers (C). Cmm1T displayed anthocyanin accumulation under leaf and in the flower.

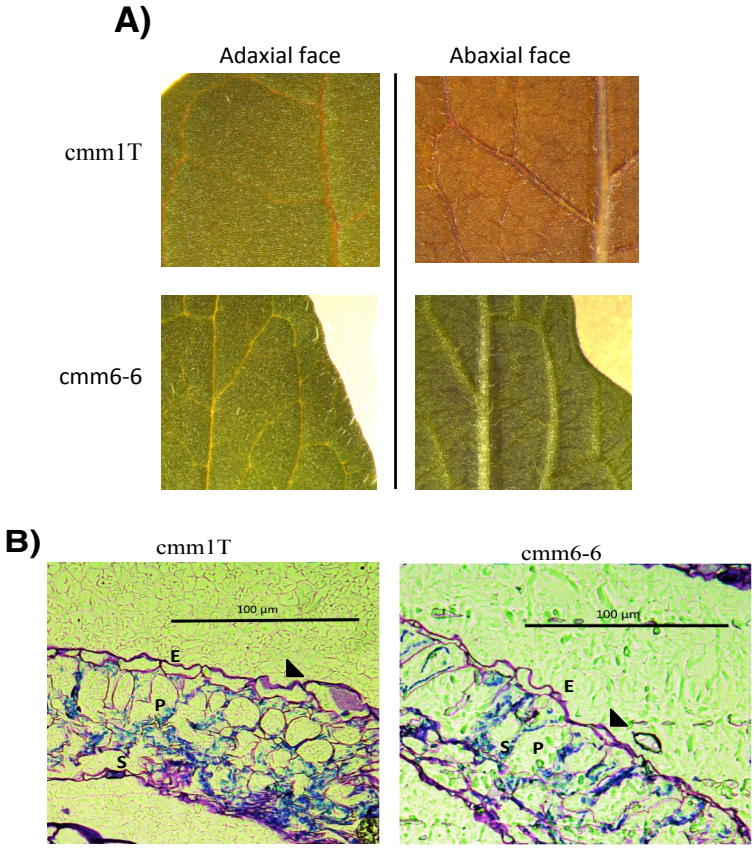


**Fig. 2** Details of cmm1t and cmm6-6 leaves observed by stereomicroscope (A). Main differences observed were attributed to the presence of trichomes on the edges of cmm6-6 abaxial face (60X magnification). Light micrographs of cross sections of leaves of cmm1T and

cmm6-6 (B). E; epidermal cells; P, palisade parenchyma cells; S, spongy parenchima cells. Differences in the thickness of cmm1T and cmm6-6 epidermal cell layers (indicated by black arrows) is particularly evident. Samples were fixed in FAA (40% formaldehyde : glacial acetic acid : 50% ethanol—5 : 5 : 90 by volume), stained with 0.5% toluidine blue and mounted in 10 % glycerol. Scale bar length is shown in the figure.

### 2.2.2 RNA extraction and sequencing

For each sample total RNA was isolated from leaf tissues using Spectrum Plant Total RNA kit (Sigma-Aldrich, St. Louis, MO, USA) according to the manufacturer's protocol. RNA concentrations were determined using NanoDrop ND-1000 spectrophotometer (Thermo Scientific, Wilmington, USA) and its integrity was verified was checked using bioanalyzer (Agilent Technologies, Santa Clara, California, USA). Three μg of total RNA of each sample was sent to UMN Genomic Center (University of Minnesota, USA) for libraries preparations. Twenty-four cDNA libraries (three biological replicates from leaves in control and stress conditions) were subsequently prepared for RNA-seq experiments with the Illumina HiSeq 2500 sequencing platform providing 125bp paired end reads, for a total of 30M reads/samples (Tab. 1). Trimming and clipping were performed with Trimmomatic-0.3330 using default parameters. The quality assessment was based on the remaining reads using the FASTQC quality control tool version 0.10.0 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc).

**Tab. 1** List of samples used to prepare cDNA libraries in RNAseq experiments. The number of biological replicates in each condition is shown. NACC= non-acclimated conditions, ACC= acclimated conditions

| Sample | Species | Condition | Tolerance/susceptibility |
|---|---|---|---|
| 1 | cmm1T | control NACC plant 1 | cold tolerant |
| 2 | cmm1T | control NACC plant 2 | cold tolerant |
| 3 | cmm1T | control NACC plant 3 | cold tolerant |
| 4 | cmm1T | NACC 30 min plant 1 | cold tolerant |
| 5 | cmm1T | NACC 30 min plant 2 | cold tolerant |
| 6 | cmm1T | NACC 30 min plant 3 | cold tolerant |
| 7 | cmm1T | control ACC plant 1 | cold tolerant |

| | | | |
|---|---|---|---|
| 8 | cmm1T | control ACC plant 2 | cold tolerant |
| 9 | cmm1T | control ACC plant 3 | cold tolerant |
| 10 | cmm1T | ACC 30 min plant 1 | cold tolerant |
| 11 | cmm1T | ACC 30 min plant 2 | cold tolerant |
| 12 | cmm1T | ACC 30 min plant 3 | cold tolerant |
| 13 | cmm6-6 | control NACC plant 1 | cold susceptible |
| 14 | cmm6-6 | control NACC plant 2 | cold susceptible |
| 15 | cmm6-6 | control NACC plant 3 | cold susceptible |
| 16 | cmm6-6 | NACC 30 min plant 1 | cold susceptible |
| 17 | cmm6-6 | NACC 30 min plant 2 | cold susceptible |
| 18 | cmm6-6 | NACC 30 min plant 3 | cold susceptible |
| 19 | cmm6-6 | control ACC plant 1 | cold tolerant |
| 20 | cmm6-6 | control ACC plant 2 | cold tolerant |
| 21 | cmm6-6 | control ACC plant 3 | cold tolerant |
| 22 | cmm6-6 | ACC 30 min plant 1 | cold tolerant |
| 23 | cmm6-6 | ACC 30 min plant 2 | cold tolerant |
| 24 | cmm6-6 | ACC 30 min plant 3 | cold tolerant |

### 2.2.3 Reads mapping to the reference genome and differential expression analysis

In order to map each cleaned library to the wild potato genome, all libraries were loaded in A.I.R. on line program (https://transcriptomics.cloud) and a new RNA-seq experiment was chosen. Gene expression levels were calculated using geometric normalization and per-condition dispersion method by quantifying the Illumina reads according to the FPKM (fragments per kilobase per million mapped fragments). These values were used to perform a principal component analysis (PCA) and to check experimental and control biological replicates. Fold-changes were reported as the log (base 2) of normalized read count abundance for the cold-stressed samples divided by the read count abundance of the control samples. To study the impact of cold stress on gene expression four different statistical approaches were used: DEseq2 (Love et al. 2014), EBseq (Leng et al. 2013), EdgeR (Robinson et al. 2010) and NOIseq (Tarazona et al. 2013). These programs were chosen because they use different ways to model the negative binomial dispersion parameter and are also the most used in literature (Cumbie et al. 2011; Rapaport et al. 2013; Soneson et al. 2013). All genes that were in common in all four tests were used for further analysis. Each dataset obtained from each considered condition was filtered according to fold change values $\geq 1.5$ and $\leq -1.5$.

### 2.2.4 Functional annotation

Gene ontology terms were examined with agriGO (http://bioinfo.cau.edu.cn/agriGO/) for GO enrichment with custom annotation. For this analysis, the following parameters were chosen: hypergeometric statistical test method, multi-test adjustment hockberg FDR, significance level: <0.05 and 3 minimum number of mapping entries. Significant values were sorted by enrichment score (Query_item/Query_total)/(Background_item/Background_total) and GO redundancy was removed with REVIGO tool (http://revigo.irb.hr). MAPMAN software was used to further understand the biological role of differentially genes. The *S. tuberosum* MAPMAN ontologies were retrieved from the GOMAPMAN web resource (Ramsak et al. 2014), and imported in the MAPMAN tool (v.3.6.0) (Thimm et al. 2004). The list of orthologous differentially expressed genes (DEGs) identified in *S. tuberosum* was then mapped to bins for data visualization and pathway analysis.

### 2.2.5 RNAseq data validation

To validate the reliability of the expression profiles observed in the RNA-seq data, 10 genes randomly chosen were used for quantitative real-time PCR (qPCR) analyses using iTaq SYBR Green supermix (Bio-Rad, Munich, Germany). The elongation factor gene *EF* was used as an internal control (Nicot et al. 2005). RNA material from the same samples employed for RNA-seq experiment was used for this validation. A 2µl aliquot of 1/10 cDNA was used in a qRT-PCR, with the addition of 0.3µM of each specific primer and FAST SYBR Green master mix (Applied Biosystems, Foster City, CA) to a final reaction volume of 20µl. The qRT-PCRs were performed using an ABI 7900HT Real Time PCR System (Applied Biosystems) and the relative expression value was calculated through the $\Delta\Delta$Ct method (Livak and Schmittgen 2001).

## 2.3. Results

### 2.3.1 RNA-seq statistics and DEGs

More than one billion reads were produced, with in average 33M reads from each sample. High-throughput sequencing and subsequent read trimming/clipping delivered a total of 850 million reads in control and stressed conditions. These were sub sequentially mapped against the *S. commersonii* scaffolds. Using default criteria, from 73% to 90% of the reads mapped to a genomic location. On

average, the uniquely mapped reads percentage was ≅50%, while that of multiple mapped reads was ≅40%.

***NACC experiment***

A total of 19.322 and 20.203 transcripts was obtained from cmm1T and cmm6-6, respectively (Fig. 3).

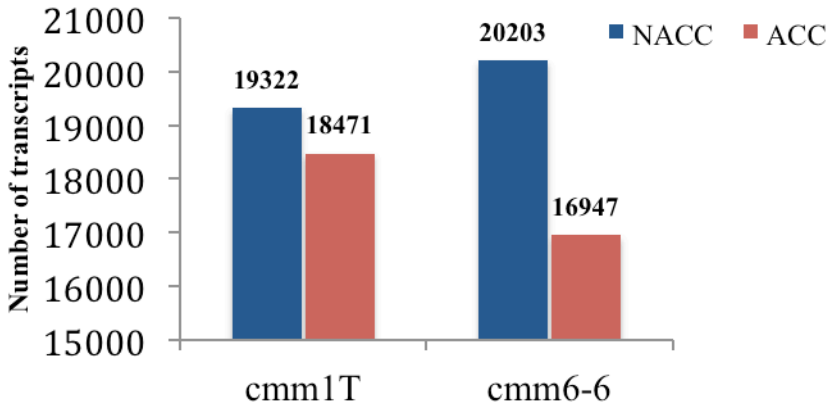


**Fig. 3** Total amount of expressed transcripts under non-acclimated (NACC) and acclimated (ACC) cold stress conditions in cmm1T and cmm6-6.

Venn diagrams were constructed to highlight uniquely and common genes identified by each statistical approach. In NACC conditions, cmm1T showed a lower number of DEGs than cmm6-6. Indeed, 193 (0.99 % of total transcripts), 206 (1.1%), 430 (2.2%) and 1.001 (5.1%) DEGs were found in cmm1T using EdgeR, DEseq2, EBseq and NOIseq, respectively. By contrast, 1.441 (7.1%), 1.135 (5.6%), 1.213 (5.9%) and 8.054 (39.8%) DEGs were found in cmm6-6 with the same statistical approaches. Only those DEGs that were consistently identified with all methods were kept for further analysis. Therefore, 71 and 883 genes were further analyzed in cmm1T and cmm6-6, respectively. All these common DEGs were annotated according to *S. commersonii* available annotations and to *A. thaliana* for those lacking annotation. The two sets were compared to reveal genes that were commonly or uniquely expressed. Of all DEGs identified, only two were in common between cmm1T and cmm6-6. The others (69 in cmm1T and 881 in cmm6-6) were unique (Fig 4).

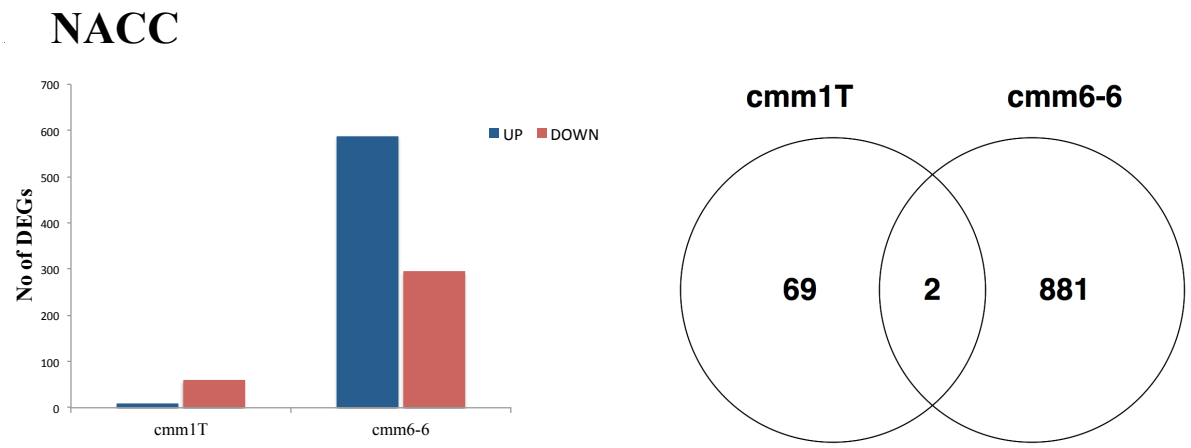


**Fig. 4** Number of total differentially expressed genes (DEGs) in cmm1T and cmm6-6 under NACC conditions and those that were in common or uniquely expressed.

An additional difference between our clones was relative to the behavior of DEGs. Indeed, out of 71 DEGs identified in cmm1T, 61 (86 %) were down-regulated while only 10 (14 %) were up-regulated (Fig 4). Further, most of the suppressed genes (33/61) showed a fold change < -1.5 (FDR value ≤ 0.05), whereas none of the induced genes had a fold change > 1.5. By contrast, in cmm6-6 the up-regulation was prevalent. Out of 883 DEGs identified, 588 (67 %) were up-regulated and 29% of them showed a fold change grater than 1.5. Two hundred and ninety five (33%) were down-regulated, with 30% of them showing a fold change grater than -1.5. MapMan software was employed to better understand the biological role of DEGs in both clones. Since a physical map for *S. commersonii* is not available yet, we used the equivalent orthologs in *S. tuberosum* to map our DEGs in MapMan bins. We found that 58 (out of 71) and 714 (out of 883) DEGs in cmm1T and cmm6-6 respectively, had at least one ortholog in *S. tuberosum*. For cmm1T, MapMan analysis showed that most of genes were involved in processes such as RNA regulation (9/58; bin27), cell wall synthesis (6/58; bin10), transport (4/58; bin34), hormone metabolism (3/58; bin17), stress (3/58) and others (Tab. 2A).

**Tab. 2** MapMan overview in NACC experiments. The number of differentially expressed genes (DEGs) in cmm1T and cmm6-6 corresponding to each MapMan bins is shown. –[a] mean not present.

**A**

| Bin | Name | Number of DEGs | |
|---|---|---|---|
| | | cmm1T | cmm6-6 |
| 1 | PS | -[a] | 46 |
| 2 | major CHO metabolism | 1 | 10 |
| 3 | minor CHO metabolism | 1 | 8 |
| 4 | glycolysis | - | 4 |
| 7 | OPP | - | 1 |
| 8 | TCA / org. transformation | - | 4 |
| 9 | mitochondrial electron transport / ATP synthesis | - | 3 |
| 10 | cell wall | 6 | 21 |
| 11 | lipid metabolism | 2 | 17 |
| 12 | N-metabolism | - | 3 |
| 13 | amino acid metabolism | - | 13 |
| 15 | metal handling | 1 | 1 |
| 16 | secondary metabolism | 3 | 22 |
| 17 | hormone metabolism | 3 | 29 |
| 18 | Co-factor and vitamine metabolism | - | 1 |
| 19 | tetrapyrrole synthesis | - | 7 |
| 20 | stress | 3 | 41 |
| 21 | redox | - | 12 |
| 22 | polyamine metabolism | - | 2 |
| 23 | nucleotide metabolism | - | 2 |
| 24 | Biodegradation of Xenobiotics | 1 | 1 |
| 25 | C1-metabolism | - | 1 |
| 26 | misc | 4 | 62 |
| 27 | RNA regulation | 9 | 91 |
| 28 | DNA | 1 | 7 |
| 29 | protein | 3 | 72 |
| 30 | signalling | 3 | 41 |
| 31 | cell | 4 | 17 |
| 33 | development | - | 13 |
| 34 | transport | 4 | 46 |
| 35 | not assigned | 9 | 129 |

Among DEGs involved in RNA regulation (bin27), 9 were transcription factors belonging to AP2/ERF, bHLH, MYB, HB families. DEGs belonging to bin10 were near identical to *AtIRX1, AtIRX3, AtIRX5* and *AtIRX6,* also called *CESA* genes (cellulose synthase). One gene was involved in lipid metabolism (acyl-carrier protein 4, similar to *AT4G25050*), confirming that changes in lipid composition in membrane might occur following cold stress. Genes with functions associated to signal perception and transduction were also found in the gene sets. For instance, two *NBS-LRR* were found to be differentially expressed in cmm1T and they might be considered as major

candidates for cold signal perception. Our data also indicated that two genes involved in signaling mechanisms, annotated as calcium binding and phosphatydilinositol-4-phosphate, were up-regulated in cmm1T. As far as MapMan analysis in cmm6-6 is concerned, in general a higher number of DEGs (129/714; bin35) did not show an annotated function (Tab. 2A), suggesting that a large amount of cold responsive genes is not yet fully characterized. The second most abundant bins were relative to RNA process (91/714DEGs; bin27). Among them, several genes were annotated as transcription factors: 12 Zn-finger (C2C2-CO like), 8 bHLH, 7 AP2/ERF, 6 HSF and 5 HB, WRKY and GARP were found. The other bins were relative to proteins (72/714; bin29), misc (62/714; bin26), transport (46/714; bin34) and signaling (41/714; bin30). Further, some bins such as PS (46/714; bin1), glycolisil mechanisms (4/714; bin4), OPP (1/714; bin7), TCA metabolism (4/714; bin8), mitochondrial electron transport (3/714; bin9), N-metabolisms (3/714; bin12), amminoacid transport (13/714; bin13), Co-factor and vitamin metabolism (1/714; bin18), tetrapyrrol synthesis (7/714; bin19), redox (12/714; bin21), polyamine metabolism (2/714; bin22), nucleotide metabolism (2/714; bin23), C1-metabolism (1/714; bin25) and development (13/714; bin33) were enriched only in cmm6-6 and not in cmm1T (Tab. 2A).

***ACC experiment***

A total of 18.471 and 16.947 transcripts were obtained from cmm1T and cmm6-6 following cold stress in acclimated conditions (Fig. 3). As in NACC, cmm1T confirmed a lower number of DEGs than cmm6-6. Indeed, a total of 101, 53, 194 and 5.651 DEGs were found in cmm1T using EdgeR, DEseq2, EBseq and NOIseq, respectively, whereas 277, 195, 115 and 8.936 DEGs were found in cmm6-6. Of all DEGs identified, 29 and 47 genes were in common to all statistical approaches in cmm1T and cmm6-6, respectively, and kept for further analysis. Most of DEGs (28 and 46 in cmm1T and cmm6-6, respectively) were unique in one or the other clone (Fig 5).

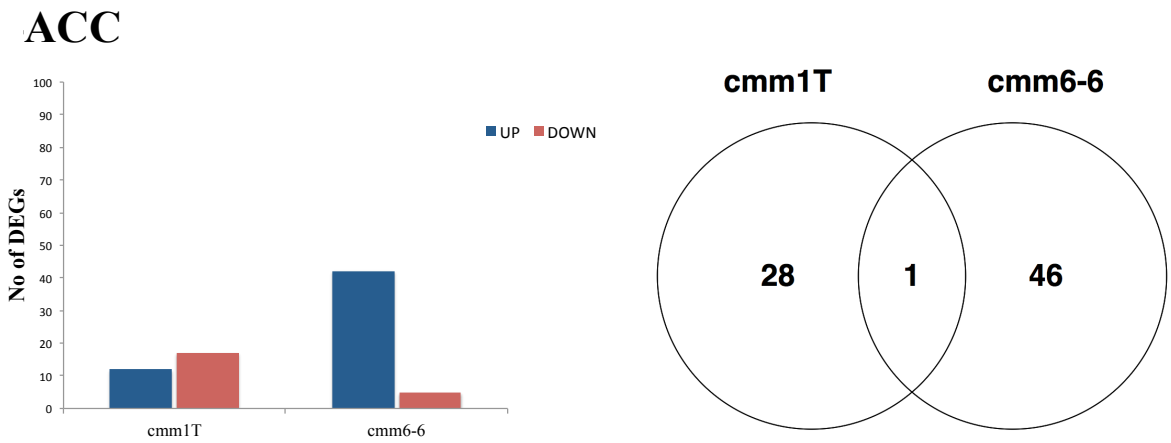


**Fig. 5** Number of total differentially expressed genes (DEGs) in cmm1T and cmm6-6 under ACC conditions and those that were in common or uniquely expressed.

Out of 29 DEGs in cmm1T, 12 (41 %) were up-regulated and 17 (59 %) were down-regulated (Fig 5). By contrast, as already found following NACC stress, in cmm6-6 most of DEGs (42/47, 89 %) were up-regulated, while only 5 (11 %) were down-regulated (Fig 5B). Orthology analysis showed that, out of 29 DEGs founded in cmm1T, 24 had an ortholog in *S. tuberosum* (Tab. 2B). Among the DEGs found in cmm1T, four were involved in secondary metabolism (bin 16). Two of them were involved in phenypropanoid and lignin pathways (PGSC0003DMP400047121 and PGSC0003DMP400037349, PAL1) and two in flavonoid synthesis (PGSC0003DMP400051588, N-idroxycinnamoil/benzoil transferase). Regarding the other DEGs, four are transcription factors (bin 27): two belong to ap2/ERF (*CBF3* and *CBF4*), one to bHLH and one to HB; three are involved in ethylene synthesis (bin 17) and two play a role in the ion transports (bin 34). As far as cmm6-6 is concerned, out of 47 DEGs, 42 had an ortholog in *S. tuberosum.* Most of them (38%) were assigned to unknown function (bin35) (Tab. 3). The other DEGs were: seven transcription factors (bin 27), six genes involved in post-translational modifications (bin 29), three involved in signaling (bin 30) and one relative to stress response (bin 20). Results from qPCR analyses evidenced that expression trends of genes analyzed were comparable to the ones obtained by RNAseq analysis, thus validating the sequencing experiment.

**Tab. 3** MapMan overview in ACC experiments. The number of differentially expressed genes (DEGs) in cmm1T and cmm6-6 corresponding to each MapMan bins is shown. –[a] mean not present.

| | | Number of DEGs | |
|---|---|---|---|
| **Bin** | **Name** | **cmm1T** | **cmm6-6** |
| 2 | major CHO metabolism | 1 | 1 |
| 3 | minor CHO metabolism | 1 | 1 |
| 10 | cell wall | 1 | 2 |
| 11 | lipid metabolism | 2 | - |
| 16 | secondary metabolism | 4 | - |
| 17 | hormone metabolism | 3 | 1 |
| 20 | stress | 1 | 1 |
| 21 | redox | - | 1 |
| 26 | misc | 2 | 2 |
| 27 | RNA regulation | 4 | 7 |
| 29 | protein | 1 | 6 |
| 30 | signalling | 1 | 3 |
| 33 | development | 1 | 1 |
| 34 | transport | 2 | - |
| 35 | not assigned | 1 | 16 |

### 2.3.2 GO classification and enrichment analysis of DEGs

Enrichment GO analysis was performed separately in cmm1T and cmm6-6. Following NACC stress, 179 and 1162 unique GO terms were found by AGRIGO analysis and associated to the significantly up-regulated or down-regulated genes in cmm1T and cmm6-6, respectively. When the FDR filter was applied to the list of GO term after REVIGO analysis to remove GO redundancy, a total of 45 (cmm1T) and 560 (cmm6-6) terms for biological process, 18 (cmm1T) and 100 (cmm6-6) for cellular component and 15 (cmm1T) and 147 (cmm6-6) for molecular functions were

identified (data not shown). For biological process, the majority of genes belong to GO:0008152-biosynthetic process (75% and 76% in cmm1T and cmm6-6, respectively), GO:0009987-cellular process (63% in both), and GO:0065007-biological regulation (7% and 6%), GO:0051179-localisation (17% in both) (Tab. 4). For cellular component, GO functions found were: GO:0031224-intrinsic component of membrane (56% and 55%), GO:0005623-cell (53% and 54%), GO:0016020-membrane (61,5% in both). For molecular function, GO functions found were GO:0003824-catalytic activity (65% and 66%), GO:0005488-binding (55% in both) GO:0001071-ion binding (4.26% and 1.92%) and GO:0005215-transporter activity (1.42% and 2.66%) for cmm1T and cmm6-6, respectively.

**Tab. 4** GO enrichment analysis in clones cmm1T and cmm6-6 of *S. commersonii* under non-acclimated (NACC) and acclimated (ACC) conditions. The percentage of genes belonging to the most represented GO categories in the two clones is reported.

| | | | **% of genes** | |
|---|---|---|---|---|
| | **GO term** | **description** | **cmm1T** | **cmm6-6** |
| NACC | GO:0008152 | biosynthetic process | 75 | 76 |
| | GO:0009987 | cellular process | 63 | 63 |
| | GO:0065007 | biological regulation | 7 | 6 |
| | GO:0051179 | localization | 17 | 17 |
| | GO:0031224 | intrinsic component of membrane | 56 | 55 |
| | GO:0005623 | cell | 53 | 54 |
| | GO:0016020 | membrane | 61.5 | 61.5 |
| | GO:0003824 | catalytic activity | 65 | 66 |
| | GO:0005488 | binding | 55 | 55 |
| | GO:0001071 | ion binding | 4.26 | 1.92 |
| | GO:0005215 | transporter | 1.42 | 2.66 |
| ACC | GO:0008152 | metabolic process | 76 | 76 |
| | GO:0009987 | cellular process | 66 | 65 |
| | GO:0044238 | primary metabolic process | 4 | 3 |
| | GO:0006950 | response to stress | 17 | 17 |
| | GO:0009059 | response to stimulus | 50 | 52 |
| | GO:0044262 | transport | 53 | 54 |
| | GO:0005623 | cell | 60 | 61.5 |

| | | | |
|---|---|---|---|
| GO:0005622 | intracellular | 60 | 60 |
| GO:0005737 | cytoplasm | 52 | 53 |
| GO:0005488 | binding | 4 | 2 |
| GO:0003824 | catalytic activity | 30 | 29 |
| GO:0009058 | biosynthetic process | 4 | 2 |

GO results indicated that cmm1T was less affected by cold than cmm6-6 after 30 min of cold stress in NACC conditions (not shown). In particular, the expression of genes that were enriched in GO categories corresponding to "phenylpropanoid metabolic process" and "protein metabolic process" were detected in cmm1T, whereas the categories "alcohol metabolic process", "cell death ", "ion transport", "DNA binding", "DNA replication", "regulation of photosynthesis", "microtubule" and "cell growth", were significantly enriched only in cmm6-6. Nine additional GO terms (e.g. biological regulation and cellular process) were enriched in both clones.

GO results from ACC experiments provided evidence that GO terms enriched in each clone were comparable and were generally related to "metabolic process". For instance, out of all GO terms, a total of 16 (in cmm1T) and 10 (in cmm6-6) for biological process, 5 (in cmm1T) and 10 (in cmm6-6) terms for cellular component and 6 (in cmm1T) and 7 (in cmm6-6) for molecular functions were identified (data not shown). The analysis of GO terms for cold-regulated genes suggested that the categories "cellular process" (66% and 65% in cmm1T and cmm6-6, respectively), "biosynthetic process" (4% and 2%), "primary metabolic process" (4% and 3%), "response to stress" (17% in both) and "response to stimuli" (50% and 52%), transport (53% and 54%) were enriched in both cmm1T and cmm6-6 (Tab 4). In the case of the GO category "phenylpropanoid metabolic process", "aromatic compound biosynthetic process" and "cellular amino acid metabolic process", significant enrichment was observed only in cmm1T (not shown), whereas the GO category "membrane-bounded organelle", "transition metal ion binding", "hydrolase activity" and "organelle" were significant enriched in cmm6-6. Twelve additional GO categories were enriched in both cmm1T and cmm6-6.

**2.4. Discussion**

Here we described for the first time the transcriptomic response of two clones of *S. commersonii* contrasting in their cold tolerance but able to cold acclimate. Studies by Stone et al. (1993) provided evidence that cold tolerance and acclimation capacity are indipendent traits with a polygenic control. Therefore, here we treated the two traits separately. As for non-acclimated cold tolerance, a

strategy focusing on differential expressed genes that were unique in tolerant cmm1T was used, whereas those that were unique in susceptible cmm6-6 were not fully discussed. Data showed that waves of changes in the composition of the transcriptome occurred within 30 min of NACC stress in both clones. However, tolerant clone cmm1T was less affected by cold treatment than susceptible cmm6-6. A similar finding was also reported in rice by da Maia et al. (2017) following RNAseq analysis, where the authors found 259 and 5579 DEGs after cold exposure of a cold tolerant and susceptible genotype, respectively. Similar to Liu et al. (2012) and Chen et al. (2015), who revealed that the cold susceptible *S. lycopersicum* showed more severe inhibition of photosynthesis than the cold tolerant *S. habrochaites,* we found that some photosynthesis-related GO terms were significantly enriched among the down-regulated genes in the susceptible clone cmm6-6 under NACC conditions. As far as cold tolerance in acclimated conditions is concerned, both genotypes showed a similar behavior, consistent with their ability to acclimate. Indeed, only few genes were differentially expressed. As already reported in NACC conditions, most of DEGs following ACC stress were genotype-specific, suggesting that different clones of one plant species might vary in their ability to respond to low temperatures stress, although their ability to cold acclimate is similar. In other words, the response of genes to cold stress may be genotype specific within the same plant species, as already shown in tomato and rice (da Maia et al. 2017; Chen et al. 2015). Although most of DEGs were different in cmm1T and cmm6-6 under ACC conditions, their GO terms were similar. Indeed, twelve GO categories comprising “response to stress”, “response to stimuli” and “transport” were enriched in both cmm1T and cmm6-6, respectively.

### *2.4.1 Cold perception and membrane modifications*

It is still unknown how plants sense cold stress. However, some evidences on the role of receptor-like protein kinases (RLKs) in the perception and transmission of an external stimulus through signaling cascades are coming out (Chae et al. 2009; Hwang et al. 2011; Marshall et al. 2012). Elucidation of the functions of these kinases in abiotic stress response will provide a better understanding of stress-sensing mechanisms in plants and help to identify potential candidate genes for breeding. In our study LRR receptors were the unique type of membrane receptors showing different expression following 30 min of exposure to NACC and ACC conditions, suggesting that they might be actively involved in signal perception and signal transduction in both cmm1T and cmm6-6. Similar findings were also shown by da Maia et al. (2017), who described the role of LRR receptor (Os02g0647300) in cold stress tolerance of rice. Likewise, Yang et al. (2014) in *Glycine soya* demonstrated that the overexpression of a specific *GsLRR* in yeast and *Arabidopsis* enhances the resistance to cold stress and increases the expression of a number of cold responsive gene

markers. Interesting features associated with cell wall and plasmatic membrane and their modifications were identified to be specific in tolerant cmm1T under NACC conditions. For example, a laccase (similar to *AtLAC4*, AT2G38080), which is involved in lignin biosynthesis, was suppressed in cmm1T. Ji et al. (2015) reported that reduction in lignin deposition in cell walls not only increases its permeability but also enhances its elasticity. These features may allow cell wall to withstand growing ice crystals, thus reducing damages. Consistent with the theory that lignin deposition is reduced under cold stress, we found also a peroxidase (similar to *AtPRX72*), which was recently identified by Herrero et al. (2013) as player in lignin biosynthesis, suppressed in cmm1T. By contrast, an Aspartic protease located in endomembrane system was induced in cmm1T. Yao et al. (2012) showed that in *A. thaliana* mutant lines, ectopically overexpressing the Aspartic protease *(ASPG1-OE*), water loss was dramatically reduced. This probably leads a minor content of free water that might be transformed in ice, which in turn might break the plasmatic membrane. Similar to the Aspartic protease, the acyl carrier protein *(ACP),* an essential cofactor carrying acyl chains of different lengths, participating in the cycles of condensation, reduction, and dehydration steps, was induced only in tolerant cmm1T. Evidence from *in vitro* and *in vivo* studies indicated that *ACP* isoforms are specific for enzymes involved in fatty acid biosynthesis, suggesting that they act changing the fatty acid composition of membrane and leaf (Guerra et al. 1986, Schütt et al. 1998, Suh et al. 1999, Branen et al. 2001, 2003). The link between *ACP* and cold tolerance was given by Tang and colleagues in 2012. For instance, the authors revealed that transgenic tobacco lines (OE-*AhACP1* or AT-*AhACP1*) overexpressing *ACP* gene, showed enhanced cold tolerance, concluding that this gene play an important role under low temperatures. Hence, molecular mechanisms involved in membrane and cell wall modification, lipid composition and lignin deposition might be essential for the cold tolerance of cmm1T. Although these modifications might be crucial for cmm1T surveillance, they may not be the unique mechanisms affected under NACC conditions in this clone. For example, a basic chitinase (maker_scaffold23738_augustus_gene_0_36) involved in resistance processes, was induced. Plant chitinases are the members of PR (Pathogenesis related) proteins family that protect plants from environmental stresses. It is known that following cold stress, plants secrete proteins in the extracellular space, mainly anti freezing and PR proteins, which represent the first line of defense (Nakamura et al. 2008). A recent work by Kashyap et al. (2017) also identified a novel chitinase (*HrCHI1*) harboring bHLH proteins and DRE elements in the promoter. The authors demonstrated its role as cold responsive gene, acting via CBF/ERF dependent cold signaling pathway.

#### *2.4.1.1 Into the nucleus: CBF-pathway*

To understand the molecular basis underlying why cmm1T can survive freezing temperatures and acclimate, whereas cmm6-6 can only acclimate, the *CBFs* pathway was analyzed. Genes associated to this molecular processes were found in the gene sets in both clones and under both NACC and ACC conditions. Our data showed that *CBF3*, *CBF4* and *ZAT12* were up-regulated compared to controls under NACC in tolerant cmm1T, but they were down-regulated in sensitive cmm6-6. By contrast, all the three genes were up-regulated under ACC in both clones. Our findings are in contrast with previous reports that revealed how *CBF1*, but not related *CBFs*, were responsive to low temperatures in both *S. commersonii* and *S. tuberosum* (Pennycooke et al. 2008; Carvallo et al. 2011). However, they are consistent to the results shown by Aversano et al. (2015), who found an induction of all *CBF* genes under non-acclimated and acclimated conditions following a microarray experiment. Our data are also consistent to those observed in tomato species, where three *CBF* genes were cold responsive in the wild cold tolerant *S. peruvianum* (Mboup et al. 2012). The different expression of *CBF* genes in our plant material may be directly responsible for enhanced cold tolerance and acclimation ability in *S. commersonii*. Therefore, probably *CBF3* and *CBF4* are the first members that play an important role in the early cold response before *CBF1* and *CBF2*. *ZAT12* is less specific, as it responds to a large number of biotic and abiotic stresses (Davletova et al. 2005). However, it has been shown that constitutive expression of *ZAT12* in *Arabidopsis* caused a small, but reproducible, increase in freezing tolerance (Vogel et al. 2005). Hence, it is another good candidate for further molecular characterization together with *CBF3* and *CBF4.*

#### *2.4.1.2 Others potential candidates*

Cold tolerance and cold acclimation in potato, as well as in other species such as rice, cabbage, wheat, and tea is based on quantitative inheritance. There are many genes and interactions, possibly epistatic mechanisms, involved in defining the different tolerance levels among genotypes. Thus, one could expect that different metabolic routes and different genes are modulated in response to cold. Transcription factors including HB, bZIP, MYB, MYC, WRKY and bHLH also play important roles in cold as well as under other abiotic stresses. For instance, previous studies showed that the overexpression of *MlNAC5* (*Miscanthus lutarioriparius*) and *SlNAC1* (*Suaeda liaotungensis*) enhanced drought and cold stress tolerance of *Arabidopsis*, respectively (Li et al. 2013; Zong et al. 2016). In our study, more than 20 different TF families were identified to respond to both NACC and ACC stress conditions in both clones of *S. commersonii*. Among them, AP2/ERF, bHLH and Zn-finger TFs genes were the most abundant. These TF families interact to regulate target genes, and several R2R3-type MYB and bHLH TFs have been reported to be involved in plant stress responses. For example, D'Amelia et al. (2017) showed how the expression

of *AN2* (a MYB transcription factor) is induced by low temperatures in wild, cold-tolerant *S. commersonii* but not in susceptible *S. tuberosum* varieties. The authors found that *AN2* is a paralog of the potato anthocyanin regulator *AN1*, showing similar interaction ability with bHLH co-partners. It is notable that some previously uncharacterized TFs are also significantly up- or down-regulated in response to cold in both cmm1T and cmm6-6, suggesting that more studies on the interactions between different TFs families under cold stresses are needed the better understand the molecular dynamics under this harmful stress.

## 2.5. Conclusions

*S. commersonii* is, among wild potatoes, the one displaying the highest tolerance to low temperatures. Our RNAseq data revealed that in NACC conditions cold induced more changes in the sensitive clone than in the tolerant one, where we identified genes mainly involved in cell wall and membrane modifications. Genes encoding proteins with a role in lignin biosynthesis, lipid composition and production of anti freezing molecules such as PR proteins might be used for further elucidation through genome editing approaches and, in the end, they might be used in breeding programs to improve the non acclimated cold tolerance of the cultivated potato. RNAseq data from ACC experiments revealed that a reduced number of genes were affected. Interestingly, we found that the function of these genes was in common between cmm1T and cmm6-6. Our study also showed that some cold responsive transcription factors genes such as CBF3, CBF4 and ZAT12 had an opposite trend (increased and decreased expression levels in the tolerant cmm1T and sensitive cmm6-6 under NACC conditions; by contrast, they were induced in both clones when acclimated). These findings demonstrate the important role of *CBF-dependent* pathway under both not acclimated and acclimated conditions. Further studies at promoter level of *CBF3*, *CBF4* and *ZAT12* of cmm1T and cmm6-6 might elucidate the differences between the two clones in term of cold tolerance. All these findings increase our understanding on the molecular basis of cold response under non- and acclimated conditions and provide a list of gene targets for marker assisted selection approaches aiming at introgressing cold tolerance in *S. tuberosum*.

**References**

Aversano R, Contaldi F, Ercolano MR, et al (2015) The *Solanum commersonii* genome sequence provides insights into adaptation to stress conditions and genome evolution of wild potato relatives. Plant Cell 27:954-968.

Badawi M, Danyluk J, Boucho B, Houde M and Sarhan F (2007). The *CBF* gene family in hexaploid wheat and its relationship to the phylogenetic complexity of cereal *CBFs*. Mol Genet Genom 277:533–554.

Bailey-Serres J, Lee SC, Brinton E (2012) Waterproofing crops: effective flooding survival strategies. Plant Physiol 160:1698–1709.

Benedict C, Skinner JS, Meng R, Chang Y, Bhalerao R, Huner NP, et al. (2006) The *CBF1*-dependent low temperature signalling pathway, regulon and increase in freeze tolerance are conserved in *Populus* spp. Plant Cell Environ 29:1259–1272

Boyer JS et al. (2013). The U.S. drought of 2012 in perspective: a call to action. Global Food Secur 2:139–143.

Carvallo MA, Pino MT, Jeknić Z, Zou C, Doherty CJ, Shiu SH, and Thomashow MF (2011). A comparison of the low temperature transcriptomes and CBF regulons of three plant species that differ in freezing tolerance: *Solanum commersonii*, *Solanum tuberosum*, and *Arabidopsis thaliana*. Journal of Experimental Botany 62:3807–3819.

Chae L, Sudat S, Dudoit S, Zhu T, Luan S (2009). Diverse transcriptional programs associated with environmental stress and hormones in the *Arabidopsis* receptor-like kinase gene family. Molecular Plant 2:84–107

Chen H, Chen X, Chen D, Li J, Zhang Y, and Wang A (2015). A comparison of the low temperature transcriptomes of two tomato genotypes that differ in freezing tolerance: *Solanum lycopersicum* and *Solanum habrochaites*. BMC Plant Biol 15:132.

Chen HH and Li PH (1980) Biochemical changes in tuber-bearing *Solanum* species in relation to frost hardiness during cold acclimation. Plant Physiol 66:414-421.

Chen X, Gu Z, Xin D, Hao L, Liu C, Huang J, et al. (2011). Identification and characterization of putative *CIPK* genes in maize. J Genet Genomics 38:77–87.

Chinnusamy V and Zhu JK (2009) Epigenetic regulation of stress responses in plants. Curr. Opin. Plant Biol 12:133–39.

Chinnusamy V, Zhu J, Zhu JK (2007) Cold stress regulation of gene expression in plants. Trends Plant Sci 12:444–451.

Cumbie JS, Kimbrel JA, Di YM, Schafer DW, Wilhelm LJ, Fox SE, et al. (2011). GENE-counter: a computational pipeline for the analysis of RNA-Seq data for gene expression differences. PLos One 6:e25279.

D'Amelia V, Aversano R, Ruggiero A, Batelli G, Appelhagen I, Dinacci C, Hill L, Martin C, Carputo D (2017) Sub-functionalization of duplicate *MYB* genes in *Solanum commersonii*

generated the cold-induced *ScAN2* and the anthocyanin regulator *ScAN1*. Plant Cell Environ, in press

da Maia LC, Cadore PRB, Benitez LC, Danielowski R, Braga EJB, Fagundes PRR and Costa de Oliveira A (2017). Transcriptome profiling of rice seedlings under cold stress. Functional Plant Biology 44:419.

Dahal K, Kane K, Gadapati W, Webb E, Savitch LV, Singh J, Sharma P, Sarhan F, Longstaffe FJ, Grodzinski B, Huner NPA (2012) The effects of phenotypic plasticity on photosynthetic performance in winter rye, winter wheat, and Dhanaraj AL, Alkharouf NW, Beard HS, Chouikha IB, Brassica napus . Physiol Plant 144:169–188

Dametto A, Sperotto RA, Adamski JM, Blasi ÉAR, Cargnelutti D, de Oliveira LFV, and Fett JP (2015). Cold tolerance in rice germinating seeds revealed by deep RNAseq analysis of contrasting indica genotypes. Plant Science, 238:1–12.
Dong XS, Im SB, Lim YP, Nou IS, Hur Y (2014). Comparative transcriptome profiling of freezing stress responsiveness in two contrasting Chinese cabbage genotypes, Chiifu and Kenshin. Genes Genom 36:215–27.

Goulas E, Schubert M, Kieselbach T, Kleczkowski LA, Gardestrom P, et al. (2006). The chloroplast lumen and stromal proteomes of *Arabidopsis thaliana* show differential sensitivity to short- and long-term exposure to low temperature. Plant J 47:720–34

Hwang SG, Kim DS, Jang CS (2011) Comparative analysis of evolutionary dynamics of genes encoding leucine-rich repeat receptor-like kinase between rice and *Arabidopsis*. Genetica 139:1023–1032.

Jaglo KR, Kleff S, Amundsen KL, Zhang X, Haake V, Zhang JZ, et al. (2001) Components of the *Arabidopsis* C-repeat/dehydration-responsive element binding factor cold-response pathway are conserved in *Brassica napus* and other plant species. Plant Physiol 127:910–7

Janskà, Marsik P, Zelenkov´ a S, Ovesn´ a J (2010) Cold stress and acclimation – what is important for metabolic adjustment? Plant Biol 12:395–405

Kanwar P, Sanyal SK, Tokas I, Yadav AK, Pandey A, Kapoor S, et al. (2014) Comprehensive structural, interaction and expression analysis of *CBL* and *CIPK* complement during abiotic stresses and development in rice. Cell Calcium 56:81–95.

Kolukisaoglu U,Weinl S,Blazevic D, Batistic O, Kudla J (2004) .Calcium sensors and their interacting protein kinases: Genomics of the *Arabidopsis* and rice *CBL-CIPK* signaling networks. Plant Physiol 134:43–58.

Kovalchuk O, Burke P, Arkhipov A, Kuchma N, James SJ, Kovalchuk I, Pogribny I (2003b) Genome hypermethylation in *Pinus silvestris* of Chernobyl—A mechanism for radiation adaptation? Mutat Res 529:13–20.

Leng N, Dawson JA, Thomson JA, Ruotti V, Rissman AI, Smits BMG, Haag JD, Gould MN, Stewart RM, Kendziorski C (2013) EBSeq: an empirical Bayes hierarchical model for inference in RNA-seq experiments . Bioinformatics 29:1035-1043.

Li J, Wang N, Xin H, and Li S (2013). Overexpression of VaCBF4, a transcription factor from *Vitis amurensis*, improves cold tolerance accompanying increased resistance to drought and salinity in *Arabidopsis*. Plant Mol Biol Report 31:1518–1528.

Lissarre M, Ohta M, Sato A, and Miura K (2010). Cold-responsive gene regulation during cold acclimation in plants. Plant Signal Beh 5:948–952.

Liu Q, Kasuga M, Sakuma Y, Abe H, Miura S, Yamaguchi-Shinozaki K, et al. (1998) Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low-temperature-responsive gene expression, respectively, in *Arabidopsis*. Plant Cell 10:1391–406.

Liu H, Ouyang B, Zhang J, Wang T, Li H, Zhang Y, … Ye Z (2012). Differential modulation of photosynthesis, signaling, and transcriptional regulation between tolerant and sensitive Tomato genotypes under cold Stress. PLoS ONE

Livak KJ and Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. Methods 25:402-408.

Love MI, et al. (2014) Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. bioRxiv, doi:10.1101/002832

Lu X, Zhou X, Cao Y, Zhou M, McNeil D, Liang S, & Yang C (2017) RNA-seq Analysis of Cold and Drought Responsive Transcriptomes of *Zea mays* ssp. mexicana L. Frontiers in Plant Science 8:1–15. doi:10.3389/fpls.2017.00136

Marshall A, Aalen RB, Audenaert D, et al. (2012). Tackling drought stress: RECEPTOR-LIKE KINASES present new approaches. The Plant Cell 24:2262–2278.

Mayer BF, Ali-Benali MA, Demone J, Bertrand A, Charron JB **(**2015) Cold acclimation induces distinctive changes in the chromatin state and transcript levels of *COR* genes in *Cannabis sativa* varieties with contrasting cold acclimation capacities. Physiologia Plantarum 155**:**281–295.

Mboup M, Fischer I, Lainer H, Stephan W (2012) Trans-species polymorphism and allele-specific expression in the *CBF* gene family of wild tomatoes. Mol. Biol. Evol. 29: 3641–3652.

Pennycooke JC, Cheng HM, Roberts SM, Yang QF, Rhee SY, Stockinger EJ (2008) The low temperature-responsive, Solanum *CBF1* genes maintain high identity in their upstream regions in a genomic environment undergoing gene duplications, deletions, and rearrangements. Plant Mol Biol 67: 483–497

Pino MT, Skinner JS, Park EJ, Jeknić Z, Hayes PM, Thomashow MF, et al. (2007) Use of a stress inducible promoter to drive ectopic *AtCBF* expression improves potato freezing tolerance while minimizing negative effects on tuber yield. Plant Biotechnol J 5:591–604.

Rapaport F, Khanin R, Liang YP, Pirun M, Krek A, Zumbo P, et al. (2013) Comprehensive evaluation of differential gene expression analysis methods for RNA-seq data. Genome Biol 14:R95

Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics., 26:139-140.

Shan X, Wang X, Yang G, Wu Y, Su S, Li S, … Yuan Y (2013) Analysis of the DNA methylation of maize (*Zea mays L.*) in response to cold stress based on methylation-sensitive amplified polymorphisms. Journal of Plant Biology 56:32–38.

Shen C, Li D, He R, Fang Z, Xia Y, Gao J, … Cao M (2014) Comparative transcriptome analysis of RNA-seq data for cold-tolerant and cold-sensitive rice genotypes under cold stress. Journal of Plant Biology, 57:337–348.

Shinozaki K, Yamaguchi-Shinozaki K, Seki M (2003) Regulatory network of gene expression in the drought and cold stress responses. Curr Opin Plant Biol 6:410–417.

Skinner JS et al. (2005) Structural, functional, and phylogenetic characterization of a large *CBF* gene family in barley. Plant Mol Biol 59:533–551.

Soneson C, Delorenzi M (2013) A comparison of methods for differential expression analysis of RNA-seq data. BMC Bioinformatics 14:91.

Song Y, Liu L, Feng Y, Wei Y, Yue X, He W, Zhang H, An L. **(**2015). Chilling- and freezing-induced alterations in cytosine methylation and its association with the cold tolerance of an alpine subnival plant, *Chorispora bungeana*. PLoS ONE 10, e0135485.

Stitt M, Hurry V (2002) A plant for all season: alternations in photosynthetic carbon metabolism during cold acclimation in *Arabidopsis*. Curr Opin Plant Biol 5:199–206

Stone JM, Palta JP, Bamberg JB, Weiss LS, Harbage JF (1993) Inheritance of freezing resistance in tuber-bearing *Solanum* species: Evidence for independent genetic control of nonacclimated freezing tolerance and cold acclimation capacity. Proc Natl Acad Sci USA 90:7869-7873

Tarazona S, Furió-Tarı P, Ferrer A, Conesa A (2013): NOISeq: Differential Expression in RNA-seq.

Thimm O, Bla¨sing O, Gibon Y, Nagel A, Meyer S, Kru¨ger P, Selbig J, Muller LA, Rhee SY, Stitt M (2004) MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. Plant J 37:914–939

Thomashow MF, Gilmour SJ, Stockinger EJ, Jaglo-Ottosen K, Zarka DG (2001) Role of the *Arabidopsis CBF* transcriptional activators in cold acclimation. Physiol Plant 112:171–175

Van Buskirk HA, & Thomashow MF (2006) *Arabidopsis* transcription factors regulating cold acclimation. Physiologia Plantarum, 126(1), 72-80. DOI: 10.1111/j.1399-3054.2006.00625.x

Vogel JT, Zarka DG, Van Buskirk HA, Fowler SG, Thomashow MF (2005) Roles of the *CBF2* and *ZAT12* transcription factors in configuring the low temperature transcriptome of *Arabidopsis*. Plant J 41: 95–211

Yang L, Wu K, Gao P, Liu X, Li G, Wu Z (2014) GsLRPK, a novel cold-activated leucine-rich

repeat receptor-like protein kinase from *Glycine soja*, is a positive regulator to cold stress tolerance. Plant Science 215–216, 19–28.

Yang T, Chaudhuri S, Yang L, Du L, Poovaiah BW (2010) A calcium/calmodulin-regulated member of the receptor-like kinase family confers cold tolerance in plants. Journal of Biological Chemistry 285:7119–7126.

Yang T, Shad Ali G, Yang L, Du L, Reddy AS, Poovaiah BW (2010b) Calcium/calmodulin-regulated receptor-like kinase CRLK1 interacts with MEKK1 in plants. Plant Signaling Behavior 5:991–994.

Yu Y, Xia X, Yin W, Zhang H (2007) Comparative genomic analysis of *CIPK* gene family in *Arabidopsis* and *Populus*. Plant Growth Regul. 52:101–110. 10.1007/s10725-007-9165-3

Yun KY, Park MR, Mohanty B, Herath V, Xu F, Mauleon R, et al (2010) Transcriptional regulatory network triggered by oxidative signals configures the early response mechanisms of japonica rice to chilling stress. BMC Plant Biol 10:16.

Ramsak Z, Baebler S, Rotter A, Korbar M, Mozetic I, Usadel B, Gruden K GoMapMan: integration, consolidation and visualization of plant gene annotations within the MapMan ontology

Zhang H, Yang B, Liu WZ, Li H, Wang L, Wang B, et al (2014) Identification and characterization of *CBL* and *CIPK* gene families in canola (*Brassica napus* L.). BMC Plant Biol 14:8. 10.1186/1471-2229-14-8

Zhang X, Fowler SG, Cheng H, Lou Y, Rhee SY, Stockinger EJ (2004) Freezing- sensitive tomato has a functional *CBF* cold response pathway, but a *CBF* regulon that differs from that of freezing-tolerant *Arabidopsis*. Plant J.39:905–19

Zong JM, Li XW, Zhou YH, Wang FW, Wang N, Dong YY et al. (2016) The *AtDREB1* transcription factor fromthe cold-tolerant plant *Adonis amurensis* enhances abiotic stress tolerance in Transgenic Plant. Int. J Mol Sci 17:611–626.

# Chapter III. Deep-sequencing of *Solanum commersonii* smallRNA libraries reveals riboregulators involved in cold stress response

**Deep-sequencing of *Solanum commersonii* smallRNA libraries reveals riboregulators involved in cold stress response**

**Abstract**

Low temperatures represent one of the most important harmful factors that limit crop production. Although the molecular mechanisms on how different plant species respond to cold stress remains to be elucidated, recent studies have shown that abiotic stresses induce aberrant expression of many small non-coding RNA (sncRNA). However, riboregulators dynamics under low temperatures remain largely unknown. Among the wild species used in potato breeding, *Solanum commersonii* is the one displaying the highest tolerance to low temperatures. It is also the first potato relative whose genome sequence has been deciphered. With the aim to understand how its sncRNAs are affected by cold stress, two clones of *S. commersonii* contrasting in their cold response were used in this study, and their sncRNAome has been analyzed through RNAseq strategy. A prevalence of 21- and 24-nt sncRNAs divided in three classes of mature miRNAs (273), tasiRNA (5.737) and other smallRNA (134.868) were annotated in the genome of *S. commersonii*. Among the miRNAs, 44 were conserved with high similarity with other plant species, and 229 were new or *S. commersonii*-specific (not reported in any database). Targets were determined by in silico prediction and several genes encoding transcription factors were identified as putative players in cold stress response. Among them, *WRKY*, *MYB* and *GRAS* were the most abundant. Differential analysis provided evidence that several miRNAs change during stress conditions, and that they are negatively correlated with their targets. We also showed that tasiRNA as well as secondary siRNA changed their expression under both stresses, leading to the idea that they play an active role under cold stress. Our results reveal possible roles of sncRNA in the regulatory networks associated with tolerance to low temperatures and provide useful information for a more strategic use of genomic resources in potato breeding efforts.

## 3.1. Introduction

The discovery of small non-coding RNAs (sncRNAs) and their widespread roles in both transcriptional and post-transcriptional gene regulation has changed our basic understanding of how genes are regulated. Importantly, there is increasing evidence indicating that sncRNAs play vital roles in stress response and during various developmental processes, including shoot apical meristem formation, leaf morphogenesis and polarity, floral organ identity, root development, vegetative phase change and vascular development (Kamthan et al. 2015; Nova-Franco et al. 2015; Damodharan et al. 2016; Li and Zhang 2016). Generally, sncRNAs are distinguished in hairpin

RNAs (hpRNAs), derived from single-stranded hairpins and small interfering RNAs (siRNAs) derived from double-stranded RNA (dsRNA). Although they are functionally similar, differences in structure and mode of biogenesis have been documented (Carthew et al. 2009). HpRNAs include microRNAs (miRNAs), a class of riboregulators that is very well known in plants. They are short (21–24nt) in length, single stranded, non-coding and produced from an RNA Polymerase II transcript with a strong secondary stem-loop structure (Bologna and Voinnet 2014). In plants, Dicer-like proteins (DCL1 and DCL3) cleave this secondary structure and produce an hairpin RNA molecule (pre-miRNA), that is successively handled by DCL1, resulting in a double stranded intermediate RNA. One strand of this RNA (guide strand) is incorporated into the effector complex (RISC, RNA Induced Silencing Complex), while the other strand is known as the 'star' sequence (miRNA*) or passenger strand and is usually degradated. Several miRNAs are evolutionary conserved among diverse plant species, while others that have recently evolved show species-specificity and are often expressed at lower levels compared to conserved miRNAs (Fahlgren et al. 2007, Ren et al. 2012). Due to their low expression levels, most of species-specific miRNAs remained unidentified. In recent years, with the advent of high-throughput sequencing technologies, both species-specific and conserved miRNAs have been identified in diverse plant species (miRBase, v. 21; Kozomara & Griffiths-Jones 2014, Megha et al. 2017). The other group of sncRNAs includes endogenous siRNA, arising from long dsRNA precursors. These come from the hybridization of sense and antisense transcripts, from the folding back of an inverted‐repeat sequence or from the hybridization of unrelated RNA molecules with strong sequence complementarity. Endogenous siRNA includes heterochromatic (het)siRNAs and secondary siRNAs (Borges et al. 2015). The former are the most abundant in plants and mediate transcriptional silencing of transposons and pericentromeric repeats. They require transcription by polymerase IV, followed by dsRNA synthesis through RNA-dependent-RNA polymerase-II (RDR2) and processed by DCL3. Secondary siRNAs, such as *trans*-acting RNA (tasiRNAs), phasedRNA (phasiRNAs) and epigenetically activated RNA (easiRNAs) are produced by DCL4 and DCL2, respectively, following Polymerase II transcription and dsRNA synthesis by RDR6.

Among the principal crops, the cultivated potato (*Solanum tuberosum*) is third in terms of production after rice and wheat. The main reasons for its increasing popularity are related to the high nutritional value of tubers combined with the simplicity of its reproduction by vegetative propagation. Unfortunately, it is susceptible to a wide range of abiotic and biotic stresses. For this reason, breeders are continuously seeking sources of genetic resistance to main stressors. A provider of allelic/gene diversity is represented by wild tuber-bearing Solanums. There are about 200 potato species distributed from the southern part of United States of America up to southern

Chile. They make the potato unrivalled within cultivated crops in terms of related species. Among them, *S. commersonii* has attracted much interest in recent years. Particularly important are its resistance to low temperatures and capacity to cold acclimate (i.e. to increase its resistance upon exposure to low but non-killing temperatures). Therefore, *S. commersonii* is a candidate species to study and clarify in details the genetic dynamics underlying these noteworthy traits. It is reported in the literature that miRNA plays an important role during cold stress. Indeed, differential profiling of low temperatures-induced miRNAs using next generation sequencing platforms has been reported in various plant species, including *Arabidopsis* (Liu et al. 2008), *Populus trichocarpa* (Zhang et al. 2009b; Chen et al. 2012), rice (Lv et al. 2010), *Hemerocallis fulva* (An et al. 2014), tomato (Cao et al. 2014), grapevine (Sun et al. 2015) and almond (Karimi et al. 2016). Knowledge about species-specific miRNA and secondary siRNA population operating in consequence to cold stress in potatoes remains largely scant. Taking advantage of the recently published genome sequence of *S. commersonii* (Aversano et al. 2015), here we used Illumina sequencing technology and bioinformatics tools to identify miRNAs, tasiRNA and other siRNAs that could be involved in the molecular mechanisms of cold resistance of this potato species. We characterized sncRNAome and identified and studied in details conserved and novel miRNAs involved in cold stress tolerance. To gain a better understanding of miRNA functions in various processes, their targets were also analyzed through GO analysis. This effort provides for the first time a comprehensive analysis of smRNA population in wild potatoes, and gives us the opportunity for further studies correlating miRNA expression with RNAseq data of cold stressed plants.

## 3.2. Material and methods

### 3.2.1 Cold stress assay

Young plants of two different clones of *S. commersonii* were used in our study: clone cmm1T of PI243503 is frost-resistant and able to cold acclimate, whereas clone cmm6-6 of PI590886 is cold-sensitive but able to cold acclimate. Plants were micro-propagated *in vitro* on Murashige and Skoog (MS) medium (Sigma-Aldrich, http://www.sigmaaldrich.com) with 1% (w/v) sucrose and 0.8% (w/v) agar, incubated at 24 ° C with irradiance of 200 $\mu$ mol m-2 s-1, under a 16/8 h (light/dark) photoperiod. Four-weeks old vitroplants were transplanted into 14-mm pots filled with sterile soil and grown for two weeks at 24° C in growth chamber prior submitting them to non-acclimated (NACC) and acclimated (ACC) stress conditions. During NACC experiments, three plants of each clone were challenged for 30min at -2° C, while three plants were chosen as controls and kept at 24° C. In the ACC experiment, 6 plants/clone were acclimated at 4° C for 2 weeks. Then, three of them were transferred for 30min at -2° C, while the others were used as control and kept at 4° C.

In both experiments, an environmentally controlled cold room was used. Young leaf samples were collected from all replicates at the end of each stress and from control plants. Samples were individually stored at -80° C before RNA extraction.

### 3.2.2 MicroRNA sequencing, identification and annotation

For each sample total RNA was isolated from leaf tissue using Spectrum Plant Total RNA kit (Sigma-Aldrich, St. Louis, MO, USA) according to the manufacturer's protocol. RNA concentrations were determined using NanoDrop ND-1000 spectrophotometer (Thermo Scientific, Wilmington, USA) and RNA integrity was verified was checked using bioanalyzer (Agilent Technologies, Santa Clara, California, USA). To identify miRNAs, 24 single-end small RNA libraries from cold stressed and not stressed leaves were sequenced using Illumina Hi-seq2500 (Tab. 1). A total of 10 million reads/sample were obtained. Quality, trimming and clipping were performed using Skewer software (Jiang et al. 2014). All reads obtained after filtering and quality check were pooled and mapped onto the *S. commersonii* genome using bowtie with default parameters (Langmead et al. 2009). The sequences that were mapped to multiple positions in the genome were discarded from the analysis, whereas those that mapped uniquely were further analyzed using miR-prefer. To identify the potential miRNA families, we used Cd-hit (Li et al. 2006) to cluster the predicted potato miRNA candidates. A word length of 7 and a sequence identity threshold of 0.9 were selected to cluster the predicted miRNA sequences. Finally, to identify conserved miRNAs, short-Blastn was used to match the predicted potato miRNA candidates to all the 5940 plant mature miRNAs from mirBase Release 19 (Kozomara et al. 2011). Default parameters were used, except the e.value that was lower than 10e-3. When blast search indicated more than one miRNA family hit, the one showing the least mismatches was selected.

**Tab. 1** List of samples used in smRNAseq experiments. The number of biological replicated in each condition is shown.

| Sample | Species | Condition | Tolerance/susceptibility |
|---|---|---|---|
| 1 | cmm1T | control NACC plant 1 | cold tolerant |
| 2 | cmm1T | control NACC plant 2 | cold tolerant |
| 3 | cmm1T | control NACC plant 3 | cold tolerant |
| 4 | cmm1T | NACC 30 min plant 1 | cold tolerant |
| 5 | cmm1T | NACC 30 min plant 2 | cold tolerant |
| 6 | cmm1T | NACC 30 min plant 3 | cold tolerant |
| 7 | cmm1T | control ACC plant 1 | cold tolerant |
| 8 | cmm1T | control ACC plant 2 | cold tolerant |
| 9 | cmm1T | control ACC plant 3 | cold tolerant |
| 10 | cmm1T | ACC 30 min plant 1 | cold tolerant |
| 11 | cmm1T | ACC 30 min plant 2 | cold tolerant |
| 12 | cmm1T | ACC 30 min plant 3 | cold tolerant |
| 13 | cmm6-6 | control NACC plant 1 | cold susceptible |
| 14 | cmm6-6 | control NACC plant 2 | cold susceptible |
| 15 | cmm6-6 | control NACC plant 3 | cold susceptible |
| 16 | cmm6-6 | NACC 30 min plant 1 | cold susceptible |
| 17 | cmm6-6 | NACC 30 min plant 2 | cold susceptible |
| 18 | cmm6-6 | NACC 30 min plant 3 | cold susceptible |
| 19 | cmm6-6 | control ACC plant 1 | cold tolerant |
| 20 | cmm6-6 | control ACC plant 2 | cold tolerant |
| 21 | cmm6-6 | control ACC plant 3 | cold tolerant |
| 22 | cmm6-6 | ACC 30 min plant 1 | cold tolerant |
| 23 | cmm6-6 | ACC 30 min plant 2 | cold tolerant |
| 24 | cmm6-6 | ACC 30 min plant 3 | cold tolerant |

### 3.2.3 Prediction of potential miRNAs target

psRNATarget program (http://plantgrn.noble.org/psRNATarget/) was employed to predict the targets of the putative miRNAs in the nucleotide sequences of all *S. commersonii* transcripts. The program predicts small RNA targets by reverse complementary matching between small RNA and target transcripts and evaluating the target site accessibility by calculating unpaired energy required

to open secondary structure around the small RNA's target site (Dai et al. 2011). It also reports the translational inhibition or cleavage degradation by presence/absence of a mismatch in the central complementary region of the small RNA sequence. Gene Ontology (http://www.Geneontology.org/) and the Kyoto Encyclopedia of Genes and Genomes (http://www.Genome.jp/kegg/) was also used to further investigate the function of target genes.

### 3.2.4 miRNA annotation and differential analysis

To find cold relative miRNA, each library was mapped onto the *S. commersonii* genome using STAR software (Dobin et al. 2013). Subsequently, only unique reads with quality greater than 30 were used in FeatureCounts to count the high quality mapped reads. After that, only miRNAs that showed at least 10 reads in at least one library were further analyzed. Differential analysis was performed using NOIseq package implemented in R v. 1.0.44.

### 3.2.5 Secondary siRNA and tasiRNA identification and annotation

To annotate the secondary siRNA we used ShortStack (Axtell et al. 2013). The program discriminates the regulatory small RNAs of interest from other small RNAs based on a user-set size range (default: 20–24 nt). Loci where the number of RNAs within the user-set size range meets a minimum threshold (default = 0.8) are annotated as Dicer-derived, and others are annotated as not Dicer-derived. ShortStack annotates and quantifies reference-aligned small RNA-seq data in different steps. During the first step, small RNA clusters are identified in a simple, two-step process. First, islands of significant alignment coverage are identified. Islands are defined as continuous genomic coordinates where the depth of alignment coverage exceeds a user-set minimum threshold. Second, the initial islands are "padded" both upstream and downstream by a user-set number of nucleotides. Regions that overlap after padding are then merged to form clusters.

## 3.3. Results

### 3.3.1 Identification and annotation of conserved and new smRNAs in *S. commersonii*

Fig. 1 shows the overall procedure for the annotation of potential miRNAs and siRNAs. Small RNA sequencing yielded a total of 402.083.098 raw reads. The raw sequences were processed and filtered through several criteria and 240.525.509 of them represented our smallRNA population

with reasonable sequencing depth. They were used to annotate putative miRNAs, tasiRNAs and other siRNAs.

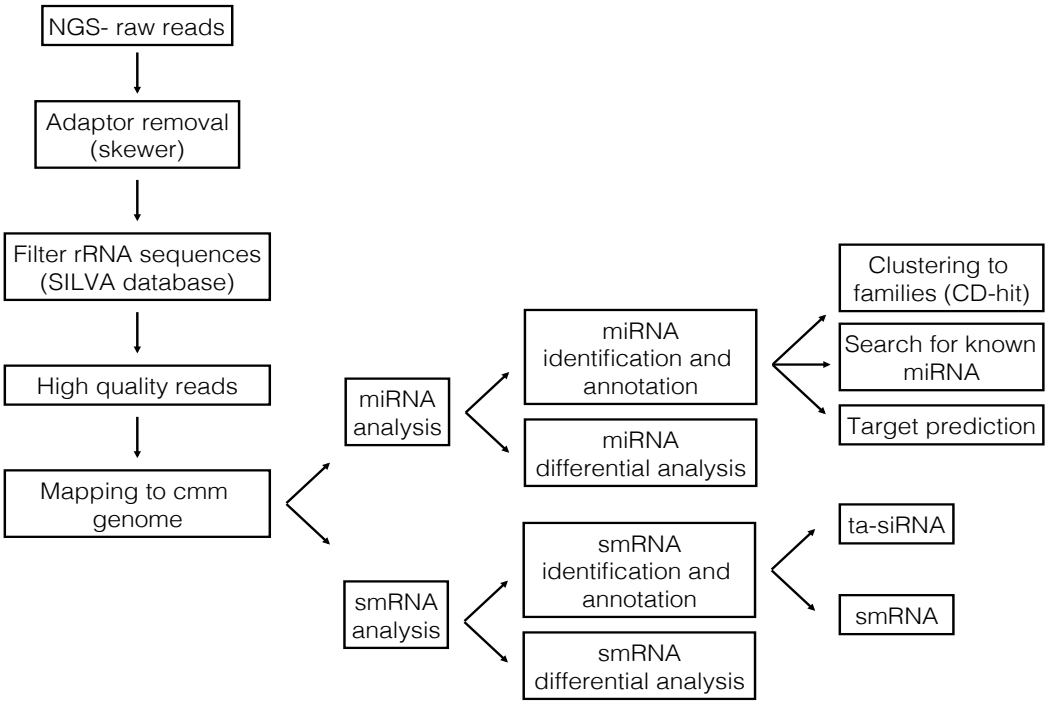


**Fig. 1** Flowchart used for the prediction of miRNAs, their targets, and smRNAs in *S. commersonii*.

A total of 273 miRNAs were identified using miR-Prefer. Among them, 24nt were the most abundant (77%), followed by 21nt (11%) (Fig. 2).

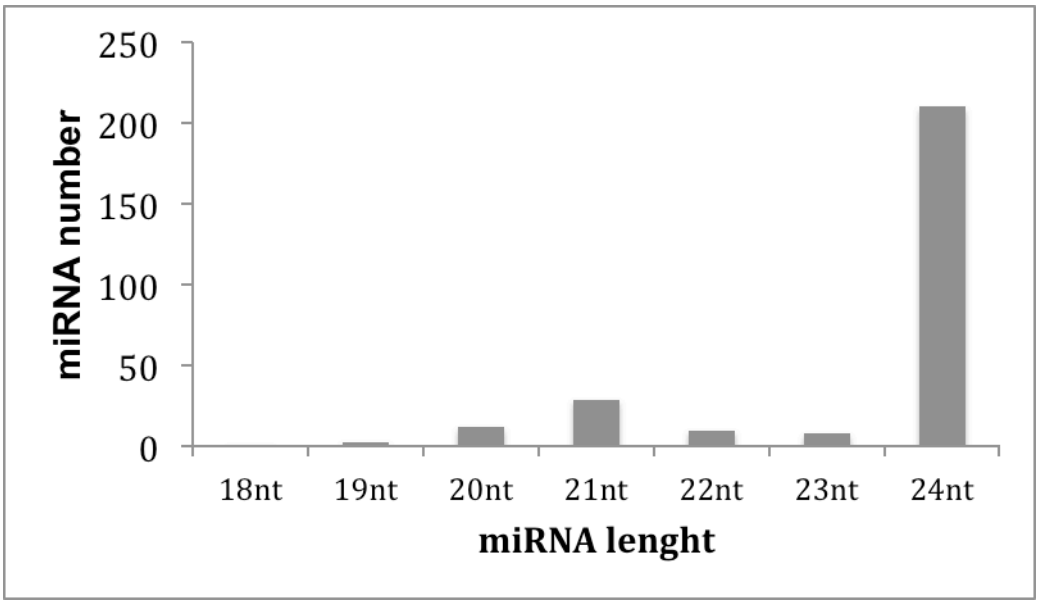


**Fig. 2** Length distribution of miRNA sequences identified in all smRNA libraries.

In order to identify the conserved miRNAs, all the predicted regions were mapped against the known plant miRNAs database and deposited in miRBase v.2.0. Out 273 miRNAs, 44 matched to known plant miRNAs, whereas 229 were considered as novel and *S. commersonii*-specific. Using a family classification made by miRBase v.2.0, conserved miRNAs falling into the same known family were merged, giving a final number of 27 conserved miRNA families (Fig. 3). MiR166 and miR169 had six and five members respectively, miRNA156 had three, while the remaining had one or two family members.

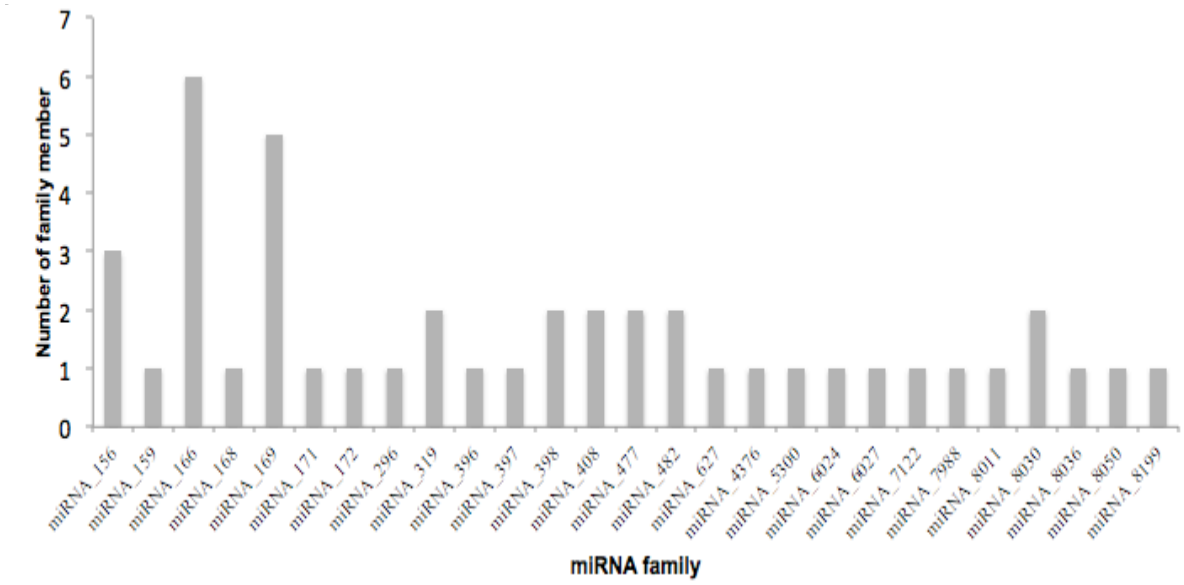


**Fig. 3** Number of different members of conserved miRNA families found in *S. commersonii* by sequencing and bioinformatics prediction.

The predicted miRNAs loci were explored further to find out their distribution among the intergenic and genic regions of the *S. commersonii* genome (Tab. 1). Two hundred and sixty six (97%) of the pre-miRNAs were located in intergenic regions, indicating that a high proportion of miRNAs were encoded by non-annotated genes or non-coding RNAs sequences. Only seven (3%) miRNA were located in introns. As far as the secondary siRNA is concerned, out of 5.737 annotated tasiRNA, 624 (11%) were located in intergenic regions, 743 (13%) in introns and 4.380 (76%) overlapped with transponsable element (TE) regions. Finally, for those smRNA different from miRNAs and tasiRNA, 44.340 were located in intergenic region (32%), 6.747 in introns (5%) and 75.600 (57%) overlapped with TEs (Tab. 2). The remaining 8.181 (6%) were located within genic regions.

**Tab. 2** Number and genomic distribution of predicted miRNAs,tasiRNAs and siRNA.

| | | Localization | | |
|---|---|---|---|---|
| | annotated | intergenic region | genic region | overlapped with TEs |
| miRNA | 273 | 266 | 7 | - |
| tasiRNA | 5.737 | 624 | 743 | 4.380 |
| siRNA | 134.868 | 44.340 | 6.747 | 75.600 |

### 3.3.2 Target predictions and functional annotation for expressed miRNAs

To explore the functional role of our miRNAs in diverse biological processes, their putative target genes were predicted using the open source web server psRNATarget (Dai et al. 2011). Most of miRNAs were found to target more than one transcript. Indeed, for the 273 miRNAs annotated, 2.316 targets were identified, of which 1.237 (53 %) had a known annotation. The functions of the target genes were different, including transcription factors, involvement in metabolism, resistance and cell differentiation. A large group of targets (91) contained the term resistance in the gene annotation (e.g. resistance to late-blight, verticillium wilt and root-knot nematodes) and 58 of them were resistance protein membrane receptors (LRR). Another group of predicted miRNA targets included transcripts encoding for transcription factors, such as R2-R3 MYB-related (four transcripts regulated by the conserved miR319 and miR159), GRAS (seven transcripts regulated by the conserved miR477 and miR171 and the novel miR9078, miR7478, miR6244 and miR1383), APETALA2 (one transcripts targeted by the conserved miR172 and novel miR1507) and SBP (six transcripts regulated by three members belonging to the conserved miRNA156 family and the novel miRNA5806). Besides resistance genes and transcription factors, 432 targets were annotated as cold responsive genes in *S. commersonii* genome (data not shown). Among them, we found five different *S. commersonii*-specific miRNA that might potentially regulate four different *CBL*-interacting serine threonine-proteins (*CIPKs*) known to be cold responsive. *CIPKs* proteins are the major candidates for cold signal transduction. For instance, different members of the *CIPK* gene family have been found to respond to different stimuli in specific plant tissues and at particular developmental stages in *Arabidopsis* and rice (Kolukisaoglu et al. 2004). However, studies on this gene family and its role under low temperatures in wild *S. commersonii* are still lacking. Using

pylomeDB (http://phylomedb.org/), we found that genemark_scaffold12273_abinit_gene_0_9 and maker_scaffold35102_snap_gene_0_78 clustered with two paralogs of *CIPK16.* (PGSC0003DMP400049575 and PGSC0003DMP400046460) in *S. tuberosum* (Fig. 4), whereas augustus_masked_scaffold8144_abinit_gene_0_0 and augustus_masked_scaffold11835_abinit_gene_0_3 were not present in any tree of the database. Through our goal to better elucidate the dynamics of cold stress signal, our next step will be the genome wide identification and characterization of *CIKPs* involved in the molecular mechanisms of cold tolerance in *S. commersonii.*

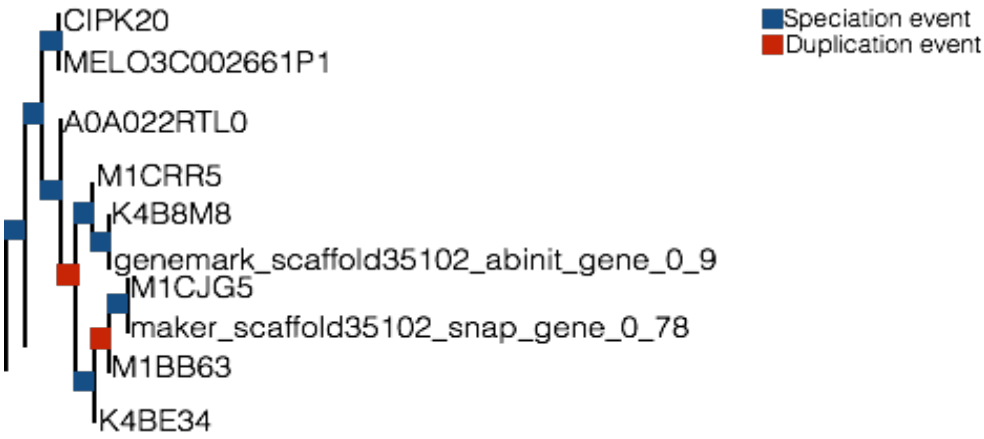


**Fig. 4**. Screenshot representing the two *S. commersonii* cbl-interacting serine threonine-protein (CIPK) annotated as CIPK16 i*n S. tubersoum*

miRNAs involved in their own biogenesis have been also found. Indeed, two different miRNAs (miR6567/6568) have as target the two isoforms of *DCL2d* (maker_scaffold2147_snap_gene_0_30), one of the members of *DCL's* gene family described in chapter IV.

### 3.3.3 miRNAs, tasiRNA and other smallRNA involved in cold stress

In non acclimated conditions (NACC), seven miRNAs were significantly down-regulated after 30min in cmm1T (Tab. 3A).

**Tab. 3** Total number (up-regulated; down-regulated) of differentially expressed (DE) miRNAs, siRNAs and tasiRNAs in cmm1T and cmm6-6 following non acclimated (NACC) and acclimated (ACC) cold stress. Only those that significantly ($P < 0.05$) changed their expression were considered.

A) NACC

| | miRNAs DE | tasiRNA DE | siRNAs DE |
|---|---|---|---|
| cmm1T | 7 (0; 7) | 102 (74; 28) | 5.758 (3.947; 1.811) |
| cmm6-6 | 8 (4; 4) | 425 (277; 148) | 15.369 (8.476; 6.893) |
| common | 0 | 56 | 1.393 |

B) ACC

| | miRNAs DE | tasiRNA DE | siRNAs DE |
|---|---|---|---|
| cmm1T | 36 (21; 15) | 797 (151; 646) | 27.132 (5.065; 22.067) |
| cmm6-6 | 4 (0; 4) | 425 (78; 347) | 20.277 (5.166; 15.111) |
| common | 1 | 138 | 6.189 |

Three of them (all novel) showed a $\log_2FC < -1.5$, whereas the others (conserved miR6027, miR408a, miR408b and the novel miRNA_5621) were weakly differentially expressed (Fig. 5).

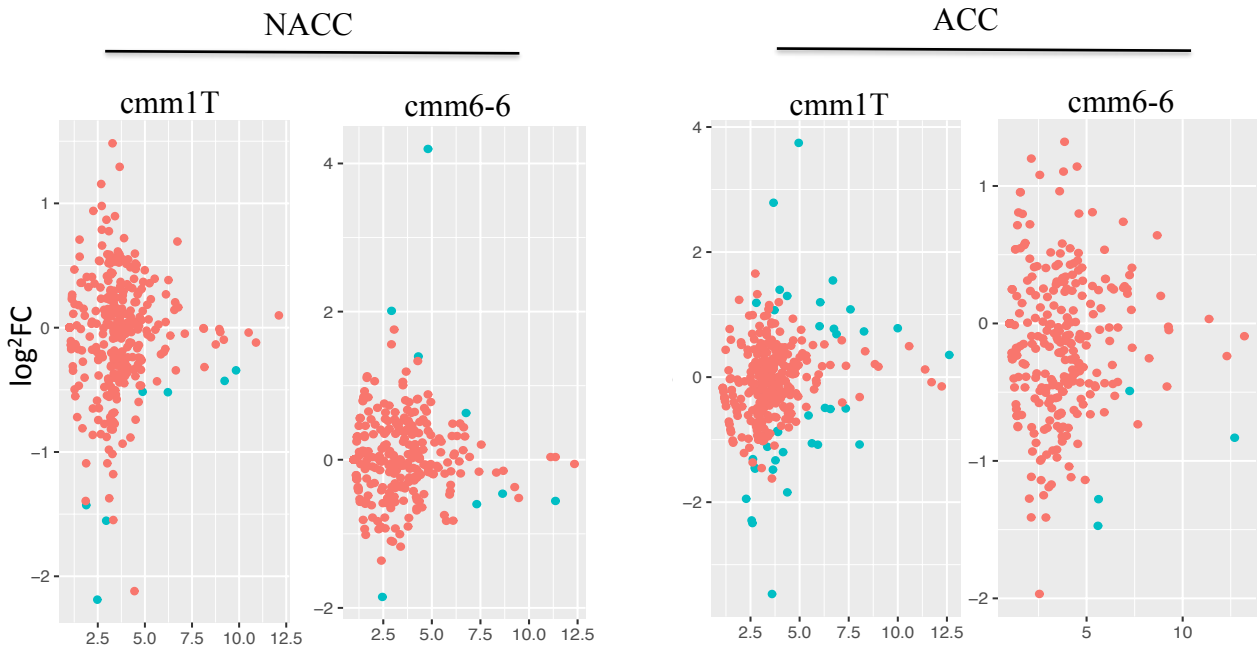


**Fig. 5** MA plots of the differentially expressed miRNAs under NACC and ACC conditions. The y-axis corresponds to the mean expression value of −log10 (q-value), and the x- axis displays the

$log^2$ fold change value. The blue dots represent significant differentially miRNAs, the red dots denote those that are not differentially expressed.

The seven differentially expressed miRNA could regulate a total of 63 putative targets, of which 38 had known GO term. Among them, we found 7 *S. commersonii* genes annotated as cold responsive: 2 Dead box ATP-binding proteins, 2 ATP-binding proteins, 1 $K^+$ channel, 1 1,3-beta glucanase and 1 alpha-L-fucosidase. By contrast, in the cold susceptible cmm6-6, eight miRNA were differentially expressed following NACC stress, of which 4 cold-induced and 4 cold-suppressed (Tab 3A). None of them was in common with cmm1T. Within miRNAs that were cold-induced, miRNA_12069, miRNA_13631 and miRNA_14664 (all *S. commersonii*-specific) showed a $log_2FC > 1.5$, whereas the other (miRNA_12366) was weakly differentially expressed. Within the down-regulated, conserved miRNAs (miR482b and miR5300) and novel miRNA_9288 had a $log_2FC$ value lower than -1, whereas the novel miRNA_11851 showed a down-regulation of two-fold compared with the control. The eight differentially expressed miRNA could potentially regulate 61 miRNA targets, of which 37 had known GO term. Sixteen were cold responsive and were all involved in different biological processes. Following NACC experiments, the differential analysis of tasiRNAs and siRNAs was also performed (Fig. 6). A total of 102 tasiRNA (74 up-regulated and 28 down-regulated) were found in cmm1T after 30min of stress, whereas a higher number was found in cmm6-6 (425, of which 277 up-regulated and 148 down-regulated) (Tab 3A).

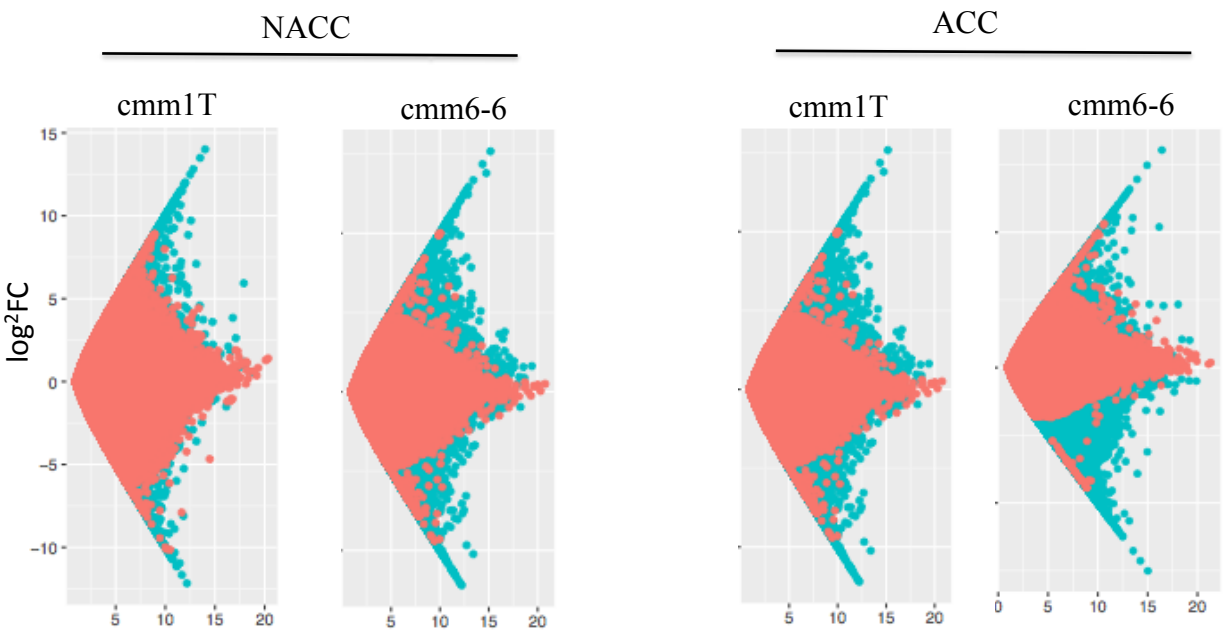


**Fig. 6** MA plots of the differentially expressed tasiRNAs and siRNAs under NACC and ACC conditions. The y-axis corresponds to the mean expression value of −log10 (q-value), and the

x- axis displays the $\log^2$ fold change value. The blue dots represent significant differentially tasiRNAs and siRNAs, the red dots denote those that are not differentially expressed.

The largest difference between the two clones was found in the expression of the secondary siRNA (different from miRNAs and tasiRNA). Indeed, 5.758 siRNAs (3.947 up-regulated and 1.811 down-regulated) were differentially expressed in cmm1T following NACC stress, whereas a higher number was found in cmm6-6 (15.369, of which 8.476 up-regulated and 6.893 down-regulated) (Tab 3A).

After acclimation (ACC), 36 miRNAs (15 cold-suppressed and 21 cold-induced) significantly changed their expression in cmm1T (Tab 3B). Among them, eight different miRNA (three cold-induced and five cold-suppressed) showed $\log_2$FC value > than |1.5|. All miRNA were annotated as *S. commersonii*-specific. As regards the down-regulated miRNA, two of them (miRNA_11331 and miRNA_2736) target two transcription factors belonging to WRKY (WRKY27) and zinc-finger families, whereas the others might potentially regulate more than one target (from two to seven). As regards cmm6-6, four miRNA (all cold-suppressed) were differentially regulated. The conserved miR6027 had ten different targets, including the same DEAD box ATP-binding protein that we have already described in cmm1T during NACC stress. MiRNA_6896 targets a l-aspartate oxidase, an enzyme involved in cellular metabolism. Following acclimation, the differential analysis of tasiRNA and siRNAs was also performed (Fig. 6). Compared to NACC experiment, cmm1T increased the number of tasiRNA differentially expressed (797, of which 151 up-regulated and 646 down-regulated) (Tab 3B). In cmm6-6 we found the same number of tasiRNA (425) detected in NACC experiments. However, compared to NACC stress, the down-regulation was more prevalent than the up-regulation. Indeed, 347 tasiRNA were down-regulated, whereas 78 were up-regulated. Finally, a strong siRNA down-regulation was observed in both clones (Tab 3B). Indeed, 22.067 of 27.132 and 15.111 of 20.277 secondary siRNA were down-regulated in cmm1T and cmm6-6, respectively.

### 3.3.4 Pearson correlations between miRNAs and RNAseq data

To determine the variation of the differentially expressed miRNAs in NACC and ACC conditions, miRNA data were analyzed by Pearson's correlation coefficient (R) with RNAseq data previously reported in chapter II (Fig. 7). Transcripts or miRNA, which were not detected as differentially expressed in any condition, were removed from downstream analyses. In detail, in cmm1T under NACC stress there was a strong negative correlation between the novel miRNA_1794 (R = - 0.8) and the conserved miR408a (R = - 0.99) and their targets (cop_1 homolog and geranial-10-hydroxilase, respectively). As far as cmm6-6 is concerned, under the same conditions three different miRNAs (novel miRNA_14664 and conserved miR482b and miR5300) showed an R-value greater than -0.9, resulting in the up-regulation of their targets. Under ACC stress conditions, among the 36 miRNAs differentially expressed, only one (novel miRNA_4540) was negatively correlated with its target (fatty acid hyperoxide lyase), whereas two miRNAs (novel miRNA_6896 and conserved miR6027) were anti-correlated. The down-regulation of the novel miRNA_6896 induced the up-regulation of two genes with unknown annotation (R = -0.92 and R = -0.83), while the conserved miR6027 regulated five different transcripts, all with R-value ranging between -0.83 and -0.99.

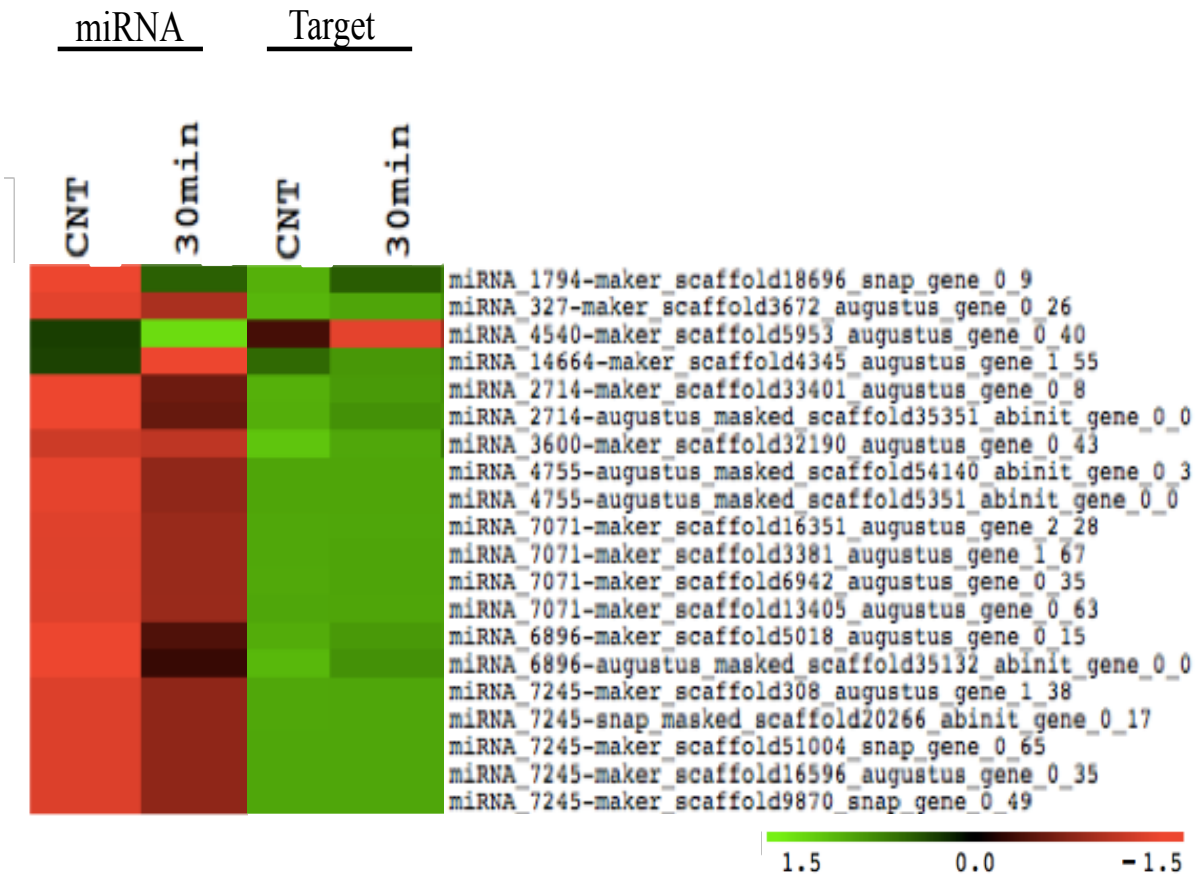


**Fig. 7** Pearson correlation between miRNAs and their targets. Only those miRNA negatively correlated with their targets (P value < 0.05 ) under both NACC and ACC conditions are shown. The expression of miRNA and their target in control condition (CNT) and 30 min following NACC and ACC stresses is shown.

### 3.4. Discussion

Increasing evidence shows that small RNA plays an important role in developmental stages as well as in gene regulation processes upon biotic and abiotic stresses (Sunkar et al. 2007; Chen et al. 2009; Ruiz-Ferrer et al. 2009). MiRNAs involved in cold stress were investigated in details only in few major plants species such as *Populus trichocarpa* (Lu et al. 2008), *B. distachyon* (Zhang et al. 2009), *Vitis vinifera* (Sun et al. 2015), *Medicago sativa* (Shu et al. 2016) and *Triticum aestivum* (Song et al. 2017). MiRNAs were also described in *S. tuberosum* (Guo et al. 2007; Yang et al. 2010; Hwang et al. 2011; Kitazumi et al. 2011; Zhang et al. 2013; Lackotia et al. 2014). However, they were identified comparing various tissues but not following cold stress. In addition, they were predicted using either previously available sequence data or large-scale data but focusing only on

conserved miRNAs. By contrast, secondary siRNAs such as tasiRNAs have never been investigated and, further, no miRNA/siRNA studies in wild potato species have been reported so far. Hence, this represents the first comprehensive analysis of sncRNAs in a wild potato with noteworthy traits for breeding efforts.

### 3.4.1 miRNAs

Our analysis showed that the 24nt length class dominates in the dataset, strongly suggesting that small RNA-directed heterochromatin-silencing plays an important role in this species, as also reported in *Arabidopsis*, tomato, pepper, cucumber, maize, peanut, pepper, citrus and rice (Fahlgren et al. 2007; Morin et al. 2008; Moxon et al. 2008; Song et al. 2011; Chi et al. 2011; Martinez et al. 2011; Li et al. 2012; Hwang et al. 2013). Our research allowed us to annotate 273 distinct miRNAs in the genome of *S. commersonii*. The majority appeared to mature from transcripts of intergenic regions or from introns of annotated genes, which is in line with previous findings (Voinnet et al. 2009). We were able to identify 44 conserved miRNA (belonging to 27 families) and 229 *S. commersonii*-specific miRNA, which are consistent with the number of miRNA identified in cultivated *S. tuberosum* (Guo et al. 2007; Yang et al. 2010; Hwang et al. 2011; Zhang et al. 2013). A recent study by Lackotia et al. (2014) led to the identification of 89 conserved miRNAs (belonging to 33 families) and 147 potato-specific miRNAs. Such disagreement could be due to the fact that we limited our analysis to only leaves, whereas Lackotia et al. (2014) analyzed different tissues. Therefore, in the dataset produced here many other miRNAs with different spatio-temporal expression might be undetected. Out of 27 conserved miRNA families identified in *S. commersonii*, 8 were the same described by the previously mentioned authors. The most abundant conserved miRNA families in *S. commersonii* were miR166, miR169 and miR156. The first family was already described (Zhang et al. 2009; Zhang et al. 2013; Lakhotia et al. 2014) and our results confirm that it is highly expressed in potato species. Our analysis showed that the members of miR166 were highly similar to those of other species such as *Arabidopsis* and tomato. Hence, the evolutionary conservation of miR166 strongly supports its important role in plant growth, development, and adaptation (Sunkar and Jagadeeswaran 2008). The ability of miR166 to down-regulate HD-ZIP-type transcription factors including *Phabulosa* (*PHB*) has been established in *Arabidopsis* and maize (Juarez et al. 2004; Chuck and O'Connor 2010) and the molecular dynamics and function of miR166/HD-ZIP-type engagement have been recently provided by Kitazumi (2016). These authors reported a model where its activity could provide a means to control growth when cellular resources and intermediates are being prioritized for defense and repair-related processes. Hence, this family might have a prominent role under stress conditions in *S.*

*commersonii*. MiR156 (three members identified) and miR169 (five members) families have been already reported in *S. tuberosum*. Three members of miR156 were described by Yang and colleagues (2010), 12 by Zhang et al. (2013) and 6 by Lackotia et al. (2014). MiR169 was also described by Zhang et al. (2013) and Lackotia et al. (2014) and their results yielded the same results as ours.

Generally, the targets of plant miRNAs have perfect or near-perfect complementary sites, allowing their identification using bioinformatic prediction methods. Here, we identified potential targets with a wide variety of predicted functions, such as transcription factors, genes with a role in defense mechanisms, kinases and ion homeostasis genes. Most of the conserved miRNA that targeted plant transcription factors (SBP, GRAS, AP2) were found to be highly similar to the conserved miRNA targets predicted in *Arabidopsis* and other plants (Chen et al. 2009; Song et al. 2010; Zuo et al. 2012). This further underlies the role of conserved miRNAs in essential biological processes. Interestingly, among the miRNA targets, a large group was relative to *NBS-LRR* receptors, suggesting a link between miRNA and defense gene expression. It is known that high expression of plant *NBS-LRR* defense genes is often lethal to plant cells (Collier et al. 2011). Our results strengthen the hypothesis that miRNAs might be considered as master regulators of the plant *NB-LRR* defense genes, as already descrbed by Zhai and colleagues (2011), and thus might be involved in the regulation of plant immunity (Fei et al. 2006). Clues regarding these interactions are also given by Zhang (2016), and consistent with their results and those reported by Gonzalez (2015), we found that different members of miR482 target *NB-LRR* genes, suggesting a conserved role of this family in plant evolution.

### 3.4.2 miRNA involved in cold stress

Our results suggested that miRNAs are important regulatory nodes for the cold response of *S. commersonii,* as already described for other plant species (Sunkar et al. 2004; Liu et al. 2008; Lu et al. 2008; Zhang et al. 2009). Although different plant species possess different set of miRNAs responding to cold, it seems that there is a core of miRNAs that are shared by most of them. For instance, among miRNA that showed differences in their expression under cold stress in *S. commersonii*, 10 were in common with those described by others authors. Great emphasis was given on the cold-induced miRNAs, whereas the cold-suppressed miRNAs have received little attention. Our study indicated that down-regulation of miRNAs, following NACC stress conditions, is more prevalent than up-regulation in this kind of stresses. These results are consistent with Zhang (2009), who showed that most of the miRNA involved in cold stress in *Brachipodium* were down-regulated. Particularly interesting is miR408 family. It has been annotated in more than 20 plant

species, making it among the most conserved miRNA families in plants. miR408 has been found to be significantly induced in senescing, accompanied by increased levels of the cleavage products generated from its validated targets in *A. thaliana* (Abdel-Ghany and Pilon, 2008; Zhang and Li, 2013; Thatcher et al. 2015). Various studies showed that it was differentially expressed in response to different abiotic stresses including cold, drought, osmotic and oxidative burdens (Sunkar and Zhu, 2004; Liu et al. 2008; Shen et al. 2010; Trindade et al. 2010; Zhou et al. 2010a,b; Mutum et al. 2013; Jovanovic et al. 2014; Zhang et al. 2014a,b). Ma et al. (2015) measured its expression in *Arabidopsis* under six different abiotic stresses, in order to obtain an overall view of its possible involvement. The authors showed that miR408 was clearly induced under cold, salinity and oxidative stresses, whereas there was no drastic changes under osmotic and drought stresses. In contrast with Ma et al. (2015), Trindade et al. (2010) reported that miR408 was induced in response to water deficit in *Medicago truncatula*. During NACC, we found that two members of miR408 family were down-regulated in tolerant cmm1T but not in susceptible cmm6-6, leading the induction of its targets (belonging to the family of laccase). These act on phenols and similar molecules, performing one-electron oxidations. It is proposed that these genes play a role in the formation of lignin by promoting the oxidative coupling of monolignols, a family of naturally occurring phenols that might confer cold tolerance (Sunkar and Zhu, 2004; Sunkar et al. 2012; Rajwanshi et al. 2014). Hence, as noted for the results showed here, miR408 could be important for the fine-tuning expression of a set of genes encoding copper-containing proteins, which are involved in different metabolic processes, as already formulated by Ma and colleagues (2015). Our results shed also a new light on the regulation of cold responsive genes under NACC stress. For instance, miRNAs that have as target the CIPK16 and other cold relative genes were identified under NACC stress in tolerant cmm1T. In light of our results, further analysis will be carried out to better clarify the interactions between CIPK16 and its targets under cold stress conditions. In our scenario, as already described for the *NBS-LRR* receptors, in normal conditions miRNA might suppress the expression of stress relative genes to induce them when they are needed, enhancing plant tolerance. If on one hand it is interesting to deep analyze the miRNAs differentially expressed in the tolerant clone cmm1T, on the other it is intriguing to analyze the behavior of those miRNAs that were cold regulated in the cold susceptible clone cmm6-6. Particularly interesting is the down regulation of the conserved miR482 under NACC stress. This miRNA has been reported to regulate several *NBS-LRR* defense genes during fungal pathogen infection in cotton (Zhu et al. 2013). MiR482-mediated target cleavage is expected to cause not only decay of their target mRNAs but also production of phaseRNAs as already described in tomato and *M. truncatula* (Zhai et al. 2011; Li et al. 2012; Shivaprasad et al. 2012). In our experiments, the down regulation of miR408 might

led the induction of several *NBS-LRR* genes that in turn might be lethal in cmm6-6 after 30 min of NACC stress (Collier et al. 2011).

Cmm1T and cmm6-6 are both able to acclimate, and as such offer an interesting model to unreveal the dynamics of cold acclimation. In this context, particularly interesting is miR4376. This miRNA was down regulated in both cmm1T and cmm6-6 under ACC conditions. Wang et al. (2011) showed that miR4376 regulates the expression of an auto inhibited $Ca^{2+}$-ATPase in *S. lycopersicum ACA10* gene, which plays a critical role in tomato reproductive growth. This likely has broad implications in light of the role of Ca2+ signaling under stress conditions in plants and other organisms (Sanders et al. 1999, 2002; Sze et al. 2000; Carafoli, 2002; Hepler, 2005; Boursiac and Harper, 2007; McAinsh and Pittman, 2009; Dodd et al. 2010; Kudla et al. 2010). Hence, this miRNA is a good candidate for further molecular studies. Another interesting feature that we found in our experiment was relative to the regulation of DEAD box proteins. It is known that these proteins may be directly involved in temperature sensing. A DEAD box RNA helicase (*LOS4*) has been found to activate the expression of *DREB1/CBF* during cold acclimation in *Arabidopsis* (Gong et al. 2005). Our results showed that miRNAs are involved in the regulation of these proteins. Their expression is induced in cmm6-6 under ACC stress, and probably it leads to an increase of the transport of cold responsive proteins into the nucleus. Since cold tolerance genes such as *CBF*, *FR2*, *COR*, and *LEA* do not seem to be the target of any miRNAs, such an observation is interesting and may be the result of cold stress. The same hypothesis has been formulated by Song et al. (2017) to explain the dynamics of cold tolerance in wheat. It seems that the drastic changes in miRNA expression levels in *S. commersonii* leaves will cause sensitive reaction of their targets which will regulate the expression of transcription factors, membrane receptors, genes involved in lignin biosynthesis and other correlate pathways. These mechanisms may be explored in breeding to enhance cold tolerance in potato.

### 3.4.3 Other secondary siRNAs

In the last few years, as a result of extensive genome sequencing in plants coupled with small RNA analysis, many new small RNAs, which are not miRNA, have been described. However, compared to miRNAs, the loci producing these endogenous regulatory siRNAs have had little systematic curation. Recent work has demonstrated an abundance of loci producing phasiRNAs in monocots, with examples in rice, maize, and *Brachypodium* (Johnson et al. 2009; International *Brachypodium* Initiative, 2010; Song et al. 2017), but nothing is known regarding their role in potato. Consistent with these authors, we found that *S. commersonii* harbors thousands of endogenous siRNAs loci. The vast majority of the clusters identified here are of the 24-nt variety, confirming the data

reported by Axtell et al. (2013). Indeed, the authors, who developed ShortStack software for comprehensive analysis of small RNA-seq data, described that 24-nt siRNAs were the most frequent type in all four plants data sets analyzed (*Arabidopsis*, tomato, rice and maize). Similar to Johnson et al. (2009), who showed that nearly half of siRNA in rice overlap with genome repeats, the siRNAs loci in *S. commersonii* matched with transposable elements, confirming that they are implicated in transcriptional silencing of repeated regions of the genome. The majority of them were expressed in our samples, suggesting that these small RNAs perform housekeeping functions in the silencing of many repeats in all tissues. Our strategy allowed us to identified and distinguish tasiRNA from other endogenous siRNAs, in order to understand and study their role under low temperatures. Our data demonstrated that both not-acclimated and acclimated conditions alter the expression of both tasiRNAs and secondary siRNAs, confirming that they might be involved in the response to abiotic stresses in plants. Clues regarding the role of tasiRNAs under abiotic stress were given by Li et al. (2014). The authors reported that two genes (HTT1 and HTT2), which are targets of TAS1 (*trans-acting* siRNA precursor-1), were highly expressed under high temperatures, whereas TAS1a was strongly suppressed. At the same time, the overexpression of HTT1 and HTT2 induced the expression of several heat shock proteins, leading plants to stronger thermo tolerance. Recently, Dutta et al. (2017) also demonstrated a role of tasiRNAs under biotic stress. They profiled the expression of *TAS* with the ability to generate four tasiRNAs, following pathogen inoculation of susceptible and resistant wheat. The authors revealed that the targets of these tasiRNAs included a-gliadin proteins, leucine rich repeat, trans-membrane proteins, glutathione-S-transferase, and fatty acid desaturase among others, which are either stress responsive genes or essential for normal growth and plants development, and they were induced under pathogenesis. Our findings are also in line with previous works by Sunkar and Zhu (2004) and Yao et al. (2010). The former gave the first evidence that siRNAs are involved in abiotic stress responses in *Arabidopsis*, whereas the latter demonstrated the role of four siRNAs in wheat seedlings under cold, heat, salt, or drought stresses. However, at present, we do not know yet the role of siRNA target genes. Allen et al. (2005) demonstrated a model in which miRNA-guided synthesis of pre-tasiRNA transcripts, followed by the formation of dsRNA that is in turn processed by DCL, yielding phased tasiRNAs biosynthesis, which negatively regulate other genes. Interestingly, we found that two well-described miRNAs (miR408 and miR4376) might induce the formation of endogenous siRNAs, which might be differentially expressed in our conditions. Overexpression or knockdown of some of these genes can be performed in the future to answer this question.

### 3.5. Conclusions

*S. commersonii* is potentially the best model species to study the molecular mechanisms involved in cold tolerance as well as the ability to cold acclimate in potato. Hence, we decided to carry out for the first time an extensive smRNA analysis to identify conserved and novel cold-responding miRNAs in its genome. This work provides the first small RNA expression profile of *S. commersonii* under cold conditions. Our small RNA sequencing data revealed regulatory roles of miRNAs, tasiRNAs and siRNAs during cold response. Our results allowed us also to identify miRNAs-target directly involved in stress response, suggesting that their differential expression in our clones might contribute to their different ability to face cold stress. It is clear from our study that *S. commersonii* has evolved sophisticated miRNA/siRNA-mediated pathways to cope with changing environments. Further study to clarify the role of these mechanisms and to identify smRNAs targets will improve our understanding on the response of plants to cold stress and will offer further approaches to breed for cold tolerance.

**References**

Abdel-Ghany SE and Pilon M (2008) MicroRNA-mediated systemic down-regulation of copper protein expression in response to low copper availability in *Arabidopsis*. J Biol Chem 283:15932–15945.

Allen E, Xie Z, Gustafson AM, and Carrington JC (2005) microRNA-directed phasing during *trans*-acting siRNA biogenesis in plants. Cell 121:207–221.

Aversano R, Contaldi F, Ercolano MR, et al (2015) The *Solanum commersonii* genome sequence provides insights into adaptation to stress conditions and genome evolution of wild potato relatives. Plant Cell 27:954-968. doi: 10.1105/tpc.114

Axtell MJ (2013) Classification and comparison of small RNAs from plants. Annu Rev Plant Biol 64:137–159

Axtell MJ (2013) ShortStack: comprehensive annotation and quantification of small RNA genes. RNA 19:740–51.

Borges F, Martienssen RA (2015) The expanding world of small RNAs in plants. Nat Rev Mol Cell Biol 16:727–41.

Boursiac Y, and Harper JF (2007) The origin and function of calmodulin regulated Ca2+ pumps in plants. J Bioenerg Biomembr 39:409–414

Carafoli E (2002) Calcium signaling: a tale for all seasons. Proc Natl Acad Sci USA 99:1115–1122.

Carthew RW, Sontheimer EJ (2009) Origins and mechanisms of miRNAs and siRNAs. Cell 136:642–655. 27.

Chen X: Small RNAs and their roles in plant development. Annu Rev Cell Dev Biol 2009:25:21–44

Chi X, Yang Q, Chen X, Wang J, Pan L, Chen M, Yang Z, He Y, Liang X, Yu S (2011) Identification and characterization of microRNAs from peanut (*Arachis hypogaea L.*) by high-throughput sequencing. PLoS One 6:e27530.

Chuck G, O'Connor D (2010) Small RNAs going the distance during plant development. Curr Opin Plant Biol 13:40–45

Collier SM, Hamel LP, Moffett P (2011). Cell death mediated by the N-terminal domains of a unique and highly conserved class of NB- LRR protein. Mol Plant Microbe Interact 24:918–931

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR (2013) STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29:15-21. 10.1093/bioinformatics/bts635.

Dodd AN, Kudla J, and Sanders D (2010) The language of calcium signaling. Annu Rev Plant Biol 61:593–620.

Fahlgren N, Howell MD, Kasschau KD, Chapman EJ, Sullivan CM, Cumbie JS, Givan SA, Law TF, Grant SR, Dangl JL, Carrington JC (2007) High-throughput sequencing of *Arabidopsis* microRNAs: evidence for frequent birth and death of MIRNA genes. PLoS One 2:e219

Fei Q, Xia R, & Meyers BC (2013) Phased, Secondary, Small Interfering RNAs in Posttranscriptional Regulatory Networks. The Plant Cell 25:2400–2415.

Gong Z, Dong C, Lee H, Zhu J, Xiong L, & Gong D (2004). A *DEAD Box RNA Helicase* is Essential for mRNA Export and Important for Development and Stress Responses in *Arabidopsis*. Plant Cell 17:1–12.

Gonzalez VM, Muller S, Baulcombe D, Puigdomenech P (2015) Evolution of *NBS-LRR* gene copies among dicot plants and its regulation by members of the miR482/2118 superfamily of miRNAs. Mol Plant 8:329–331.

Guo Q, Xiang A, Yang Q, & Yang Z (2007). Bioinformatic identification of microRNAs and their target genes from *Solanum tuberosum* expressed sequence tags. Chinese Science Bulletin 52: 2380–2389.

Hepler PK (2005) Calcium: A central regulator of plant growth and development. Plant Cell 17: 2142–2155.

Hwang D-G, Park JH, Lim JY, Kim D, Choi Y, Kim S, Reeves G, Yeom SI, Lee JS, Park M, Kim S, Choi IY, Choi D, Shin C (2013) The hot pepper (*Capsicum annuum*) MicroRNA transcriptome reveals novel and conserved targets: a foundation for understanding MicroRNA functional roles in hot pepper. PLoS ONE 8:e64238.

Jiang HS, Lei R, Ding SW, Zhu SF (2014) Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. BMC Bioinformatics 15:182

Johnson C, Kasprzewska A, Tennessen K, Fernandes J, Nan GL, Walbot V, Sundaresan V, Vance V, Bowman LH (2009) Clusters and superclusters of phased small RNAs in the developing inflorescence of rice. Genome Res 19:1429–1440

Jovanovic Z, Stanisavljevic N, Mikic A, Radovic S and Maksimovic V (2014) Water deficit down-regulates miR398 and miR408 in pea (*Pisum sativum L.*). Plant Physiol. Biochem. 83:26–31.

Juarez MT, Kui JS, Thomas J, Heller BA, Timmermans MC (2004) MicroRNA- mediated repression of rolled leaf1 specifies maize leaf polarity. Nature 428:84–8.

Kitazumi A, Kawahara Y, Onda TS, Koeyer DDe, & de Los Reyes BG (2015) Implications of miR166 and miR159 induction to the basal response mechanisms of an andigena potato (*Solanum tuberosum* subsp. andigena) to salinity stress, predicted from network models in *Arabidopsis*. Genome / National Research Council Canada = Genome / Conseil National de Recherches Canada, 58:13–24.

Kozomara A, Griffiths-Jones S (2011) miRBase: integrating microRNA annotation and deep-sequencing data. Nucleic Acids Res,39:152–157.

Kudla J, Batistic O, and Hashimoto K (2010) Calcium signals: The lead currency of plant information processing. Plant Cell 22:541–563.

Lakhotia N, Joshi G, Bhardwaj AR, Katiyar-Agarwal S, Agarwal M, Jagannath A, … Kumar A (2014). Identification and characterization of miRNAome in root, stem, leaf and tuber developmental stages of potato (*Solanum tuberosum* L.) by high-throughput sequencing. BMC Plant Biology 14:1–6.

Langmead B, et al. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome Genome Biol 10:25

Li D, Wang L, Liu X, Cui D, Chen T, Zhang H, Jiang C, Xu C, Li P, Li S, Zhao L, Chen H (2013) Deep sequencing of maize small RNAs reveals a diverse set of MicroRNA in dry and imbibed seeds. PLoS One 8:e55107.

Li F, Pignatta D, Bendix C, Brunkard JO, Cohn MM, Tung J, Sun H, Kumar P, Baker B (2012): MicroRNA regulation of plant innate immune receptors. Proc Natl Acad Sci USA 109:1790–1795.

Li W, and A Godzik (2006) CD-HIT: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22:1658–1659.

Liu HH, Tian X, Li YJ, Wu CA and Zheng CC (2008) Microarray- based analysis of stress-regulated microRNAs in *Arabidopsis thaliana*. RNA 14:836–843.

Ma C, Burd S, & Lers A (2015) miR408 is involved in abiotic stress responses in *Arabidopsis*. The Plant Journal 84:169–187

Martinez G, Forment J, Llave C, Pallas V, Gomez G (2011) High-throughput sequencing, characterization and detection of new and conserved cucumber miRNAs. PLoS One 6:e19523. 11.

McAinsh MR, and Pittman JK (2009) Shaping the calcium signature. New Phytol 181:275–294. Meyers,

Meng Y, Shao C, Wang H, Chen M (2011) The regulatory activities of plant microRNAs: a more dynamic perspective. Plant Physiol 157:1583–1595.

Morin RD, Aksay G, Dolgosheina E, Ebhardt HA, Magrini V, Mardis ER, Sahinalp SC, Unrau PJ (2008) Comparative analysis of the small RNA transcriptomes of *Pinus contorta* and *Oryza sativa.* Genome Res 18:571–584. 16.

Moxon S, Jing R, Szittya G, Schwach F, Rusholme Pilcher RL, Moulton V, Dalmay T (2008) Deep sequencing of tomato short RNAs identifies microRNAs targeting genes involved in fruit ripening. Genome Res 18:1602–1609

Mutum RD, Balyan SC, Kansal S, Agarwal P, Kumar S, Kumar M and Raghuvanshi S (2013) Evolution of variety-specific regulatory schema for expression of osa-miR408 in indica rice varieties under drought stress. FEBS J 280:1717–1730.

Sunkar R, Chinnusamy V, Zhu JH, Zhu JK (2007) Small RNAs as big players in plant abiotic stress responses and nutrient deprivation Trends Plant Sci 12:301–309

Rajagopalan R, Vaucheret H, Trejo J, Bartel DP (2006) A diverse and evolutionarily fluid set of microRNAs in *Arabidopsis thaliana*. Genes Dev, 20:3407–3425.

Rajwanshi R, Chakraborty S, Jayanandi K, Deb B and Lightfoot DA (2014) Orthologous plant microRNAs: microregulators with great potential for improving stress tolerance in plants. Theor Appl Genet 127:2525–2543.

Lu SF, Sun YH, Shi R, Clark C, Li LG, Chiang VL (2005) Novel and mechanical stress-responsive microRNAs in *Populus trichocarpa* that are absent from *Arabidopsis* Plant Cell 17:2186–2203.

Sanders D, Brownlee C, and Harper JF (1999) Communicating with calcium. Plant Cell 11:691–706.

Sanders D, Pelloux J, Brownlee C, and Harper JF (2002). Calcium at the crossroads of signaling. Plant Cell 14:S401–S417.

Shen J, Xie K. and Xiong L. (2010) Global expression profiling of rice microRNAs by one-tube stem-loop reverse transcription quantitative PCR revealed important roles of microRNAs in abiotic stress responses. Mol Genet Genomics 284:477–488.

Shivaprasad PV, Chen H, Patel K, Bond DM, Santos BACM, & Baulcombe DC (2012) A MicroRNA Superfamily Regulates Nucleotide Binding Site – Leucine-Rich Repeats and Other mRNAs. The Plant Cell 24:859–874.

Shu Y, Liu Y, Li W, Song L, Zhang J, & Guo C (2016) Genome-Wide Investigation of MicroRNAs and Their Targets in Response to Freezing Stress in *Medicago sativa* L., Based on High Throughput Sequencing. G3: Genes Genomes Genetics 6:755–765.

Shukla LI, Chinnusamy V, Sunkar R (2008) The role of microRNAs and other endogenous small RNAs in plant stress responses. Biochim Biophys Acta 1779:743–748

Song C, Wang C, Zhang C, Korir NK, Yu H, Ma Z, Fang J (2010) Deep sequencing discovery of novel and conserved microRNAs in trifoliate orange (*Citrus trifoliata*). BMC Genomics 11:431

Song QX, Liu YF, Hu XF, Zhang WK, Ma B, Chen SY, Zhang JS (2011) Identification of miRNAs and their target genes in developing soybean seeds by deep sequencing. BMC Plant Biol, 11:5. 42.
Song G, Zhang R, Zhang S, Li Y, Gao J, Han X, … Li G (2017) Response of microRNAs to cold treatment in the young spikes of common wheat. BMC Genomics 18:212

Sun X, Fan G, Su L, Wang W, Liang Z, Li S, & Xin H (2015) Identification of cold-inducible microRNAs in grapevine. Frontiers in Plant Science 6:1–13. doi:10.3389/fpls.2015.00595

Sunkar R, Jagadeeswaran G (2008) In silico identiWcation of conserved microRNAs in large number of diverse plant species. BMC Plant Biol 8:37

Sunkar R and Zhu JK (2004) Novel and stress-regulated microRNAs and other small RNAs from *Arabidopsis*. Plant Cell 16:2001–2019.

Sunkar R, Li YF and Jagadeeswaran G (2012) Functions of microRNAs in plant stress responses. Trends Plant Sci 17:196–203.

Sze H, Liang F, Hwang I, Curran AC, and Harper JF (2000) Diversity and regulation of plant Ca2+ pumps: Insights from expression in yeast. Annu Rev Plant Physiol Plant Mol Biol 51:433–462.

Thatcher SR, Burd S, Wright C, Lers A and Green PJ (2015) Differential expression of miRNAs and their target genes in senescing leaves and siliques: insights from deep sequencing of small RNAs and cleaved target RNAs. Plant Cell Environ 38:188–200.

Trindade I, Capitao C, Dalmay T, Fevereiro MP and dos Santos DM (2010) miR398 and miR408 are up-regulated in response to water deficit in *Medicago truncatula*. Planta 231:705–716.

Ruiz-Ferrer V, Voinnet O (2009) Roles of plant small RNAs in biotic stress responses Annu Rev Plant Biol 60:485–510

Vaucheret H (2006) Post-transcriptional small RNA pathways in plants: mechanisms and regulations. Genes Dev 20:759–771.

Voinnet O (2009): Origin, biogenesis, and activity of plant microRNAs. Cell 136:669–687.

Wang Y, Itaya A, Zhong X, Wu Y, Zhang J, van der Knaap E, … Ding B (2011) Function and evolution of a microRNA that regulates a Ca2+-ATPase and triggers the formation of phased small interfering RNAs in Tomato Reproductive Growth. The Plant Cell Online 23:3185–3203.

Chen X (2009) Small RNAs and their roles in plant development Annu Rev Cell Dev Biol 25:21–44

Yao Y, Ni Z, Peng H, Sun F, Xin M, Sunkar R, Zhu JK, Sun Q (2010) Non-coding small RNAs responsive to abiotic stress in wheat (Triticum aestivum L.) Funct Integr Genomics 10:187–190

Yang W, Liu X, Zhang J, Feng J, Li C, & Chen J (2010). Prediction and validation of conservative microRNAs of *Solanum tuberosum* L. Molecular Biology Reports 37:3081–3087.

Zhai J, Jeong DH, De Paoli E, Park S, Rosen BD, Li Y, Gonzalez AJ, Yan Z, Kitto SL, Grusak MA, et al. (2011) MicroRNAs asmaster regulators of theplant NB-LRR defense gene family via the production of phased, *trans*-acting siRNAs. Genes Dev 25:2540–2553.

Zhang HY and Li L (2013) SQUAMOSA promoter binding protein-like7 regulated microRNA408 is required for vegetative development in *Arabidopsis*. Plant J 74:98–109.

Zhang HY, Zhao X, Li JG, Cai HQ, Deng XW and Li L (2014a) MicroRNA408 is critical for the HY5-SPL7 gene network that mediates

Zhang J, Xu Y, Huan Q, & Chong K (2009) Deep sequencing of *Brachypodium* small RNAs at the global genome level identifies microRNAs involved in cold stress response. BMC Genomics 10: 449.

Zhang R, Marshall D, Bryan GJ, & Hornyik C (2013). Identification and Characterization of miRNA Transcriptome in Potato by High-Throughput Sequencing. PLoS ONE 8: doi:10.1371/journal.pone.0057233

Zhang XN, Li X and Liu JH (2014b) Identification of conserved and novel cold-responsive microRNAs in trifoliate orange (*Poncirus trifoliata (L.) Raf*.) using high-throughput sequencing. Plant Mol Biol Rep 32:328–341.

Zhang Y, Xia R, Kuang H, & Meyers BC (2016). The Diversification of Plant NBS-LRR Defense Genes Directs the Evolution of MicroRNAs That Target Them. Molecular Biology and Evolution 33:2692–2705.

Zhou L, Liu Y, Liu Z, Kong D, Duan M and Luo L (2010a) Genome- wide identification and analysis of drought-responsive microRNAs in *Oryza sativa*. J Exp Bot 61:4157–4168.

Zhou M, Gu L, Li P, Song X, Wei L, Chen Z and Cao X (2010b) Degradome sequencing reveals endogenous small RNA targets in rice (*Oryza sativa* L. ssp. indica). Front Biol 5:67–90.

Zhu QH, Fan L, Liu Y, Xu H, Llewellyn D, & Wilson I (2013). miR482 regulation of *NBS-LRR* defense genes during fungal pathogen infection in cotton. PLoS ONE 8. doi:10.1371/journal.pone.0084390

Zuo J, Zhu B, Fu D, Zhu Y, Ma Y, Chi L, Ju Z, Wang Y, Zhai B, Luo Y (2012): Sculpting the maturation, softening and ethylene pathway: The influences of microRNAs on tomato fruits. BMC Genomics 13:7-10.

# Chapter IV. *Dicer-like* and *RNA-dependent RNA polymerase* gene families identification and annotation in the cultivated *Solanum tuberosum* and its wild relative *S. commersonii*

Submitted to Planta: Esposito Salvatore[1], Aversano Riccardo[1*], D'Amelia Vincenzo [1], Villano Clizia[1], Alioto Daniela[1], Mirouze Marie[2], Carputo Domenico[1*]

## ***Dicer-like*** **and** ***RNA-dependent RNA polymerase*** **gene families identification and annotation in the cultivated** ***Solanum tuberosum*** **and its wild relative** ***S. commersonii***

**Main conclusion**

**We provide advances in understanding RNA silencing in cultivated *Solanum tuberosum* and wild *S. commersonii.* We shed new lights on *DCL* and *RDR* gene evolution, localization and expression.**

**Abstract**

*Dicer-like* (*DCL*) and *RNA-dependent RNA polymerase* (*RDR*) genes form the core components to trigger RNA silencing in plant and their concerted activities are required in mediating plant tolerance. Little information is available on the *DCLs* and *RDRs* in non-model species. The cultivated potato (*Solanum tuberosum*) and its cold-tolerant wild relative *S. commersonii* offer a valuable opportunity to advance our understanding on these genes, as their genome sequences are available. To determine the extent of diversification and evolution of *DCLs* and *RDRs* in these species, we performed a comparative analysis among their orthologs. Seven *DCLs* were identified in the two species, whereas seven and six *RDR* genes were found in *S. tuberosum* and *S. commersonii*, respectively. Based on phylogenetic analysis, an increase in the number of *DCL* and *RDR* paralogs in both species occurred, highlighting a possible diversification of their functions through duplication events. *DCL* and *RDR* expressions were investigated in different tissues and under cold and virus stresses. Divergent expression profiles of the tandem duplicated genes in different tissues were found. *DCL* paralogs showed a contrasting expression in *S. tuberosum* and *S. commersonii* following cold stress and virus infection. By contrast, no change in the *RDR* transcript activity was detected following both stresses. Overall, this study provides the first comparative genomic analysis of the core components of the RNAi machinery in *Solanaceae*.

**Keywords**

**gene-silencing, comparative genomics, microRNA biogenesis, cold stress, PVY**

### 4.1. Introduction

RNA interference (RNAi) has the potential to control gene expression either through the repression of transcription initiation (transcriptional gene silencing, TGS) or through mRNA degradation (post-transcriptional gene silencing, PTGS) (Matzke et al. 2015). TGS and PTGS depend on small non-coding RNA (ncRNA), especially small interfering RNAs (siRNA) or microRNAs (miRNA).

Once produced from double-stranded RNA (dsRNA) precursors, they are loaded into RNA-induced silencing complexes (RISCs) to act in gene silencing mechanisms. The biogenesis and function of plant small ncRNAs involve various protein families, such as DICER-like (DCLs), HYPONASTIC LEAVES1 (HYL1), C2H2 Zn-finger protein SERRATE (SE), HEN1, HASTY, RNA-dependent RNA polymerases (RDRs) and ARGONAUTES (AGOs). Among them, DCLs and RDRs represent the core components for small ncRNA biogenesis and recently their structure and functions have been clarified (Matzke et al. 2015). In particular, DCLs are endoribonucleases consisting of six domains, namely DExD-helicase (DExDc), helicase-C (HelC), Dicer dimer, PAZ, RNaseIII C (RIBO III) and double stranded RNA-binding (dsRB) (Margis et al. 2006). Multiple copies of *DCL* genes exist in eukaryotes and produce siRNAs with different length (Shabalina et al. 2008). They can play distinct roles in plant development and environmental interactions (Qi et al. 2005; Kapoor et al. 2008; Bai et al. 2012). The action of DCLs is complement by RDRs, which use single-stranded RNAs as templates to generate dsRNA intermediates, which are later processed by DCLs. RDRs are defined by the presence of a conserved RDR catalytic domain and exist in multiple copies. The combination of different copies of *DCL* and *RDR* genes underpins distinct and overlapping processes. For example, the most abundant small ncRNA class (24 nucleotides) arises through *RDR2*/*DCL3*/DNA-dependent RNA polymerase IV (PolIV) activities (Pontier et al. 2005; Tran et al. 2005), whereas the *RDR6*/*DCL4* activity catalyzes the formation of trans-acting siRNAs (tasiRNAs), which are endogenous regulators of several mRNAs (Xie et al. 2005; Yoshikawa et al. 2005).

TGS and PTGS regulatory networks have been reported to mediate plant tolerance to different stresses. For instance, recent studies have shed some lights on the contribution of RNAi against virus attacks (Palaez and Sanchez 2013). Likewise, mechanisms of PTGS mediate mRNA translation of numerous genes involved in abiotic stresses such as drought, salinity, cold, heat, light, and oxidation (Kraiwesh et al. 2012; Crisp et al. 2016). Despite the pivotal role played by small ncRNAs in plant response to environmental constraints, there is little information available on proteins controlling their biogenesis in non-model wild species, which are known to often tolerate biotic and abiotic stresses better than their closely related cultivated species (Carputo et al. 2013; Yoo et al. 2014; Chen et al. 2015). In particular, the diversity between wild and cultivated species in terms of candidate orthologous gene pairs with important role in RNAi has not received much attention, although several studies have shown that differences between wild and cultivated plants in adapting to stress conditions lies on their different gene expression regulation (Besser et al. 2009; Ghorecha et al. 2014; Yoo et al. 2014). Nowadays, the availability of the genome sequence of several species provides an unprecedented opportunity for exploring gene family evolution by

comparative analyses. In this regard, a valuable example is given by the potato. Indeed, the genomes of cultivated *Solanum tuberosum* and wild tuber-bearing *S. commersonii* are available (Potato Genome Sequencing Consortium 2012; Aversano et al. 2015). Here we report the identification and the characterization of *DCL* and *RDR* genes in these two species. To determine whether they possess a diverse regulation of these genes, the expression profiles of *DCLs* and *RDR*s in different tissues and after both cold stress and Potato Virus Y (PVY) infection were investigated. *S. commersonii* is tolerant to cold and susceptible to PVY, whereas *S. tuberosum* is susceptible to both stresses. Overall, this study provides the first comparative genomic analysis of the core components of the RNAi machinery in *Solanaceae*.

## 4.2. Materials and methods

### 4.2.1. Identification of candidate DCL and RDR genes and their respective regulatory elements

We used the known protein sequences of four *Arabidopsis thaliana* and seven *S. lycopersicum DCL* genes as queries to search for the amino acid orthologs in *S. commersonii* (PI243503) and *S. tuberosum* Group Phureja (clone DM1-3 516 R44) through the Blastp tool of SpudDB (Hirsch et al.2013). In particular, amino acid sequences of *DCL* genes of *A. thaliana* were downloaded from the National Center for Biotechnology Information (NCBI) database, while the proteins of *S. lycopersicum* were those identified by Bai et al. (2012). Conserved domains were searched using the Pfam protein family database (Pfam 24.0) (Finn, 2011). The newly identified genes were named based on the nomenclature used for the previously identified genes and on their phylogenetic relatedness to other members of the same family. A similar strategy was used to identify *RDR* genes in *S. commersonii*. Amino acid sequences of *StRDR* genes identified by Hunter et al. (2016) and those of *S. lycopersicum* were used. The exon-intron organization of *DCL* and *RDR* genes was determined using the online Spidey program (Wheelan et al. 2001) (https://www.ncbi.nlm.nih.gov/spidey/) by comparing their full-length coding sequences (CDS) with the corresponding genomic sequences downloaded from each database. For each *DCL* and *RDR* gene, 1,500 nucleotides upstream the translation initiation codon were extracted using a custom PERL script. They were further used for the transcription factor binding sites (TFBSs) analysis, using the PlantCARE tool (Lescot et al. 2002).

### 4.2.2. Sequence alignments and phylogenetic analysis

*S. tuberosum, S. commersonii, S. lycopersicum* and *A. thaliana* candidate protein sequences were retrieved from the dedicate databases. Phylogenetic analyses were conducted using MEGA7

(Tamura et al. 2007). Multiple sequence alignments were carried out using CLUSTALW (Larkin et al. 2007). The evolutionary history was inferred by using the Neighbor-Joining (NJ) model. The bootstrap consensus tree was built using 1000 replicates. Branches corresponding to partitions reproduced in less than 30% of bootstrap replicates were collapsed. Initial trees for the heuristic search were obtained automatically to a matrix of pairwise distances estimated using a Jones–Thornton–Taylor (JTT) model, and then selecting the topology with superior log likelihood value.

### 4.2.3 Public RNAseq-based expression analysis

The transcriptional activity of *DCL* and *RDR* genes in the cultivated potato was estimated using the publically available RNAseq dataset deposited in SpudDB. Briefly, raw single-end fastq files (ERR029909, ERR029910, ERR029911, ERR029914, ERR029916 and ERR029917) were downloaded from the study named "Transcriptome Analysis of the potato" retrieved in NCBI SRA database (https://trace.ncbi.nlm.nih.gov/Traces/sra/). To remove unwanted sequences originating from organelles, reads were mapped against the mitochondrial (S_tuberosum_Group_Phureja_mitochondrion_DM1-3-516-R44) and chloroplast (S_tuberosum_Group_Phureja_chloroplast_DM1-3-516-R44) genomes using BOWTIE2 2.2.2 (Langmead and Salzberg, 2012) with sensitive local mapping. Unmapped reads were considered for the next analysis and were mapped against the *S. tuberosum* genome (ITAG annotation v1). The BAMs files were then analyzed using Cufflinks-Cuffquant software (version 2.2.1) to assemble the aligned reads and to access the transcriptome complexity. Expression values for each gene were estimated based on RPKM (Reads Per Kilobase of transcript per Million mapped reads) using the default options. For *S. commersonii* we used fastq data related to four tissues (flower, leaf, tuber and stolon) and deposited under study SRP050412.

### 4.2.4 Cold stress assay

Young plants of *S. commersonii* (clone cmm1T of PI243503) and *S. tuberosum* (variety Blondy), respectively tolerant and susceptible to cold (Carputo et al. 2007), were micro-propagated *in vitro* as described by D'Amelia et al. (2017). Four-weeks old vitroplants were transplanted into 14-mm pots filled with sterile soil and grown for two weeks at 24°C prior submitting them to non-acclimated (NACC) and acclimated (ACC) stress conditions. In particular, in NACC experiment three plants of each genotype were challenged for 30min at -2°C, while three plants were chosen as controls and kept at 24°C. In the ACC experiment, 6 plants/genotype were acclimated at 4°C for 2 weeks. Then, three of them were transferred for 30min at -2°C, while the others were used as control and kept at 4°C. In both experiments, an environmentally controlled cold room was used.

Young leaf samples were collected from all replicates at the end of each stress and from control plants. Samples were stored at -80°C before RNA extraction.

### 4.2.5 Potato Virus Y inoculation

Young plants of clone cmm1T and cv. Blondy were mechanically inoculated with Potato Virus Y tuber necrotic strain ($PVY^{NTN}$). The virus, isolated in Italy from *S. tuberosum* (Barone et al. 2009), was reactivated from dehydrated infected tissues grinding them in extraction buffer (10 mM phosphate, 1% carborundum, pH 7-7.2). The virus was maintained on *Nicotiana glutinosa*, under greenhouse conditions (20–24°C). Extract was prepared by grinding *N. glutinosa* symptomatic leaves (1g) in 10 ml of extraction buffer. Inoculum was applied on ten plantlets of each species at the 2–3 leaf stage rubbing sap with a latex-gloved finger. Ten plants of each species were inoculated with buffer (mock control). Young leaf samples were collected from virus infected and mock control plants after 8 hours and stored at -80°C before RNA extraction.

### 4.2.6 RNA isolation and quantitative Real-Time PCR

Total RNA was isolated from leaves using Spectrum™ Plant Total RNA kit (Sigma-Aldrich, St. Louis, MO, USA) following the manufacturer's instructions. For quantitative real-time PCR (qRT-PCR) experiments, samples were treated with RNase-free DNase and one microgram of RNA was reverse transcribed using the Superscript III (Life Technologies, Carlsbad, California, USA) following the manufacturer's protocol. The 20µl final reaction was diluted in 200µl using sterile water. A 1µl aliquot of cDNA was used in a qRT-PCR, with the addition of 0.3µM of each specific primers and FAST SYBR Green master mix (Applied Biosystems, Foster City, CA) to a final reaction volume of 20µl. The qRT-PCRs were performed using an ABI 7900HT Real Time PCR System (Applied Biosystems). Primer sequences are reported in the Table S1. Relative expression was calculated using the adenine phosphoribosyl transferase (*aprt*) and the elongation factor (*EF*) genes (Nicot et al. 2005) through the $\Delta\Delta$Ct method (Livak and Schmittgen 2001). To find differentially expressed genes (DEGs) between *S. commersonii* and *S. tuberosum*, a Student's t-test with 100 permutations and a critical P-value < 0.05 was performed.

## 4.3. Results

### 4.3.1. Comparative analysis highlights structural differences among *DCL* and *RDR* orthologs in Solanum

We carried out a comparative analysis to unearth possible specificities and functional differences between *DCL* and *RDR* orthologs in *S. commersonii* and *S. tuberosum*. Based on *Arabidopsis* and

tomato sequence similarity, six and seven *DCL* genes with a functional structure were identified in *S. tuberosum* and *S. commersonii*, respectively (Table 1a). In addition, through the exploitation of ITAG annotation v1 we found a genomic region of 10 kb on chromosome 11 of potato (present in the PGSC data), downstream to Sotub11g010160.1.1 and Sotub11g010170.1.1, where three short sequences (Sotub11g010180.1.1, Sotub11g010190.1.1, Sotub11g010200.1.1) were separately annotated as Dicer-like (data not shown). No orthologs of these three sequences could be found in other sequenced species. Specifically, Sotub11g010180.1.1 encodeS for PAZ and Dicer dimerization domains, whereas Sotub11g010190.1.1 and Sotub11g010200.1.1 encode partial Helicase domains. To solve this ambiguity, we analyzed in depth the genomic region harboring the three annotated genes. We re-annotated these genes as one single gene containing 23 exons. Based on the distance homology with *A. thaliana* and *S. lycopersium* genes, we named it *StDCL2c*. Hence, we found a seventh *DCL* gene in *S. tuberosum*. On average, our *DCL*s showed putative protein lengths ranging from 707 (augustus_masked_scaffold13333_abinit_gene_0_0, ScDCL4) to 2208 amino acids (maker_scaffold2147_snap_gene_0_30, ScDCL2d) (Table 1a).

**Table 1** Identification of *DCL* and *RDR* genes in *Solanum tuberosum*, *S. commersonii*, *S. lycopersicum* and *S. pennellii*. The accession number corresponds to the annotation provided by ITAG2.3.2 for *S. lycopersicum* and ITAG1.0 for *S. tuberosum*.

| | Species | Annotation | Length (aa) |
|---|---|---|---|
| *Dicer-like genes* | | | |
| DCL1 | *S. tuberosum* | Sotub10g006390.1.1 | 1889 |
| | *S. commersonii* | maker_scaffold5659_augustus_gene_1_36 | 1918 |
| | *S. lycopersicum* | Solyc10g005130.2.1 | 1888 |
| | *S. pennellii* | Sopen10g001150 | 1914 |
| DCL2a | *S. tuberosum* | Sotub06g012900.1.1 | 1401 |
| | *S. commersonii* | maker_scaffold17562_snap_gene_1_46 | 1288 |
| | *S. lycopersicum* | Solyc06g048960.2.1 | 1399 |
| | *S. pennellii* | Sopen06g015290 | 1354 |
| DCL2b | *S. tuberosum* | Sotub11g010160.1.1 | 1400 |
| | *S. commersonii* | maker_scaffold2147_snap_gene_0_28 | 1359 |
| | *S. lycopersicum* | Solyc11g008540.1.1 | 1352 |
| | *S. pennellii* | Sopen11g004330 | 1405 |
| DCL2c | *S. tuberosum* | This study | 861 |
| | *S. commersonii* | maker_scaffold1113_snap_gene_0_59 | 1054 |
| | *S. lycopersicum* | Solyc11g008520.1.1 | 1428 |
| | *S. pennellii* | Sopen11g004310 | 1499 |
| DCL2d | *S. tuberosum* | Sotub11g010170.1.1 | 1372 |
| | *S. commersonii* | maker_scaffold2147_snap_gene_0_30_1 | 2208 |
| | *S. lycopersicum* | Solyc11g008530.1.1 | 1317 |
| | *S. pennellii* | Sopen11g004320 | 1499 |
| DCL3 | *S. tuberosum* | Sotub08g016690.1.1 | 1368 |
| | *S. commersonii* | maker_scaffold6660_snap_gene_2_27 | 1678 |
| | *S. lycopersicum* | Solyc08g067210.2.1 | 1431 |
| | *S. pennellii* | Sopen08g021150 | 1658 |
| DCL4 | *S. tuberosum* | Sotub07g007080.1.1 | 1538 |
| | *S. commersonii* | augustus_masked_scaffold13333_abinit_gene_0_0 | 707 |
| | *S. lycopersicum* | Solyc07g005030.2.1 | 1536 |
| | *S. pennellii* | Sopen07g001140 | 1690 |
| *RNA-dependent RNA polymerase genes* | | | |
| RDR1a | *S. tuberosum* | Sotub05g010490.1.1 | 1116 |
| | *S. tuberosum* | Sotub05g010500 | 1116 |
| | *S. commersonii* | maker_scaffold34693_augustus_gene_0_18 | 754 |
| | *S. lycopersicum* | Solyc05g007510.2.1 | 1115 |
| | *S. pennellii* | Sopen05g003390 | 1114 |
| RDR1b | *S. tuberosum* | NA | NA |
| | *S. commersonii* | maker_scaffold9269_augustus_gene_0_64 | 1115 |
| | *S. lycopersicum* | NA | NA |
| | *S. pennellii* | NA | NA |
| RDR2 | *S. tuberosum* | NA | NA |
| | *S. commersonii* | maker_scaffold14992_augustus_gene_0_25 | 1119 |
| | *S. lycopersicum* | Solyc03g114140.2.1 | 1120 |
| | *S. pennellii* | Sopen03g033200 | 1123 |
| RDR3f | *S. tuberosum* | Sotub06g014160.1.1 | 554 |
| | *S. commersonii* | maker_scaffold20276_snap_gene_0_42 | 971 |
| | *S. lycopersicum* | Solyc06g051170.2.1 | 831 |
| | *S. lycopersicum* | Solyc06g051190.2.1 | 1014 |
| | *S. pennellii* | Sopen06g017040 | 728 |
| RDR3e | *S. tuberosum* | Sotub12g009540.1.1 | 335 |
| | *S. commersonii* | NA | NA |
| | *S. lycopersicum* | NA | NA |
| | *S. pennellii* | Sopen12g003370 | 246 |
| RDR3d | *S. tuberosum* | Sotub12g009540.1.1 | 554 |
| | *S. commersonii* | augustus_masked_scaffold6297_abinit_gene_0_1 | 529 |
| | *S. lycopersicum* | Solyc12g008410.1.1 | 831 |
| | *S. lycopersicum* | Solyc04g014870.2.1 | 1198 |
| | *S. pennellii* | NA | NA |
| RDR6a | *S. tuberosum* | Sotub04g012000.1.1 | 1199 |
| | *S. commersonii* | maker_scaffold1433_augustus_gene_0_57 | 1198 |
| | *S. pennellii* | Sopen04g007000 | 1197 |
| RDR6b | *S. tuberosum* | Sotub08g020870.1.1 | 732 |
| | *S. commersonii* | maker_scaffold9967_snap_gene_0_63 | 1068 |
| | *S. lycopersicum* | Solyc08g075820.2.1 | 1180 |
| | *S. lycopersicum* | Solyc08g075825.1.1 | 748 |
| | *S. pennellii* | Sopen08g024460 | 1162 |

Comparing *S. tuberosum* and *S. commersonii* DCL proteins, we found high similarities in sequence length and composition of DCL1, DCL3 and DCL4. Their similarity percentage was comprised between 96.8% (DCL3) and 97.8% (DCL1). By contrast, higher variability was found in the proteins belonging to the DCL2 group, resulting in similarity percentage ranging from 79.1% (DCL2a) to 80% (DCL2d). The protein sequence of *S. commersonii* DCL2d (maker_scaffold2147_snap_gene_0_30, ScDCL2d) was much longer compared with its orthologs in *S. tuberosum*. However, our data revealed that except for DCL2d, *DCL* loci have been substantially conserved between the two species. To investigate the evolution of multiple *DCLs* in potato, we analyzed their genomic distribution by localizing genes on potato chromosomes (data not shown). A total of seven *S. tuberosum DCL* genes were distributed on five chromosomes. Three of them (*StDCL2b*, *StDCL2c* and *StDCL2d*) were on chromosome 11, while the remaining were located on chromosomes 6, 7, 8 and 10 and they represented single copy *StDCL* genes. To localize *DCL* genes in *S. commersonii* (for which a physical map is not available yet), we blasted the respective scaffolds versus *S. tuberosum* chromosomes, confirming that *DCL* genes remain in corresponding chromosomes (syntheny) and orders (collinearity) in both species (data not shown). To broaden our understanding into possible trans-species polymorphisms between wild and cultivated contexts, we extended our analysis also to the cultivated tomato (*S. lycopersicum*) and its wild relative *S. pennellii*. They are members of the *Solanaceae* family, closely related to potato species, and their genomes have been recently sequenced (The Tomato Genome Consortium 2012, Bolger et al. 2014). Common features shared by the two cultivated species compared to their respective wild relatives were also found. Interestingly, within the DCL1 subfamily, both *S. commersonii* and *S. pennellii* harbored a 26 amino acids insertion in RiboIII functional domain (Fig. S1). This insertion makes the domain longer in wild species compared to the cultivated counterparts. Similar differences were also found in DCL3, where the typical DExD domain was lacking in the two cultivated species compared with their respective wild orthologs. To gain more insights into *DCL* structural evolution, their exon-intron structure was examined (Fig. 1a). Intron and exon number was generally conserved within members of the same group in all four genomes analyzed, varying from 16 (DCL2a) to 25 (DCL3 and DCL4). However, differences in exon number and gene length were identified between the two potatoes. *S. commersonii* had the lowest (11) number of exons in *ScDCL4* and the highest (39) in *ScDCL2d* compared with the respective orthologs of *S. tubersoum* and the other two species. This confirmed our previous observation on predicted amino acid sequences.

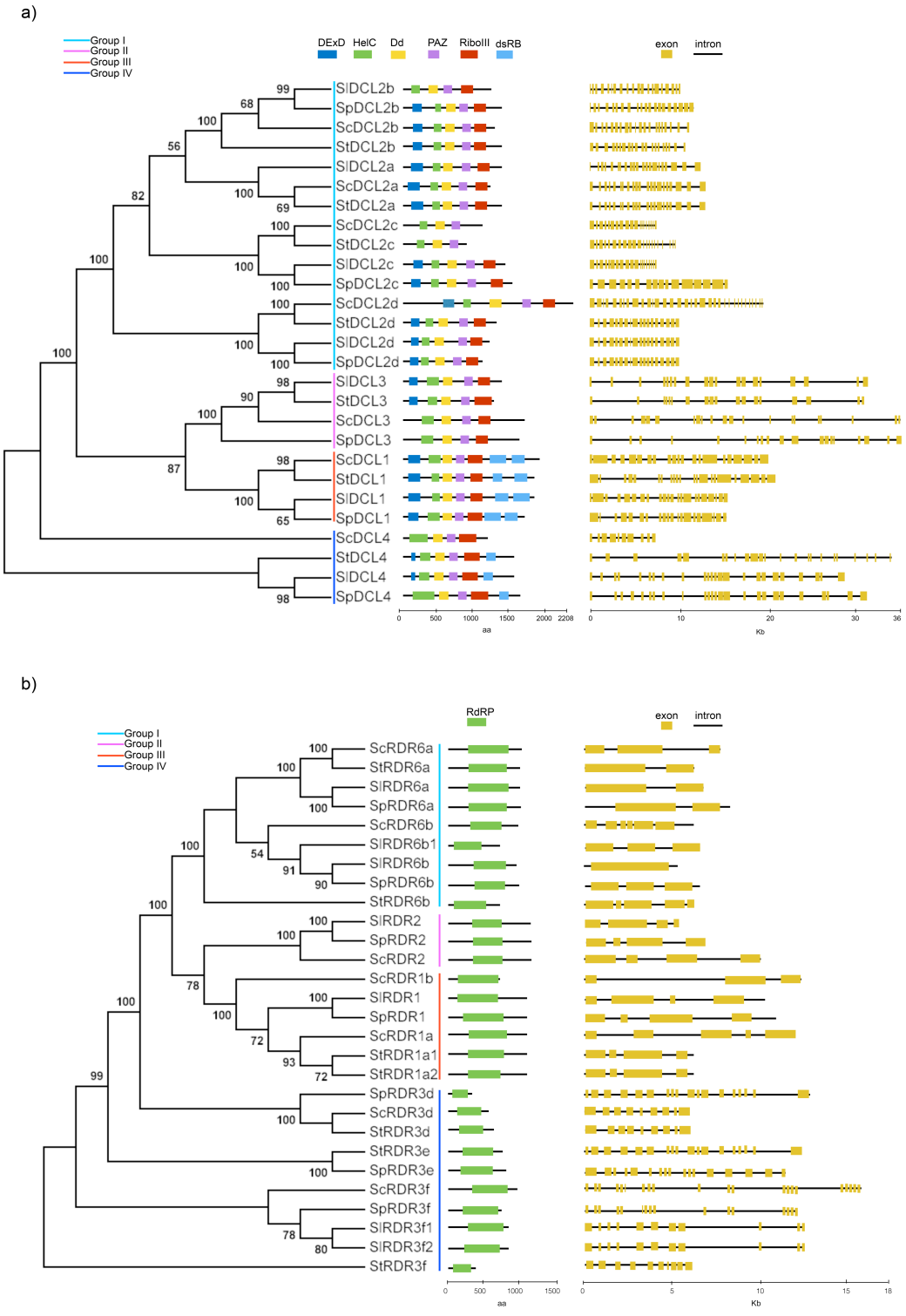


**Fig. 1** Domain distribution of *Solanum commersonii*, *S. tuberosum*, *S. lycopersicum* and *S. pennellii* Dicer-like (a) and RDR (b) proteins. The following conserved domains are present: DExD (DEAD/DEAH box helicase), HelC (Helicase conserved C-terminal domain), Dd (Dicer dimer), PAZ, RiboIII (Ribonuclease III domain), and dsRB (double-strand RNA-binding) in StDCL, ScDCL and SlDCL proteins; RdRP (RNA-dependent RNA polymerase) in StRDR, ScRDR and SlRDR proteins and SpRDR.

We performed the same comparative analysis on *RDR* genes. We started from the previously identified *RDRs* in *S. tuberosum* and *A. thaliana*. Using them as baits, we found eight *RDR* putative genes in *S. commersonii* and *S. lycopersicum*, seven in *S. tuberosum* and six in *S. pennellii* (Table 1b). The classification of the orthologous sequences was difficult due to the ambiguity of nomenclature in literature. Indeed, the naming of tomato *RDRs* reflected that of *Arabidopsis* (Bai et al. 2012), but no direct orthology was found with the nomenclature available for potato *RDRs* (Hunter et al. 2016). Therefore, a new unique nomenclature for *RDR* gene classes in the genus *Solanum* was established based on phylogenetic relationships and leveraging on the most recent annotations of wild and cultivated genomes (Potato Genome Sequencing Consortium 2011; The Tomato Consortium 2012; Bolger et al. 2014; Aversano et al. 2015) (Table 1). In particular, the *SlRDR3a* (Solyc12g008410) and *SlRDR3b* (Solyc06g051170) described by Bai (2012) were renamed *SlRDR3d* and *SlRDR3f*, respectively (Table 1; Fig. 1). *RDR* genes were distributed on six chromosomes (chr3, chr4, chr5, chr6, chr8 and chr12) in *S. tuberosum*. Blast analysis of *S. commersonii* scaffolds versus *S. tuberosum* chromosomes confirmed that also *RDR* loci have been substantially conserved among the two species. Moreover, our alignment in tomato allowed us to identify two additional *RDR* genes (Solyc06g051190.2.1 and Solyc08g075825.1.1, belonging to SlRDR3f and SlRDR6b clade, respectively) never described before. The full protein sequence of these genes ranged from 246 amino acids in *S. pennellii* (Sopen12g003370, SpRDR3e) to 1180 in *S. lycopersicum* (Solyco8g075820.2.1, SlRDR6b) (Table 1b, Fig. S1). As observed for *DCL* genes, differences at gene level between *S. tuberosum* and *S. commersonii* were also found for *RDRs*. For example, the latter had higher number of exons in *ScRDR6a* and *ScRDR6b* compared with the cultivated potato. Furthermore, common features shared by two cultivated *S. tuberosum* and *S. lycopersicum* compared with their respective orthologs in *S. commersonii* and *S. pennellii* were also found. For example, the proteins encoded by Sopen08g024460 in *S. pennellii* and maker_scaffold9967_snap_gene_0_63_mRNA_1 in *S. commersonii* were roughly 1100 amino acids, whereas the orthologs in the two cultivated species were shorter (732 and 748 amino acids in *S. tuberosum* and *S. lycopersicum*, respectively). In each *RDR* type, the intron-exon number was generally conserved within members of the same group in all genomes analyzed, varying from one (*SlRDR6b*) to 21 (*ScRDR3f*).

### 4.3.2. Phylogenetic analysis reveals evolutive differentiation within *DCL* and *RDR* clades

A phylogenetic analysis was conducted to define evolutionary relationships of *S. tuberosum* and *S. commersonii DCLs* with those of *S. lycopersicum* and *S. pennellii* (Figure 1a). The 26 DCL protein sequences formed 7 clades; four groups were identified in the neighbor-joining tree generated.

DCL2a, DCL2b, DCL2c and DCL2d candidate proteins were respectively assigned to clades 1-4 and, altogether, included in group I; DCL3, DCL1 and DCL4 subfamilies occupied distinct clades (5, 6 and 7, respectively) and comprised groups II, III and IV. The clusterization validated the gene exon structure and the protein domain analysis we previously reported. Differences were observed in all groups and within wild and cultivated species. As regard Group I (in particular DCL2a relatives), the PAZ domain was lacking in *S. commersonii* and *S. pennellii*, but it was present in their respective cultivated forms. By contrast, the DExD domain was present only in the two wild species. In addition, the number or the presence of DCL domains varied among the subfamilies. For example, the ResIII domain was found only within the DCL1 clade. Likewise, the dsRB domain was absent in DCL2 and DCL3 proteins, but it was present twice in DCL1 (Fig. 1).

Regarding RDRs, the neighbor-joining tree derived from protein sequences was resolved into nine clades; four groups were identified (RDR1, RDR2, RDR3 and RDR6) (Fig. 1b). RDR6a and RDR6b (clades 7, 8 and 9, respectively) were comprised in group I. RDR2 and RDR1 occupied distinct clades (6 and 5 respectively) and formed group II and III, while RDR3e, RDR3f and RDR3d candidate genes were assigned to clades 1, 2, 3 and 4, respectively, and were included in group IV. We found that they all shared a common and unique RNA-dependent RNA Polymerase (RdRP) domain. Two new loci in *S. lycopersicum*, namely *SlRDR6b1* (Solyc08g075825.1.1) and *SlRDR3f2* (Solyc06g051190.2.1), were identified. The former is located close to Solyc08g075820.2.1 (*RDR1*), the latter upstream *RDR3f1* (Solyc06g051190.2.1). Further differences among species were observed. For example, duplication of *RDR1* occurred in *S. tuberosum* and *S. commersonii* but not in cultivated and wild tomatoes (Fig. 1b).

### 4.3.3 *Cis*-acting elements are abundant in *DCL* and *RDR* promoters

Numerous cis-acting regulatory elements were identified in the upstream region of *DCLs* (Fig. 2a). We classified them in five classes based on their biological function. The most redundant motifs were the light-responsive elements, widely present in the upstream regions of most *DCL* genes. However, we sought annotated regulatory motifs involved also in stress response. Among them TC-rich repeats, HSE (heat stress element), WUN (responsive to wounding), as-2, LTR (low temperature responsive) and Box-W1. TC-rich repeats were present in the promoter regions of all *DCL* genes except for*DCL2b* and *DCL2c* in both *S. commersonii* and *S. tuberosum*. The HSE motif was found in the upstream regions of *DCL1* and *DCL2b* in both potato species, whereas the as-2-box was found in *DCL2a* and *DCL2c*. The promoter regions of *ScDCL1* and *StDCL1* contained the GCCGAC motif of the LTR elements, that forms the core of DRE sequence. This element was found also in the up-stream region of *StDCL3*, but not in *S. commersonii* due to the lack of the

*ScDCL3* upstream region. In general, *S. commersonii* had a higher number of cis-acting elements compared with the cultivated potato, ranging from 19 (*ScDCL2b*) to 31 (*ScDCL2a*). Overall, the wild potato harbored 166 regulatory elements whereas *S. tuberosum* 147 (Fig. 2a).

Regarding *RDR* genes, we found fewer motifs compared to those identified in *DCLs* (Fig. 2b). As in *DCLs*, the most abundant ones were the light-responsive motifs, whereas the stress-responsive elements were the least frequent. Among them, the LTR motif was identified in *RDR6a* of both species and in *ScRDR3d* and *StRDR3f*, whereas the Box-W1 (TTGAC), known to participate in disease response, in *RDR6a* and *RDR6b* of both potato species and in *StRDR1a*, *StRDR3d*, *ScRDR3e* and *ScRDR3f*. The second most abundant class of motifs identified in *RDRs* was that related to plant hormone response. Indeed, motifs involved in gibberellin (GARE), ABA (ABRE and MBSI) as well as ethylene (ERE) pathways were widely present. Some unknown cis-elements were found, with the Unnamed_4 elements being the most redundant in all the *RDR* promoters. Comparing the two species, we found that they shared the same number of elements in each paralog promoter, except for *RDR1a*. Indeed, only four motifs were identified in *ScRDR1a* promoter, whereas 22 were found in *StRDR1a*. In total, 114 cis-element motifs were identified in both potato species.


Motif
ScDCL1
StDCL1
ScDCL2b
StDCL2a
ScDCL2a
StDCL2b
ScDCL2c
StDCL2c
ScDCL2d
StDCL2d
ScDCL3
StDCL3
ScDCL4
StDCL4
light
ATC
ATCT
Box 4
MRE
ACE
G-box
GATA
I-box
TCT
box II
rbcS-CMA7a
GA
chs-CMA1a
AT1
Sp1
chs-CMA2a
Box I
Box II
GT1
TCCC
as-2-box
AE-box
stress
Box-W1
TC-rich repeats
HSE
WUN
LTR
MBS
hormones
TGA-box
TCA-element
ERE
AuxRR-core
CGTCA
TGACG
GARE
ABRE
Box III
other
CAAT-box
CCAAT-box
O2-site
5UTR Py-rich stretch
AT-rich element
OBP-1 site
Skn-1
GCN4
ARE
circadian
TATCCAT/C
TCCACCT
Unnamed 1

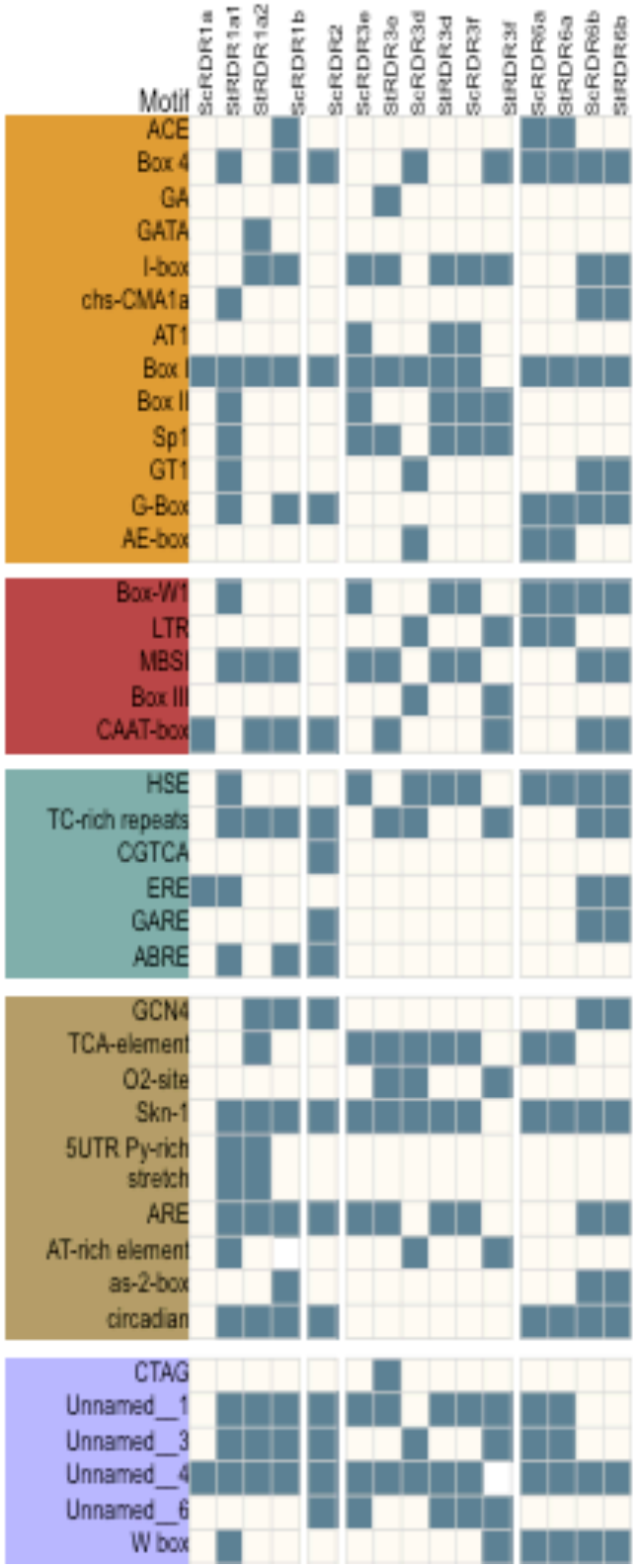


**Fig. 2** List of cis-elements in DCL (a) and RDR (b) promoters of *Solanum tuberosum* and *S. commersonii*. The presence of the elements is reported in blue

### 4.3.4 Expression profiles of *DCL* and *RDR* genes change in different tissues and after stress

Two types of expression analyses were carried out. In the first one the objective was to get a gene expression overview of *DCLs* and *RDRs* in different tissues using both publicly accessible RNAseq data and those available in our laboratory. These data allowed us to curate and validate gene models and to further investigate the role of these genes in different tissues. Gene structure previously described was confirmed for all *DCL* genes, except for the putative *StDCL2c*. Its expression was not supported by RNAseq data, suggesting that *StDCL2c* might be a pseudogene. Indeed, using RepeatMasker we found that the genomic region of *StDCL2c* was interrupted by several transposon insertions (data not shown). All the other *S. commersonii* and *S. tuberosum DCLs* were expressed in flower, leaf, stolon, root and tuber tissues (Fig. 3). The expression patterns of *ScDCL1* and *ScDCL4* matched with those of *S. tuberosum* in all tissues, whereas *ScDCL2b* and *StDCL2b* showed similar patterns in all samples except leaves, where they were expressed at high level in *S. commersonii* but not in *S. tuberosum*. As far as *RDR* genes are concerned, transcript abundance analysis showed a general expression of *RDR* genes in the tissues taken into consideration (Fig. 3), with only exception of *ScRDR6b*, which was not expressed at all. However, *RDR* genes differ in their expression patterns between *S. tuberosum* and *S. commersonii*. For example, *RDR1a, RDR1b, RDR3d* and *RDR6a* were expressed at high level only in *S. commersonii* roots and tubers. Furthermore, the expression level of *ScRDR3d* was higher in all tissues compared with its ortholog in *S. tuberosum*.

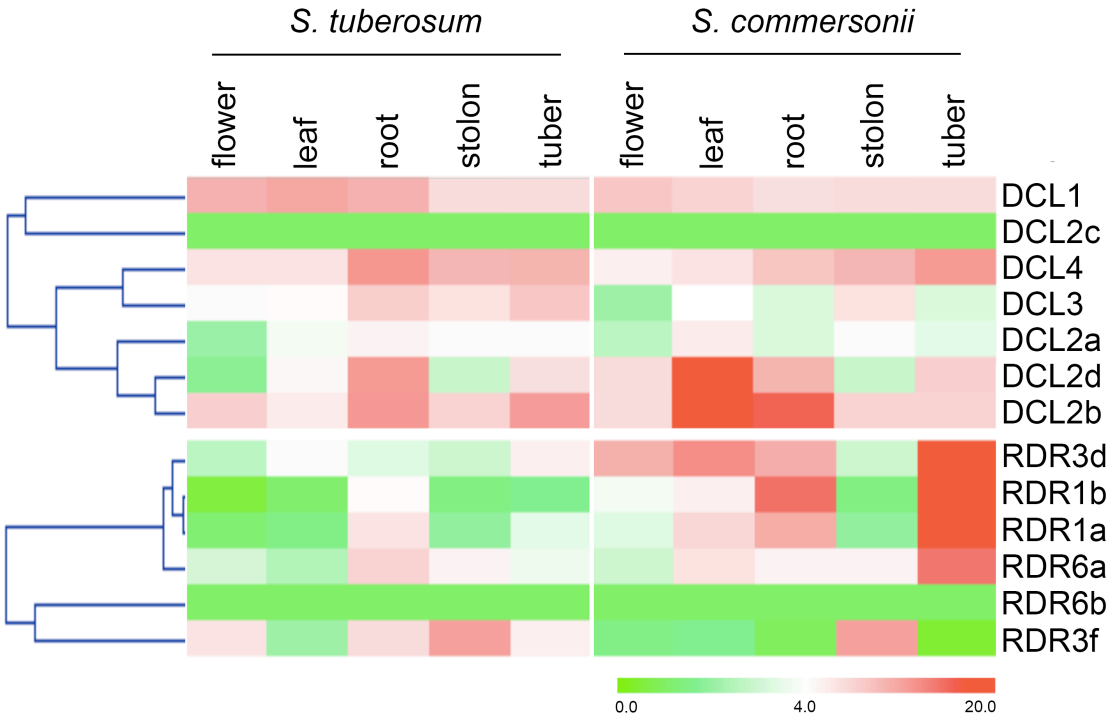


**Fig. 3** Heatmap showing the expression pattern of *DCL* and *RDR* genes in five tissues of cultivated *Solanum tuberosum* and wild *S. commersonii*. The color scales fold change values are shown at the bottom right of the figure

In a second analysis, we performed a RT-qPCR analysis to monitor the expression *DCL* and *RDR* genes after cold stress and PVY infection (Fig. 4). As far as cold stress is concerned, in NACC experiment all *DCL* genes were differentially expressed, with $\log_2$FC varying from -0.3 to -2 in stressed plants. The only exception was *DCL2c*. In detail, *DCL1* and *DCL4* were down-regulated in both *S. commersonii* and *S. tuberosum*. *DCL2b* was up-regulated in *S. commersonii* but it did not change its expression in *S. tuberosum*. *DCL3*, instead, was up-regulated in *S. tuberosum* but down-regulated in *S. commersonii*. *DCL2a/d* were activated in *S. commersonii* but repressed in *S. tuberosum*. Following ACC experiment, five of seven *DCLs* were down-regulated in *S. commersonii*, while only *ScDCL2d* was activated ($\log_2$FC = 4.3). On the other hand, in *S. tuberosum DCL1* and *DCL3* were weakly up-regulated while *DCL2a/b/d* and *DCL4* were down-regulated. As for virus infection experiments, qRT-PCR showed significant differences in the expression of *DCL* genes between control and stressed plants in both species (Fig. 4). For example, *DCL2a* and *DCL4* were up-regulated in both *S. commersonii* and *S. tuberosum* in comparison with control plants, whereas *DCL1* and *DCL2b* were up-regulated in *S. commersonii* and *DCL3* and *DCL2d* were up-regulated only in *S. tuberosum*. Regarding *RDRs* expression under cold and virus stresses, the preliminary survey through semi-quantitative RT-PCR revealed no

significant changes (data not shown). Therefore, further analyses through RT-qPCR were not carried out.

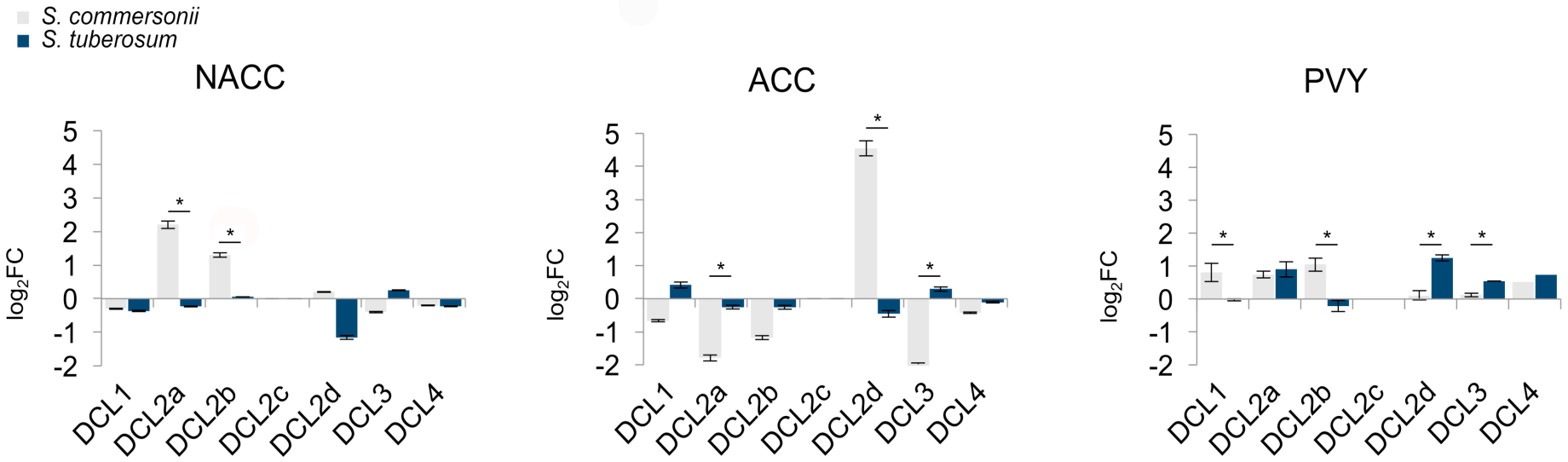


**Fig. 4** Differential analysis of *DCL* and *RDR* expression during cold stress in not acclimated (NACC) and acclimated (ACC) conditions and following virus infection in *Solanum tuberosum* (blue) and *S. commersonii* (grey). The values are reported as $log_2$FC. *EF* was used as endogenous gene. * Different at $P < 0.05$ (Student's t-test).

## 4.4. Discussion

Plants require elaborate mechanisms to produce specific regulators of gene expression, namely miRNA and siRNA. The molecular process leading to their formation is regulated by several proteins. Among them, DCLs and RDRs are key components trigging the gene silencing pathway. We have carried out an exhaustive analysis of *DCL* and *RDR* genes of *S. tuberosum* and *S. commersonii* genomes to add a contribution into understanding their molecular diversification.

### 4.4.1 Dicer-like genes in *S. tuberosum* and *S. commersonii*

Data obtained highlighted that the *S. tuberosum* and *S. commersonii* genomes do not differ in the degree to which *DCL* genes remain on corresponding chromosomes (synteny) and in corresponding orders (collinearity). Indeed, we identified seven *DCLs* both in *S. tuberosum* and *S. commersonii*, implying that the *DCL* loci have been substantially conserved between the two species after their evolutionary divergence. Conservation of *DCL* number and localization was observed also in respect to *S. lycopersicum* and *S. pennellii*, suggesting that the seven *DCL* genes come from events involving a single common ancestor. With respect to Solanums studied here, the number of *DCL* paralogs described in *Arabidopsis*, cucumber, maize and grapevine (Wassenegger et al. 2006; Qian et al. 2011; Zhao et al. 2015; Gan et al. 2017) is lower. By contrast, in rapeseed, rice, and foxtail millet a higher number of orthologs was identified (Kapoor et al. 2008; Yadav et al. 2015; Cao et al. 2016). In as much reciprocal homolog loss/gain in different lineages can affect a species' biology, degree of evolvability, and adaptability to changing environments (Zong et al. 2009), it would be interesting to study whether the *DCL* paralogs retained or lost in the *Solanaceae* can be adaptive and thus relevant for the evolution of this family. We have also found that the *DCL2* clade underwent an expansion in *S. commersonii* and *S. tuberosum*, since four paralogous copies were identified. Such duplication may suggest that this clade could provide specific and indispensable functions (Deleris et al. 2006). Since in potato the role of *DCL2* is still unclear, it is difficult to infer such specific function. However, in *Arabidopsis*, secondary *DCL2*-dependent 22-nt siRNAs are involved in viral targeting activities as well as to promote in cell-to-cell spread of VIGS (Garcia-Ruiz et al. 2010; Qin et al.2017). Given that, it is possible that the peculiar duplications of *DCL2* clade within the *Solanaceae* served to expand *DCL2* functions in mediating systemic silencing effects, perhaps through mechanisms of neo-functionalization and sub-functionalization (Panchy et al. 2016). In support of this hypothesis, we observed differences in gene and protein structures among DCL2 homologs, including important regulatory sequences. Rearrangement of paralogs at protein level leading to functional divergences has been observed also for Receptor-Like protein Kinases (RLK) and genes involved in secondary metabolism such as *TPS4* and *TPS5* (Feuilletet al.

2001; Ober 2005), possibly as consequence of adaptation to environment. Comparative-functional studies might further illuminate the details on how evolution has shaped *DCL2* function in *Solanaceae*. Therefore, we explored whether *DCL* paralogs show differences in their expression patterns both in various tissues and under stress. Our transcriptional analyses revealed that out of seven genes identified, six were expressed at different levels in one or more tissues analyzed. In particular*DCL2a*, *DCL2d* and *DCL3* had different patterns in *S. tuberosum* vs. *S. commersonii*. Also in respect to stress response, *DCL2* genes showed a different regulation between the cultivated and the wild species. Specifically, at low temperatures, *DCL2* genes were generally up-regulated in cold-tolerant *S. commersonii*, but almost unresponsive in cold-susceptible *S. tuberosum*. In addition, we showed that depending on the cold stress imposed (with or without acclimation) different *DLC2* copies were activated. A contrasting expression of the two *DCL2* paralogs was also observed in *S. commersonii* vs. *S. tuberosum* under PVY infection. Since both species are susceptible to this virus (Carputo et al. 2013), this may suggest a different evolutive diversification in virus-mediating response. To explain the divergent dynamic of *DCL2s* gene expression at least two hypotheses can be made. First, *S. tuberosum* and *S. commersonii* may have a different ability to program cellular transcriptional responses following stresses. Such different ability has been reported by D'Amelia et al. (2017), who found a divergent paralog expression after cold stress for *AN2*, a R2R3 MYB transcription factor which is induced by low temperatures in *S. commersonii* but not in *S. tuberosum*. Second, the polyploid nature of the cultivated potato might explain the diversification in gene expression between the two species. Indeed, it is becoming increasingly clear that the rewiring of the regulatory network following whole genome duplications is more important than functional divergence of the coding regions of individual genes (Aversano et al. 2012; De Smet and Van de Peer 2012; Chen et al. 2011; Osborn et al. 2003,).

### 4.4.2 RDR like-genes in *S. tuberosum* and *S. commersonii*

*RDRs* complement has been previously analyzed in cultivated tomato and potato (Hunter et al. 2016; Bai et al. 2012), but no information is available on non-model *Solanum* species. For this reason, we explored the extent of the potato *RDR* gene family also in *S. commersonii*. Our results provided evidence that this wild species harbors a total of eight *RDR* genes, including *RDR2*, which has never been described in *S. tuberosum*. However, the number of *S. tuberosum* and *S. commersonii RDRs* can be considered equal since expression data by Hunter et al. (2016) hinted the presence of *RDR2* also in the cultivated potato. We expect that future improvements of the current PGSC DM v4.03 pseudomolecule annotation will lead to its identification also in *S. tuberosum*. To our surprise, we discovered two new *RDRs* (*SlRDR3f* and *SlRDR6b*) homologs in the cultivated

tomato, leading to increase the number of this gene family from six to eight in *S. lycopersicum*. The identification of new genes can be explained by the different version of the tomato genome annotation we used (ITAG v3.2), which allowed us to identify *RDR* loci more accurately than Bai and colleagues (2012), who used the v2.3. It can be hypothesized that, along with cucumber (Gan et al. 2017) and foxtail millet (Yadav et al. 2015), the complement of *RDR* genes in *Solanaceae* is one of the most represented in plants. Indeed, six *RDRs* were identified in *Arabidopsis* (Wassenegger et al. 2006), five in rice and maize (Kapoor et al. 2008; Qian et al. 2011), seven in sorghum and soybean (Liu et al. 2014). Two copies of *RDR1* were found in both potato species. By contrast, only one copy in *S. lycopersicum* and *S. pennellii*. This potato-specific duplication may have derived after the divergence of potato and tomato from the common ancestors occurred 7,3 Mya (Tomato Sequence Consortium, 2012). The variable number of *RDR1* genes in different *Solanum* species is intriguing, in particular because this variability has not been found for the other *RDRs*. It is known that *RDR1* is involved not only in basal defense responses, but more in general in induced resistance mechanisms, such as the systemic acquired resistance (Yu et al. 2003; Muangsan et al. 2004). Since potato and tomato had different life histories after their divergence (TGSC 2012), the nature of selection pressure imposed by their environmental conditions might explain the *RDR1* clade variability, as already proposed for gene families involved in plant stresses (Hanada et al. 2008).

Our transcriptional study provided evidence that no differences in *RDR* gene expression after cold stress occurred in either species. This is comparable to findings by Kapoor and colleagues (2008), who reported no *RDR* transcript accumulation in rice after 4°C treatment. In contrast to our observations, the induction of five RDR genes has been recently reported in *Cucumber* (Gan et al. 2017) following cold stress. To data, the role of *RDRs* under low temperatures stress has received scant attention and the contrasting data available are puzzling. Hence, further investigations on various species under different experimental conditions (e.g. temperatures, time points) may lead to a better understanding of *RDR*'s role under cold stress. No changes in transcript activity of *S. tuberosum* and *S. commersonii RDR* loci has been detected also following virus infection. This contrasts with a number of studies concerning the induction of *RDR* genes to prevent virus invasion in different species (Yu et al.2003; Donaire et al. 2008; Qi et al.2009; Bai et al. 2012). Recently, Hunter et al. (2016) reported that suppression of *StRDR1* gene expression did not increase the susceptibility of PVY-infected potatoes, suggesting that the two *RDR1* paralogs may have lost their roles in antiviral resistance. However, functional analyses on the other members of *RDR* clades need to be carried out to shed lights on the lack of expression changes we observed.

## 4. 5. Conclusions

The importance of *DCL* and *RDR* genes in potato is not well understood, and differences among wild and cultivated species have not been described yet. Here, we reported that both *S. tuberosum* and *S. commersonii* harbor seven *DCL*, whereas seven *RDRs* were found in the former and six in the latter. Phylogenetic analysis revealed that duplication events have contributed to increase the number of *DCL* and *RDR* paralogs in both species. However, their genomes do not differ in the degree to which *DCL* and *RDR* genes remain on corresponding chromosomes, suggesting that their loci have been substantially conserved after their evolutionary divergence. This work has also pointed out new insights into the evolution of these gene families in tomato and potato species. For instance, lineage-specific duplications and retentions in *RDR* and *DCL* families were present in potatoes and tomatoes. To clarify whether the *DCL* and *RDR* paralogs retained or lost in *Solanaceae* can be adaptive and thus relevant for the evolution of these gene families, further studies will be carried out. For this purpose, transgenic plants defective in different members of *DCL* and *RDR* will be produced.

**Acknowledgments**

This work was carried out with the project "Development of potato genetic resources for sustainable agriculture" (PORES) founded by the University of Naples Federico II. We thank Mr. Raffaele Garramone for technical assistance. No conflict of interest declared.

**Author Contributions**

S.E., D.C. and R.A. planned and designed the research. S.E. and C.V. performed experiments. D.C. and R.A. contributed reagents/materials/analysis tools. S.E., M.M., V.D.A., R.A. and D.C analyzed data. S.E., D.C. and R.A. wrote the manuscript with the input from all the authors.

## References

Aversano R, Contaldi F, Ercolano MR, et al (2015) The *Solanum commersonii* genome sequence provides insights into adaptation to stress conditions and genome evolution of wild potato relatives. Plant Cell 27:954-968. doi: 10.1105/tpc.114

Aversano R, Ercolano MR, Caruso I, Fasano C, Rosellini D, Carputo D (2012) Molecular tools for exploring polyploid genomes in plants. Int J Mol Sci 13:10316-10335. doi: 10.3390/ijms130810316

Bai M, Guo-Shun Y, Wen-Ting C, Zhen-Chuan M, Hou-Xiang K, Guo-Hua C, Yu-Hong Y, Bing-Yan X (2012) Genome-wide identification of *Dicer-like*, *Argonaute* and *RNA-dependent RNA polymerase* gene families and their expression analyses in response to viral infection and abiotic stresses in *Solanum lycopersicum*. Gene 501:52-62. doi: 10.1016/j.gene.2012.02.009

Barone M, Carrieri R, Cardamone F, Ragozzino A, Alioto (2009) Potato tuber necrotic ringspot disease: severe outbreaks in Calabria. J Plant Pathol 91:S4.48.

Besser K, Harper A, Welsby N, Schauvinhold I, Slocombe S, Li Y, Dixon RA, Broun P (2009) Divergent regulation of terpenoid metabolism in the trichomes of wild and cultivated tomato species. Plant Physiol 149:499–514. doi:10.1104/pp.108.126276

Bolger A, Scossa F, Bolger ME et al (2014) The genome of the stress-tolerant wild tomato species *Solanum pennellii*. Nat Genet 46:1034–1038. doi:10.1038/ng.3046

Cao JY, Xu YP, Li W, Li SS, Rahman H, Cai XZ (2016) Genome-wide Identification of *Dicer-Like*, *Argonaute*, and *RNA-Dependent RNA Polymerase* gene families in *Brassica* species and functional analyses of their *Arabidopsis* homologs in resistance to *Sclerotinia sclerotiorum*. Front Plant Sci 7:1-17. doi:10.3389/fpls.2016.01614

Carputo D, Alioto D, Aversano R, Garramone R, Miraglia V, Villano C, Frusciante L (2013) Genetic diversity among potato species as revealed by phenotypic resistances and SSR markers. Plant Genet Resour-C 11:131–139. doi:10.1017/S1479262112000500

Carputo D, Castaldi L, Caruso I, Aversano R, Monti L, Frusciante L (2007) Resistance to frost and tuber soft rot in near-pentaploid *Solanum tuberosum* - *S. commersonii* hybrids. Breeding Sci 57:145–151.

Chen TW, Wu TH, Ng WV, Lin W-C (2011) Interrogation of alternative splicing events in duplicated genes during evolution. BMC Genomics 12 (supplement 3): S16. doi: 10.1186/1471-

2164-12-S3-S16

Crisp PA, Ganguly D, Eichten SR, Borevitz JO, Pogson BJ (2016) Reconsidering plant memory: intersections between stress recovery, RNA turnover, and epigenetics. Sci Adv 2:1-14. doi: 10.1126/sciadv.1501340.

D'Amelia V, Aversano R, Batelli G et al (2014) High *AN1* variability and interaction with basic helix-loop-helix co-factors related to anthocyanin biosynthesis in potato leaves. Plant J 80: 527–540. doi:10.1111/tpj.12653.

D'Amelia V, Aversano R, Ruggiero A, Batelli G, Appelhagen I, Dinacci C, Hill L, Martin C, Carputo D (2017) Sub-functionalization of duplicate *MYB* genes in *Solanum commersonii* generated the cold-induced *ScAN2* and the anthocyanin regulator *ScAN1*. Plant Cell Environ, in press. doi:10.1111/pce.12966.

De Smet R, Van De Peer Y (2012) Redundancy and rewiring of genetic networks following genome-wide duplication events. Curr Opin Plant Biol 15:168–176. doi: 10.1016/j.pbi.2012.01.003

Deleris A, Gallego-Bartolome J, Bao J, et al (2006) Hierarchical action and inhibition of plant Dicer-like proteins in antiviral defense. Science 313:68–71 doi:10.1126/science.1128214

Donaire L, Barajas D, Martinez-Garcia B, Martinez-Priego L, Pagan I, Llave C (2008) Structural and genetic requirements for the biogenesis of tobacco rattle virus-derived small interfering RNAs. J Virol 11:5167–5177. doi: 10.1128/JVI.00272-08

Feuillet C, Penger A, Gellner K, Mast A, Keller B (2001) Molecular evolution of receptor-like kinase genes in hexaploid wheat. Independent evolution of orthologs after polyploidization and mechanisms of local rearrangements at paralogous loci. Plant Physiol 125**:**1304–1313

Punta M, Coggill PC, Eberhardt RY, et al (2012) The Pfam protein families database. Nucleic Acids Res 1–12

Gan D, Zhan M, Yang F, Zhang Q, Hu K, Xu W, Lu Q, Zhang L, Liang D (2017) Expression analysis of *argonaute*, *Dicer-like*, and *RNA-dependent RNA polymerase* genes in cucumber (*Cucumis sativus* L.) in response to abiotic stress. J Gen 96:235–249. doi:10.1007/s12041-017-0758-y

Garcia-Ruiz H, Takeda A, Chapman EJ, Sullivan CM, Fahlgren N, Brempelis KJ, Carrington JC (2010) *Arabidopsis* RNA-dependent RNA polymerases and Dicer-like proteins in antiviral defense and small interfering RNA biogenesis during turnip mosaic virus infection. Plant Cell 22:481–496. doi: 10.1105/tpc.109.073056

Ghorecha V, Patel K, Ingle S, Sunkar R, Krishnayya NSR (2014) Analysis of biochemical variations and microRNA expression in wild (*Ipomoea campanulata*) and cultivated (*Jacquemontia pentantha*) species exposed to *in vivo* water stress. Physiol Mol Biol Plants 20:57–67. doi: 10.1007/s12298-013-0207-1

Goyer A, Hamlin L, Crosslin J, Buchanan A, Chang J (2015) RNA-Seq analysis of resistant and susceptible potato varieties during the early stages of potato virus Y infection. BMC Genomics 16:1–13. doi: 10.1186/s12864-015-1666-2

Hanada K, Zou C, Lehti-Shiu MD, Shinozaki K, Shiu S-H (2008) Importance of lineage-specific expansion of plant tandem duplicates in the adaptive response to environmental stimuli. Plant Physiol 148:993-1003. doi:10.1104/pp.108.122457

Hirsch CD, Hamilton JP, Childs KL, Cepela J, Crisovan E, Vaillancourt B, Hirsch CN, Habermann M, Neal B, Buell CR (2014) Spud DB: a resource for mining sequences, genotypes, and phenotypes to accelerate potato breeding. Plant Genome-US 7:1–12. doi: 10.3835/plantgenome2013.12.0042

Hunter LJR, Brockington SF, Murphy AM, Pate AE, Gruden K, MacFarlane SA, Palukaitis JPC (2016) *RNA-dependent RNA polymerase1* in potato (*Solanum tuberosum*) and its relationship to other plant *RNA-dependent RNA polymerases*. Sci Rep 6:23082. doi: 10.1038/srep23082

Kapoor M, Arora R, Lama T, Nighawan A, Khurana JP, Tyagi AK, Kapoor S (2008) Genome-wide identification, organization and phylogenetic analysis of *Dicer-like*, *Argonaute* and *RNA-dependent RNA polymerase* gene families and their expression analysis during reproductive development and stress in rice. BMC Genomics 9:451. doi: 10.1186/1471-2164-9-451

Kraiwesh B, Zhu JK, Zhu J (2012) Role of miRNAs and siRNAs in biotic and abiotic stress responses in plants. Biochim Biophys Acta 1819:137–148

Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357-359. doi: 10.1038/nmeth.1923

Larkin MA, Blackshields G, Brown NP et al (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23:2947–2948

Larkin, M. A., Blackshields, G., Brown, N. P.,et al. (2007) Clustal Lescot M, Dehais P, Thijs G, Marchal K, Moreau Y, van de Peer Y, Rouze P, Rombauts S (2002) PlantCARE, a database of plant *cis*-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Res 30:325-327

Livak K.J. & Schmittgen T.D. (2001) Analysis of relative gene expression datLivak KJ and Schmittgen TD (2001) Analysis of relative gene expression data using real‐time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. Methods 25:402–408

Liu X, Lu T, Dou Y, Yu B, Zhang C (2014) Identification of RNA silencing components in soybean and sorghum. BMC Bioinformatics 15:1-13. doi: 10.1186/1471-2105-15-4

Margis R, Fusaro AF, Smith NA, Curtin SJ, Watson JM, Finnegan EJ, Waterhouse PM (2006) The evolution and diversification of *Dicers* in plants. FEBS Lett 10:2442–2450

Matzke MA, Kanno T, Matzke AJM (2015) RNA-Directed DNA Methylation: The evolution of a complex epigenetic pathway in flowering plants. Annu Rev Plant Biol 66:243–267. doi:10.1146/annurev-arplant-043014-114633

Nicot N, Hausman JF, Hoffmann L, Evers D (2005) Housekeeping gene selection for real-time RT-PCR normalization in potato during biotic and abiotic stress. J Exp Bot 56:2907–2914

Ober D (2005) Seeing double: gene duplication and diversification in plant secondary metabolism. Trends Plant Sci 10:445–449

Osborn TC, Butrulle DV, Sharpe AG, Pickering KJ, Parkin IAP, Parker JS, Lydiate DJ (2003) Detection and effects of a homoeologous reciprocal transposition in *Brassica napus*. Genetics 165:1569-1577

Panchy N, Lehti-Shiu MD, Shiu SH (2016) Evolution of gene duplication in plants. Plant Physiol 171:2294–2316. doi:10.1104/pp.16.00523

Peláez P, Sanchez F (2013) Small RNAs in plant defense responses during viral and bacterial interactions: similarities and differences. Front Plant Sci 4:1-16

Pontier D, Yahubyan G, Vega D, Bulski A, Saez-Vasquez J, Hakimi MA, Lerbs-Mache S, Colot V, Lagrange T (2005) Reinforcement of silencing at transposons and highly repeated sequences requires the concerted action of two distinct RNA polymerases IV in *Arabidopsis*. Gene Dev 19:2030–2040

Potato Genome Sequencing Consortium (2011) Genome sequence and analysis of the tuber crop potato. Nature 475:189-95. doi: 10.1038/nature10158

Qi Y, Denli AM, Hannon GJ (2005) Biochemical specialization within *Arabidopsis* RNA silencing pathways. Mol Cell 3:421–428

Qian Y, Cheng Y, Cheng X, Jiang H, Zhu S, Cheng B (2011) Identification and characterization of *Dicer-like*, *Argonaute* and *RNA-dependent RNA polymerase* gene families in maize. Plant Cell Rep 30:1347–1363

Qin C, Li B, Fan Y et al (2017) Roles of Dicer-Like Proteins 2 and 4 in intra- and intercellular antiviral silencing. Plant Physiol 174:1067-1081. doi: 10.1104/pp.17.00475

Shabalina SA, Koonin EV (2008) Origins and evolution of eukaryotic RNA interference. Trends Ecol Evol 23:578–587

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol 8:1596–1599

The Tomato Genome Consortium (TGC) (2012) The tomato genome sequence provides insights into fleshy fruit evolution. Nature 485:635-641. doi:10.1038/nature11119

Tran RK, Zilberman D, de Bustos C et al (2005) Chromatin and siRNA pathways cooperate to maintain DNA methylation of small transposable elements in *Arabidopsis*. Genome Biol 6:1-11

Wassenegger M, Krczal G (2006) Nomenclature and functions of *RNA-directed RNA polymerases*. Trends Plant Sci 11:142–151

Wheelan SJ, Church DM, Ostell JM (2001) Spidey: a tool for mRNA-to-genomic alignments. Genome Res 11:1952–1957

Xie Z, Johansen LK, Gustafson AM, Kasschau KD, Lellis AD, Zilberman, D, Jacobsen SE, Carrington JC (2005) Genetic and functional diversification of small RNA pathways in plants. Plos

Biol 2:0642-0652

Yadav CB, Muthamilarasan M., Pandey G, Prasad M (2015) Identification, characterization and expression profiling of *Dicer-Like*, *Argonaute* and *RNA-dependent RNA polymerase* gene families in foxtail millet. Plant Mol Biol Rep 33:43–55

Yoo MJ, Liu X, Pires JC, Soltis PS, Soltis DE (2014) Non additive gene expression in polyploids. Annu Rev Genet 48:485–517

Yoshikawa M, Peragine A, Park MY, Poethig RS (2005) A pathway for the biogenesis of trans-acting siRNAs in *Arabidopsis*. Gene Dev 19:2164-2175

Yu D, Fan B, MacFarlane, SA, Chen Z (2003) Analysis of the involvement of an inducible *Arabidopsis RNA-dependent RNA polymerase* in antiviral defense. Mol Plant Microbe In 16:206–216

Zhang D, Ma H (2009) Evolution of the *RNA-dependent RNA polymerase* (*RdRP*) genes: duplications and possible losses before and after the divergence of major eukaryotic groups. Gene 447:29–39

Zhao HL, Zhao K, Wang J, Chen X, Chen Z, Cai RH, Xiang Y (2015) Comprehensive analysis of *Dicer-like*, *Argonaute*, and *RNA-dependent RNA polymerase* gene families in grapevine (*Vitis vinifera*). J Plant Growth Regul 34:108–121

# Chapter V. General conclusions

**Conclusions**

It has been estimated that, with climate changes, plants will be exposed to more unexpected events, ranging from abiotic to biotic stress conditions. These will impact crop production, quality and geographical distribution. Among the abiotic stresses, cold is one of the suboptimal conditions that are more harmful to the crop that is being the subject of several studies at the Department of Agricultural Sciences: the potato. Indeed, *S. tuberosum* is classified as frost-sensitive and incapable of cold acclimate. By contrast, some wild potato species are much more frost hardy than *S. tuberosum* and capable of cold acclimation. Thus, they are a potential genetic resource for introgressing freezing tolerance into cultivated varieties. It has been estimated that only 10% of these species has been used to develop new improved potato varieties (Gavrilenko, 2011). This is due to the fact that there are many difficulties in their introduction and use in breeding programs, which often require long and laborious methods for obtaining new superior interspecific hybrids. A particularly negative aspect is related to linkage drag, often resulting from the lack of both effective selection criteria and genomic tools to assist breeders. For this reason traditional breeding methods for improvement of freezing resistance in potato plants have achieved only limited success. With all these thoughts in mind, to broaden the genetic base of the cultivated potato and to transfer interesting genes from wild species, in the scientific community it is increasingly felt the need to develop new data based on structural and functional genomics, allowing a more efficient exploitation of the wide genetic background of wild potatoes, and therefore, to improve the current strategies of potato breeding and selection.

Among the wild potato species, *S. commersonii* is the one displaying the highest tolerance to low temperatures. It can survive to about -5 °C and to as low as -11 °C after acclimation (Costa & Li 1993). It also represents the first potato relative whose genome sequence has been deciphered (Aversano et al. 2015). Hence, *S. commersonii* is potentially the best model species to study the molecular mechanisms involved in cold tolerance and, consequently, to deploy for breeding pouropses. Until today most studies were focused on specific metabolic pathways involved in freezing tolerance. One example of this study is that reported by Palta et al. (1993), where the authors determined differentiation of plasma membrane lipid changes associated with increased freezing tolerance following acclimation. The molecular basis of tolerance per se is also poorly understood, although it has been reported that it may be genetically determined by loci independent of acclimated tolerance not only in potato but also in other crops.

Today, with the availability of genome sequence of *S. commersonii*, further studies have been carried out. In fact, Aversano et al. (2015) reported a comparison of cold-responsive gene expression profiles between not acclimated and acclimated stressed plants and highlighted remarkable features of cold-responsive genes known to be critical in cold sensing and signaling pathways. For example, their results shed new lights on the role of CBF genes, showing that all ScCBFs were up regulated under all tested conditions relative to controls. These findings were in contrast with previous reports that CBF1, but not related CBFs, were responsive to low temperatures in both *S. commersonii* and *S. tuberosum* (Pennycooke et al. 2008; Carvallo et al. 2011). In our work we obtained new and interestingly result that extend the knowledge on the genetic regulation of the cold stress signaling and tolerance mechanisms. Thanks to the annotation of *S. commersonii* genes it was possible, for the first time, to identify a complete set of putative cold responsive genes in this species using a RNA-seq approach. Particularly useful has been the availability of two different clones of *S. commersonii* contrasting in their cold tolerance. Overall, the whole-transcriptomic expression data highlighted an extensive reorganization of the transcriptome under cold stress, with enhanced expression of genes involved in signal perception, transduction and genes involved in cell repair (such as heat-shock proteins, HSPs and dehydrins, DHNs). One notable observation was that several genes were responsive to cold when compared to control conditions, but with contrasting kinetics under different conditions. Our data showed that CBF3, CBF4 and ZAT12 were up-regulated under not acclimated conditions in the cold tolerant clone, but they were down-regulated in the cold susceptible. By contrast, they were all up-regulated under acclimated conditions in both clones. Our findings are in contrast with previous reports that revealed how CBF1, but not related CBFs, were responsive to low temperatures in both *S. commersonii* and *S. tuberosum* (Pennycooke et al. 2008; Carvallo et al. 2011) but our observations were similar to the results showed by Aversano et al. (2015), that highlighted the induction of all CBF genes under not-acclimated and acclimated conditions. Our data are also consistent to those observed in tomato species, where three CBF genes were cold responsive in the cold-tolerant *S. peruvianum* (Mboup et al. 2012). The different expression of CBF genes in our plant material may be directly responsible for enhanced cold tolerance and acclimation ability in this species. Therefore, probably CBF3 and CBF4 are the first members that play an important role in the early cold response before CBF1 and CBF2. Cold tolerance and cold acclimation in potato, as well as in other species such as rice, cabbage, wheat, and tea is based on quantitative inheritance, since there are many genes and interactions (possibly epistatic mechanisms) involved in defining the different tolerance levels among genotypes. Thus, one could expect that different metabolic routes and different genes are modulated in response to cold. Our data demonstrated also that transcription

factors including HB, bZIP, MYB, MYC, WRKY and bHLH play important roles in cold stress. Previous studies showed that the overexpression of MlNAC5 (*Miscanthus lutarioriparius*) and SlNAC1 (*Suaeda liaotungensis*) enhanced drought and cold stress tolerance of *Arabidopsis*, respectively (Li et al. 2013; Zong et al. 2016). In our study, more than 20 different TF families were identified to respond to cold in *S. commersonii*. Among them AP2/ERF, bHLH and Zn-finger TFs genes were the most abundant. These TF families interact to regulate target genes during plant stress responses (D'Amelia et al. 2017).

Although the molecular mechanisms on how different species respond to cold stress remains to be elucidated, recent studies have shown that abiotic stress induces aberrant expression of many small non-coding RNA (sncRNA) in several plant species. For this reason, in the third chapter of the thesis we provide, for the first time, a comprehensive analysis of smRNA population in *S. commersonii*. This analysis gave us the opportunity to correlate miRNA expression with RNAseq data of cold stressed plants. Although miRNAs and their targets have already been identified in *S. tuberosum*, their participation during cold stress in potato remains unknown. We believe that this work will contribute to extend the information available on the genomic structure of *S. commersonii* and all the data obtained will provide additional genomic tools for an efficient exploitation of the cold resistance traits in this species. Overall, our research allowed to annotate in *S. commersonii* genome 273 distinct miRNAs. The majority appeared to mature from transcripts of intergenic regions or from introns of annotated genes, which is in line with previous findings. Most of them were *S. commersonii*-specific, underlying the need to improve the genomic data of *S. commersonii* in order to provide as much as possible information on the gene and miRNA specificity of this wild potato. Small RNA sequencing data revealed also regulatory roles of miRNAs during cold stress. Our results are consistent with Zhang (2009), who provided evidence that most of the miRNA involved in cold stress in *Brachipodium* were down-regulated. In other words, the dramatic change in miRNA expression levels in *S. commersonii* leaves causes a sensitive reaction of their targets, which regulate the expression of transcriptions factors, membrane receptors, genes involved in lignin biosynthesis and other correlated pathways. Further, the cross-response of miRNAs to multi-biotic and abiotic stresses indicated that *S. commersonii* has evolved sophisticated miRNA-mediated pathways to cope with changing environments. Such mechanisms may be explored in the future to efficiently breed for cold tolerance in potato.

The production of miRNA and their expression relies on the accurate functioning of Dicer-like (DCL), Argonaute (AGO) and RNA-dependent RNA polymerases (RDR) proteins, whose genes are present in multiple copies in eukaryotic genomes. DCL and RDR comprise the core components of RNA-induced silencing complexes, which trigger RNA silencing. These proteins are

prominent players in the post-transcriptional control of gene expression, as they control small RNA-mediated gene silencing pathways and function in the epigenetic regulation of the genome under various environmental stresses (Yadav et al. 2015). Since no molecular information was available on these gene families, we performed a genome-wide identification of *DCL* and *RDR* genes in *S. tuberosum* and *S. commersonii*. These results have been submitted to Planta. Our findings demonstrated that duplication events have contributed to increase the number of *DCL* and *RDR* paralogs in both species. However, *S. tuberosum* and *S. commersonii* genomes do not differ in the degree to which *DCL* genes remain on corresponding chromosomes and in corresponding orders, suggesting that their loci have been substantially conserved after their evolutionary divergence. To clarify whether the *DCL* and *RDR* paralogs retained or lost in *Solanaceae* can be adaptive and thus relevant for the evolution of these gene families, we also explored whether *DCL* paralogs showed differences in their expression patterns both in various tissues and under stress. Our transcriptional analyses revealed that out of seven genes identified, six were expressed at different levels in one or more tissues analyzed. Also in respect to stress response, *DCL2* genes showed a different regulation between the cultivated and the wild species. Specifically, at low temperatures, *DCL2* genes were generally up-regulated in cold-tolerant *S. commersonii*, but almost unresponsive in cold-susceptible *S. tuberosum*. In addition, we showed that depending on the cold stress imposed (with or without acclimation) different *DLC2* copies were activated.

We strongly believe that, combined with strategies including the efficient use of genome sequences, genome-wide association studies, mutation detection, gene discovery and regulation, and –omics databases, results reported here may represent a starting point for additional investigations and future breeding applications not only in potato but in several other crops.

**References**

Aversano R, Contaldi F, Ercolano MR, et al (2015) The *Solanum commersonii* genome sequence provides insights into adaptation to stress conditions and genome evolution of wild potato relatives. Plant Cell 27:954-968.

Carvallo MA, Pino MT, Jeknić Z, Zou C, Doherty CJ, Shiu SH, and Thomashow MF (2011) A comparison of the low temperature transcriptomes and CBF regulons of three plant species that differ in freezing tolerance: *Solanum commersonii*, *Solanum tuberosum*, and *Arabidopsis thaliana*. J of Exp Bot 62:3807–3819. doi:10.1093/jxb/err066

Costa A, and Li P, (1993) Development of cold hardiness in *Solanum tuberosum* by abscisic acid and mefluidide. In: Li P. and Christersson L. (eds.) Advance in Plant Cold Hardiness. CRC Press, Boca Raton FL USA 139–140.

D'Amelia V, Aversano R, Ruggiero A, Batelli G, Appelhagen I, Dinacci C, Hill L, Martin C, Carputo D (2017) Sub-functionalization of duplicate *MYB* genes in *Solanum commersonii* generated the cold-induced *ScAN2* and the anthocyanin regulator *ScAN1*. Plant Cell Environ in press. doi:10.1111/pce.12966.

Gavrilenko (2011) Application of molecular cytogenetics in fundamental and applied research of potato. In Genetics, genomics and breeding of potato. Edited by Bradeen J.M. and Kole C. CRC Press New York. pp. 184-206.

Li J, Wang N, Xin H, and Li S (2013) Overexpression of *VaCBF4*, a transcription factor from *Vitis amurensis*, improves cold tolerance accompanying increased resistance to drought and salinity in *Arabidopsis*. PlantMol Biol Report 31:1518–1528. doi: 10.1007/s11105-013-0627-7

Mboup M, Fischer I, Lainer H, Stephan W (2012) Trans-species polymorphism and allele-specific expression in the *CBF* gene family of wild tomatoes. Mol Biol Evol 29:3641–3652.

Palta JP and Simon G (1993) Breeding potential for improvement of freezing stress resistance: genetic separation of freezing tolerance, freezing avoidance, and capacity to cold acclimate. In Advances in Plant Cold Hardiness, P.H. Li, ed (CRC Press: Boca Raton, Florida, USA) 299–310

Pennycooke JC, Cheng HM, Roberts SM, Yang QF, Rhee SY, Stockinger EJ (2008) The low temperature-responsive, Solanum *CBF1* genes maintain high identity in their upstream regions in a genomic environment undergoing gene duplications, deletions, and rearrangements. Plant Mol Bio l67:483–497

Yadav CB, Muthamilarasan M, Pandey G, Prasad M (2015) Identification, characterization and expression profiling of *Dicer-Like*, *Argonaute* and *RNA-dependent RNA polymerase* gene families in foxtail millet. Plant Mol Biol Rep 33:43–55

Zhang J, Xu Y, Huan Q, & Chong K (2009) Deep sequencing of *Brachypodium* small RNAs at the global genome level identifies microRNAs involved in cold stress response. BMC Genomics, 10: 449. doi:10.1186/1471-2164-10-449

Zong JM, Li XW, Zhou YH, Wang FW, Wang N, Dong YY, et al. (2016) The *AaDREB1* transcription factor fromthe cold-tolerant plant *Adonis amurensis* enhances abiotic stress tolerance in Transgenic Plant Int J Mol Sci 17:611–626. doi: 10.3390/ijms17040611