COMPARATIVE PROTEOMIC ANALYSIS APPLIED FOR STUDY OF THE HETEROSIS AND ACQUISITION OF THE EMBRYOGENIC COMPETENCE IN PAPAYA

ELLEN DE MOURA VALE

UNIVERSIDADE ESTADUAL DO NORTE FLUMINENSE
DARCY RIBEIRO – UENF

CAMPOS DOS GOYTACAZES – RJ
Março - 2017

COMPARATIVE PROTEOMIC ANALYSIS APPLIED FOR STUDY OF THE HETEROSIS AND ACQUISITION OF THE EMBRYOGENIC COMPETENCE IN PAPAYA

ELLEN DE MOURA VALE

"Tese apresentada ao Centro de Ciências e Tecnologias Agropecuárias da Universidade Estadual do Norte Fluminense Darcy Ribeiro, como parte das exigências para obtenção do título de Doutor em Genética e Melhoramento de Plantas."

Orientador: Prof. Vanildo Silveira

CAMPOS DOS GOYTACAZES – RJ
MARÇO - 2017

COMPARATIVE PROTEOMIC ANALYSIS APPLIED FOR STUDY OF THE HETEROSIS AND ACQUISITION OF THE EMBRYOGENIC COMPETENCE IN PAPAYA

ELLEN DE MOURA VALE

"Tese apresentada ao Centro de Ciências e Tecnologias Agropecuárias da Universidade Estadual do Norte Fluminense Darcy Ribeiro, como parte das exigências para obtenção do título de Doutor em Genética e Melhoramento de Plantas."

Aprovada em 29 de março de 2017.

Comissão Examinadora:

Profª. Neusa Steiner (D.Sc.,Recursos Genéticos Vegetais) – UFSC

Prof. Messias Gonzaga Pereira (Ph.D. Melhoramento de Plantas) – UENF

Prof. Jurandi Gonçalves de Oliveira (D.Sc., Biologia Vegetal) – UENF

Prof. Vanildo Silveira (D.Sc., Biotecnologia) – UENF
(Orientador)

ACKNOWLEDGMENTS

The research and scholarships were supported by the Carlos Chagas Foundation for Research Support in the State of Rio de Janeiro (FAPERJ), the Coordination for the Improvement of Higher Education Personnel (CAPES), and the National Council for Scientific and Technological Development (CNPq).

I would like to thank the Universidade Estadual do Norte Fluminense Darcy Ribeiro and the Genetic Plant Breeding Graduate Program for the free quality education.

To Professor Vanildo Silveira, for guidance, confidence and patience.

To the teachers Messias Gonzaga Pereira and Claudete Santa Catarina for the collaboration and the councils in the development of this work.

To the group of vegetable biotechnology for the scientific support in the developed activities.

SUMMARY

RESUMO vi
ABSTRACT viii
1. INTRODUCTION 1
2. OBJECTIVES 3
2.1. General objective 3
2.2. Specific objectives 3
3. CHAPTERS 4
3.1. COMPARATIVE PROTEOMIC ANALYSIS OF HETEROSIS PHENOMENON IN PAPAYA ROOTS 4
3.1.1. INTRODUCTION 4
3.1.2. REVIEW 6
3.1.2.1. *Carica papaya* 6
3.1.2.2. Heterosis 8
3.1.2.3. Plant proteomics 11
3.1.2.4. Proteomics applied to plant breeding 12
3.1.3. MATERIALS AND METHODS 14
3.1.3.1. Plant Material 14
3.1.3.2. Protein extraction and quantification 14
3.1.3.3. Protein digestion 15
3.1.3.4. Mass spectrometry analysis 15
3.1.3.5. Protein identification and quantification 16

3.1.3.6. Statistical analysis 17
3.1.4. RESULTS 18
3.1.4.1. Heterosis effects in root growth 18
3.1.4.2. Proteomic profiles of the heterosis response 18
3.1.4.3. Non-additively accumulated proteins in seminal roots 19
3.1.4.4. Functional classification of all non-additively accumulated proteins in hybrid roots 20
3.1.5. Discussion 25
3.1.5.1. Energy-related proteins 26
3.1.5.2. Protein metabolism-related proteins 27
3.1.5.4. Reactive oxygen species (ROS) metabolic process-*related proteins* 29
3.1.6. CONCLUSION 31
3.2. DIFFERENTIALLY ABUNDANT PROTEINS ASSOCIATED WITH THE ACQUISITION OF THE EMBRYOGENIC COMPETENCE OF PAPAYA CALLUS 32
3.2.1. INTRODUCTION 32
3.2.2. REVIEW 34
3.2.2.1. Papaya somatic embryogenesis 34
3.2.2.2. Acquisition of embryogenic competence 36
3.2.3. Materials and methods 39
3.2.3.1. Somatic embryogenesis induction 39
3.2.3.3. Microscopy analysis 40
3.2.3.4. Protein extraction 41
3.2.3.5. Protein Digestion 41
3.2.3.6. Mass spectrometry analysis 42
3.2.3.7. Bioinformatics 43
3.2.4. RESULTS 44
3.2.4.1. Morphological and cellular characterization 44
3.2.4.2. Protein identification by LC–MS/MS 46
3.2.5. DISCUSSION 59
3.2.5.1. Stress and detoxification-related proteins 59
3.2.5.2. Response to Lipid-related proteins 61
3.2.5.3. Auxin regulation-related proteins 62
3.2.5.4. Gene regulation-related proteins 63

3.2.5.5. Proteins not associated with the selected functional groups 64
3.2.6. CONCLUSION 65
REFERENCES.....66

RESUMO

VALE, Ellen de Moura; D.Sc.; Universidade Estadual do Norte Fluminense Darcy Ribeiro; Março, 2017; Comparative proteomic analysis applied for study of the heterosis and acquisition of the embryogenic competence in papaya; Orientador: Prof. Dr. Vanildo Silveira; Conselheiros: Prof. Dr. Gonçalo Apolinário de Souza Filho e Prof. Dr. Messias Gonzaga Pereira.

Ferramentas biotecnológicas apresentam alto potencial de inserção em programas de melhoramento genético de várias espécies, inclusive no mamoeiro (*Carica papaya*L.). Neste contexto, a proteômica comparativa pode fornecer informações importantes sobre os mecanismos moleculares que governam os processos biológicos, como a heterose ea embriogênese somática. O objetivo desse trabalho foi desenvolver uma abordagem proteômica comparativa aplicada ao estudo do fenômeno da heterose e aquisição de competência morfogênica durante o amadurecimento de embriões somáticos de mamoeiro, visando a integração dessas abordagens nos programas de melhoramento desta espécie.No primeiro capítulo desta tese foram analisadas as proteínas das raízes primárias do híbrido JS12 × São Mateus do mamão e suas linhagens parentais puras utilizando análises proteômicas combinando ao método shotgun ea tecnologia nanoESI-HDMSE. Um total de 955 proteínas foram identificadas pelo método da shotgun, entre os quais 261 apresentaram abundância não aditiva, quando sua abundância diferiu estatisticamente da média dos pais.As proteínasnão aditivas foram divididas de acordo com o padrão de abundânciaem "acima do maior pai" (16,1%), "igual ao maior pai" (6,5%), "igual ao menor pai" (22,2%) e "menor que o menor pai" (55,2%).

Os resultados revelaram uma diminuição em proteínas envolvidas em processos de consumo de energia no híbrido, como o metabolismo de proteínas e um aumento de proteínas relacionados ao desenvolvimento de raiz, tais como aquelas envolvidas no transporte polar da auxina e regulação de sinalização. Os resultados sugerem que o híbrido possui um mecanismo de otimização para a síntese protéica que resulta em melhorias substanciais na eficiência energética celular e no desempenho fenotípico. Portanto, este estudo pode contribuir para uma melhor compreensão da base molecular da heterose no mamão. No segundo capítulo dessa tese foi investigada as proteínas de calos embriogênicos e não embriogênicos de mamoeiro. Para identificar proteínas específicas envolvidas na competência embriogênica, utilizou-se análise proteômica de culturas embriogênicas e não embriogênicas de mamoeiro. Foram identificadas 668 proteínas, das quais 157 foram significativamente abundantes diferencialmente. Comparando a abundância das proteínas de calo embriogênicas em relação ao calo não embriogênico, 83 foram up-regulados e 75 foram down-regulados. Observou-se um aumento na abundância de proteínas relacionadas à resposta ao estímulo, à resposta ao estresse, ao processo de desenvolvimento, à resposta ao hormônio, à regulação do processo biológico, à resposta a lípidos, à morfogênese celular e à resposta a espécies reativas de oxigênio nos calos embriogênicos. Entre as proteínas identificadas, a Auxin GH3 pode ser indispensável para manter a homeostase auxina e parece desempenhar um papel funcional na competência embriogênica e no desenvolvimento adicional do embrião. As espécies reativas de oxigênio (ROS) têm um efeito central na regulação do desenvolvimento, a família Glutathione S-transferase desempenhar um papel no mecanismo de defesa para proteger as plantas de danos celulares de espécies excessivamente reativas de oxigênio. A maior abundância de proteínas relacionadas com o metabolismo lipídico, como a bifunctional inhibitor lipid-transfer seed storage 2S superfamily albumin, parece ser fundamental para a aquisição da competência embriogênica do mamão. Estas proteínas identificadas mais abundantes em calos embriogênicos fornecem pistas para a compreensão dos processos que ocorrem na aquisição da competência embriogênica de calos de mamão.

ABSTRACT

VALE, Ellen de Moura; D.Sc.; Universidade Estadual do Norte Fluminense Darcy Ribeiro; Março, 2017; Comparative proteomic analysis applied for study of the heterosis and acquisition of the embryogenic competence in papaya; Advisor: Prof. Dr. Vanildo Silveira; Committee members: Prof. Dr. Gonçalo Apolinário de Souza Filho and Prof. Dr. Messias Gonzaga Pereira.

Biotechnological tools present high insertion potential in breeding programs of several species, including papaya (*Carica papaya* L.). In this context, comparative proteomics can provide important information about the molecular mechanisms that govern biological processes, such as heterosis and somatic embryogenesis. The objective of this work was to develop a comparative proteomic approach applied to the study of the heterosis phenomenon and acquisition of morphogenic competence during the maturation of papaya somatic embryos, aiming the integration of these approaches in breeding programs of this species.In the first chapter of this thesis were analyzed the proteins of the primary roots of the hybrid JS12 × São Mateus and its parental inbred lines were analyzed using proteomic analyses combining the shotgun method and nanoESI-HDMSE technology.A total of 955 proteins were identified by the shotgun method, among which 261 exhibited non-additive expression, when their expression differed statistically from themid-parents. Non-additive proteins were divided into “above high-parent” (16.1%), “high-parent”(6.5%), “low-parent” (22.2%), and “below low-parent” (55.2%) abundance patterns. The results revealed adecrease in proteins involved in energy-consuming processes in hybrid such as protein metabolism and an increasein root development

proteins such as those involved in auxin polar transport and signaling regulation. The findings suggest that the hybrid possesses an optimization mechanism for protein synthesis that results in substantial improvements in cellular energy efficiency and phenotypic performance. Therefore, this study may contribute to a better understanding of the molecular basis of heterosis in papaya. In the second chapter of this thesis it was investigated proteins of embryogenic callus and non-embryogenic of papaya. To identify specific proteins involved in embryogenesis competence was used proteomic analyses combining the shotgun method and nanoESI-HDMSE technology of callus embryogenic and non-embryogenic of *C. papaya*. Total of 668 proteins were identified, among which 157 proteins were significantly differentially abundant. Comparing the abundance of the embryogenic callus proteins in relation to the non-embryogenic callus, 83 were up-regulated and 75 were down-regulated. An increase in the abundance of proteins related to response to stimulus, response to stress, developmental process, response to hormone, regulation of biological process, response to lipid, cell morphogenesis and response to reactive oxygen species was observed in embryogenic callus. Among these identified proteins, Auxin GH3 may be indispensable maintaining homeostasis auxin and play a functional role in embryogenic competence and further embryo development. Reactive oxygen species (ROS) have a central effect in development regulation, Glutathione S-transferase family play a role in the defense mechanism to protect plants from cell damage from excessive reactive oxygen species. The greater abundance of proteins related to lipid metabolism such as bifunctional inhibitor lipid-transfer seed storage 2S superfamily albumin seems to be fundamental for the acquisition of the embryogenic competence of papaya. These identified more abundant proteins in embryogenic callus provide clues to understanding of the processes that occur in the acquisition of the embryogenic competence of papaya callus.

1. INTRODUCTION

The species *Carica papaya* L., belonging to the family Caricaceae, is a fruit of great economic importance cultivated mainly in tropical countries and with great acceptance in the world market. Brazil is currently the second largest producer of fruit (FAO 2016).

Food production uses phenomena known for decades, such as heterosis and somatic embryogenesis, however, its use is often carried out empirically (Duncan 2011; Cooper et al. 2014; Zhao et al. 2015b). Molecular analyzes can contribute to the understanding of these phenomena in order to contribute with new analysis tools and in the development of optimized protocols.

Among them, the heterosis,isa phenomenon that has been used for more than a hundred years and is responsible for the superior performance of F_1 hybrids in relation to their parental lines (Falconer and Mackay 1996). Additionally, somatic embryogenesis is a biotechnological tool for efficient propagation technique that reflects the totipotency of plant cell (Kaplan and Cooke 1997).

Despite great advances in genomics it was on the understanding of the mechanisms responsible for the expression of characteristics (phenotype) of agronomic interest, this isolated tool may not be enough to answer all the questions about the transmission of genetic information.

In this way the proteomic analysis in plants has been emerging as an extremely useful tool in plant breeding (Eldakak et al. 2013). Since it reflects the variability that is actually expressed by genes (Pennington and Dunn 2001).

Proteomic analysis is responsible for several studies that bring a new light about heterosis (Marcon et al. 2010; Marcon et al. 2013; Mohayeji et al. 2014) and somatic embryogenesis of several species (Sun et al. 2013; Varhaníková et al. 2014; Heringer et al. 2015; dos Santos et al. 2016), including papaya (Vale et al. 2014). According to (Cramer et al. 2013) in a special edition on proteomics in plant breeding, proteomics is a fundamental research tool in the development of new technologies to promote agricultural and its sustainability. In the same way proteomic studies are an important tool for the development of new strategies for the papaya plant breeding.

2. OBJECTIVES

2.1. General objective

To develop a comparative proteomic approach applied to the study of the heterosis phenomenon and the acquisition of morphogenic competence during the multiplicação of somatic embryos of papaya (*Carica papaya* L.), aiming the integration of these approaches in the breeding programs of this species.

2.2. Specific objectives

- To identify and quantify the differentially abundant proteins of primary roots of hybrids and their parental lines in order to identify non-additive proteins accumulated at a very early stage of the heterosis manifestation, besides identifying candidate proteins to favor the hybrid vigor in papaya
- To identify and quantify differentially abundant proteins in embryogenic and non-embryogenic callus of papayaduring multiplication phase, besides identifying the molecular mechanisms involved in the competence acquisition during early somatic embryogenesis development.

3. CHAPTERS

3.1. COMPARATIVE PROTEOMIC ANALYSIS OF HETEROSIS PHENOMENON IN PAPAYA ROOTS

3.1.1. INTRODUCTION

Heterosis, or hybrid vigor, is a genetic phenomenon involving the superior performance of F_1-hybrid heterozygous plants in terms of increased biomass, size, yield, growth rate, fertility, disease resistance, or resistance to environmental stress compared with the average performance of their homozygous parental lines (Falconer and Mackay 1996; Mohayeji et al. 2014).

Heterosis was first described by Charles Darwin in 1876 and rediscovered independently by George H. Shull and Edward M. East in 1908 (Hochholdinger and Hoecker 2007). In papaya, heterosis was first observed by Lassoudiére (1968) in an F_1 hybrid derived from a cross between the genotypes Philippine x Solo, which made the hybrid more vigorous and with early flowering. Meaningful results have been achieved in the development of competitive hybrids with important agronomic traits, including fruit production (Marin et al. 2006; Cardoso et al. 2014) and disease resistance (Vivas et al. 2012; Vivas et al. 2014). The hybrid UENF/Caliman04 (UC04) shows high heterosis for important characteristics such as productivity,

soluble solids content and commercial fruit quality (Cardoso et al. 2014).Because of its great economic and scientific importance, heterosis has been studied through various methods, such as quantitative genetics, physiology and molecular biology. However, because of the great complexity of heterosis, its regulatory mechanisms remain poorly understood.

Despite the major advances of genomics in understanding the mechanisms responsible for the expression of characteristics (phenotypes) of agronomic interest, genomics alone may be insufficient to answer all questions regarding the transmission of genetic information. For instance, studies have demonstrated a substantial difference between gene expression and the abundance of mRNAs, which are constantly under the influence of different mechanisms of regulation and epigenetic control of gene expression(Banks et al. 2000). Gene expression and metabolic studies in *Zea mays*, *Oryza sativa*, and other species suggest that protein metabolism is involved in the growth differences observed between hybrids and inbreds (Goff 2011). Thus, proteomic analysis in plants is emerging as an important tool in plant breeding because it reflects the variability in gene expression (Pennington and Dunn 2001; Cramer et al. 2013; Eldakak et al. 2013).

Recent studies have used proteomic approaches to improve our understanding of heterosis during several stages of plant development, such as in seeds (Marcon et al. 2010), during germination (Fu et al. 2011), and during root (Marcon et al. 2013) and leaf (Mohayeji et al. 2014) development. However, young roots have been considered a model for the study of the molecular basis of heterosis (Yao et al. 2005; Hoecker et al. 2008). To investigate heterosis at the protein level, two strategies have been used. Two-dimensional gel electrophoresis (2-DE) is one technique used to detect differences in protein abundance; this is done by comparing stained protein spot volumes followed by protein identification using mass spectrometry. The gel-free shotgun technique uses bioinformatic tools and computational algorithms to measure quantitative differences at the protein level (Mohayeji et al. 2014). The gel-free shotgun technique has the following advantages: increased sensitivity, the identification of very high- or low-molecular-weight proteins, and the detection of highly acidic, basic, or hydrophobic proteins (Domon and Aebersold 2006; Panchaud et al. 2008).

Advances in mass spectrometry have enabled the generation of high-quality and reliable data for the analysis of complex mixtures of proteins. In particular, the

use of MS^E acquisition generates multiplex fragmentation data for peptides of precise mass; from this, both quantitative and qualitative characterizations of complex proteomic samples can be achieved (Silva et al. 2005; Chakraborty et al. 2007).

To date, proteomic studies in papaya have primarily analyzed responses to disease, fruit development (Rodrigues et al. 2009; Rodrigues et al. 2011; Angel Huerta-Ocampo et al. 2012; Rodrigues et al. 2012), and somatic embryogenesis (Vale et al. 2014). Proteomic studies may contribute to the development of papaya, a fruit of great economic importance worldwide. Papaya is grown primarily in tropical countries and is rich in nutrients, vitamins A and C, niacin and calcium (Ming et al. 2008). Various products can be extracted from papaya, such as papain and carpain (Oliveira et al. 1994).

Thus, the aim of this study was to identify and quantify the differentially abundant proteins in the roots of the papaya hybrid from their parental lines in order to identify non-additive proteins accumulated at an early stage of heterosis, identify candidate protein biomarkers and investigate the molecular mechanisms that promote hybrid vigor in papaya.

3.1.2. REVIEW

3.1.2.1. *Carica papaya*

The papaya (*Carica papaya* L.), a member of the Caricaceae family, originates in southern Mexico and northern Central America (Badillo 1993) and is widely distributed in several tropical regions (Schmildt et al. 2005). It is a diploid species with 18 chromosomes and a relatively small genome of 372 Mb (Zhang et al. 2008), polygamous, preferably self-pollinated.

Wild papaya populations are dioecious, with one-half male and one-half female plants, whereas cultivated papaya is predominantly gynodioecious, with two-thirds hermaphrodite and onethird female plants, though dioecious varieties do exist (VanBuren et al. 2015). The center of papaya origin is probably the Northwest of

South America, more precisely in the Amazon Basin, where the genetic diversity of this species is greater (Farias et al. 1994).

It is one of the most cultivated and consumed fruits in the tropical and subtropical regions of the world (Chen et al. 1991). It can be consumed not only in natura but also industrialized, in a range of products and by-products that can be used by the food, pharmaceutical and animal feed industries (Ruggiero et al. 2011). This is because in addition to being a fruit rich in nutrients, vitamins A, B_1, B_3, C, and calcium , from it can be extracted various products such as papain and carpaine, which is an alkaloid used as a cardiac activator (Oliveira et al. 1994).

From the introduction of papaya cultivation in Brazil, many advances have been achieved in particular because of the scientific research that leveraged their production not only at national level but also at international level (Ruggiero et al. 2011) and made Brazil a great world producer of this fruit.

Papaya fruit world production in 2014 was about 12.6 million t in an area of 411.3 thousand hectares representing a yield of 30.8 t/ha. The five largest countries producing this fruit in 2014 were India, Brazil, Nigeria, Indonesia and Mexico (FAO 2016). Brazil presented a production of 1.60 million tons in 2014 (FAO, 2014). The cultivation of papaya is carried in almost all Brazilian territory, concentrated in Bahia (49.5%) and in Espírito Santo (24.9%) (IBGE, 2016). Higher yields are observed in Bahia (64.2 t ha-1), Espírito Santo (63.0 t ha-1) and Ceará (39.8 t.ha-1).

The papaya varieties can be grouped in two main groups: Solo and Formosa (Dias et al. 2011). In the Solo group, some cultivars stand out as more planted: Sunrise Solo, Improved Sunrise Solo, Baixinho de Santa Amália, Sunrise Golden and Taiwan. In the Formosa group, the following cultivars are planted in Brazil: Tainung n.1 and Uenf/Caliman 01 (UC01) (COSTA et al. 2003).

Increasing the productivity of a crop depends on a number of factors, among which methods of breeding and selection of varieties with higher yields can contribute decisively to the breeding of papaya (Dantas and Lima 2001).

The development of varieties of papaya with good agronomic characteristics, such as high fruit quality and resistance to diseases, requires the use of crosses between contrasting genotypes. Because of this, one of the main demands for the breeding of the culture, part of the collection, introduction, characterization and germplasm conservation of gender and culminates in the evaluation and development of varieties for different ecosystems (Dias et al. 2011).

For a long time, the production of inbred papaya cultivars was the main objective of plant breeding work in several countries, including Brazil. This is because papaya does not suffer from the loss of vigor due to endogamy (Dantas and Lima 2001).

The plant breeding of the culture is mainly developed based on intraspecific hybridization between the groups Solo and Formosa, or within each of these groups.The first studies aimed at the development of hybrids in Brazil had begun in the 70's (Ferreguetti 2003). Currently Brazil has thirteen hybrids registered in the National Register of Cultivars (RNC), being the majority of the Formosa group, whose main characteristic is the mass of the fruit around 1.5 kg, and only two hybrids of the soil group, that attend the demand of fruits for the export, that is of fruits type papaya or havaí (Luz et al. 2015). In this way, the availability of new cultivars, productive and adapted to the wide range of cultivation environments and to diverse market demands is of great importance for the development of this culture in Brazil.

Among them, the UENF/CALIMAN04 hybrid from the cross between JS-12 and genotype São Mateus presents short plants with short stature, high productivity, little susceptibility to carpelloidy and pentandry. In addition, it presents high external and internal firmness of the fruits (Ide et al. 2009). This hybrid still stands out in the sugar content (Cardoso 2012) and in the greater resistance to phoma leaf spot (Vivas et al. 2014).

3.1.2.2. Heterosis

The term heterosis or hybrid vigor is the genetic phenomenon, which refers to the superior performance of heterozygous F1-hybrid plants in terms of biomass increase, size, yield, rate grown, fertility and disease resistance to environmental stress as compared with the average of its homozygous parent lines (Falconer and Mackay 1996).

Heterosis was first described by Charles Darwin in 1876, nevertheless, it was from 1908 with the study of George H. Shull and Edward M. East (Hochholdinger and Hoecker 2007), that his potential has been extensively exploited in agriculture, especially in maize, due to the high productivity gains achieved by hybrids (Schnable and Springer 2013).

Heterosis is more evident in the characteristics of the adult phase, however it already manifests during the initial phase of the development of the embryo (Meyer et al. 2007). In papaya heterotic effect was already observed during the germination of hybrid seeds, including the hybrid UC01, through the analysis of the percentage of germination and more vigorous growth of the hybrids in relation to the parents (Martins 2007).

In this sense, due to its great economic and scientific importance, it has been studied through several approaches such as quantitative genetics, physiology and molecular approaches. However, due to the great complexity of this phenomenon, the mechanisms that regulate it are still poorly understood.

The first hypotheses to explain heterosis are derived from quantitative genetics. The first one was the theory of dominance (complementation), which proposes that heterosis is a result of the complementation of the deleterious recessive alleles that are present in the consanguineous parents (Bruce 1910). The second is the overdominance hypothesis, which proposes that allelic interactions occur in the hybrid such that the heterozygote class performs better than any homozygous class (Shull 1908). However, none of these models alone can adequately explain the evidence about heterosis.

The study of quantitative loci (QTLs) was the first advance in molecular understanding of heterosis (Lippman and Zamir 2007). However, despite the several studies using QTLs associated with heterotic traits (hQTLs), to study the roles of dominance and overdominance in heterosis, there was no common finding to explain heterosis. Some authors concluded that the hQTLs acted through mechanisms of domination, whereas other authors concluded by mechanisms of overdominance (Schnable and Springer 2013).

Subsequently, a range of studies on the gene expression associated with heterosis in plants was started, using from very young embryos to adult plants. And again were not identified global uniform standards of gene expression in relation to the standard of additivity and non-additivity of the hybrid in relation to the parental lines in these studies, as well as between different tissues used (Hochholdinger and Hoecker 2007).

With advances in molecular biology and biochemistry of plants, heterosis is seen as a result of a complex network of interaction not only at the genetic level, as epigenetic, biochemical and regulatory (Chen 2013).

Among the advances found, epigenetic modifications of genes seem to play a key role in hybrids, which can alter complex regulatory networks of physiology and metabolism, thus modulating biomass increment and leading to heterosis (Birchler et al. 2003).

In this sense, studies at the proteome fieldmay be important for the understanding of heterosis, since they represent what is actually produced by the genes. Studies on seminal maize roots (Marcon et al. 2013), sunflower inflorescence (Mohayeji et al. 2014), maize seed (Marcon et al. 2010), germination of maize seeds (Fu et al. 2011), shoots of field-grown of maize (Dahal et al. 2016), leaves of sorghum-sudangrass (Han et al. 2016) and mature rice embryos (Wang et al. 2008) already bring important information, showing non-additive character changes that do not necessarily correspond to the non-additive expression expressed by the genes.

In the study conducted by Marcon et al. (2013) of heterosis in seminal roots of maize allowed the identification of an increase in the rate of protein synthesis in hybrids that may contribute to the early manifestation of heterosis in the seminal roots of maize.

Mohayeji et al. (2014) studying the inflorescence of sunflower hybrids and their respective parents observed that 38 proteins differentially expressed in the hybrid are directly associated with heterosis.These heterosis-related proteins are basically associated with the increase of the energy input mechanisms of the plant with reinforcement of the carbon fixation pathway and reduction of consumed energy directed to the production of superior hybrids (Mohayeji et al. 2014).

These results already demonstrate the possibilities and advantages in the use of the proteomic approach in the study of heterosis and reinforce the importance of this tool in the analysis of plants as a whole.

Another justification for this type of approach comes from information already available on the transcriptome of some species, as well as between hybrids and their parents (Springer and Stupar 2007). These studies have shown a great difference between the expression of genes and the abundance of mRNAs (Schnable and Springer 2013). Also this show the importance of studies at the protein level, since it demonstrates what was actually expressed by the genes, which are constantly under the influence of several regulatory mechanisms besides the epigenetic control of gene expression (Banks et al. 2000).

3.1.2.3. Plant proteomics

Proteomics is the large-scale study of proteins from a complex biological sample(Wilkins et al. 1996). It deals with the analysis of proteins in organisms, organs, cell populations and subcellular compartments, all under diverse developmental and environmental conditions (Takac et al. 2011).

The research on plant proteomics presents some limitations when compared to the biomedical area. It faces the intrinsic challenge more complex and less susceptible to data analysis, especially due to greater genomic diversity of the plant kingdom and less coverage of genomic sequences (Cramer et al. 2013). In particular with respect to the proteomic investigation of hormonal pathways regulating the development of plants (Takac et al. 2011).

Proteomic studies in papaya are still scarce and focus on disease response and fruit development (Rodrigues et al. 2009; Rodrigues et al. 2011; Angel Huerta-Ocampo et al. 2012; Rodrigues et al. 2012; Soares et al. 2017). In a proteomic study during the somatic embryogenesis of papaya, our group identified proteins important for the maturation of somatic embryos of papaya (Vale et al. 2014). Currently there is no study of heterosis in initial stages of development using this approach.

One of the factors that hinder the development of proteomic protocols in plants is that plant tissues present large amounts of different metabolites, such as mono and polysaccharides, phenolic compounds and oils (Fröhlich and Lindermayr 2011), besides presenting low protein content, due to the presence of the cell wall and vacuoles that represent the majority of the cellular mass (Carpentier et al. 2005).In this way, plant proteomics often requires efforts in the development of adequate protocols for each tissue and species study.

The ideal extractive selection is one that solubilizes the largest amount of proteins and depends on the species, tissue and the proteins of interest. The different extractive solutions have affinities with specific classes of proteins, which allows a differential extraction according to the method used (Carpentier et al. 2005). The most used for extraction of total proteins in plants, the chaotropic ones, as: urea and thiourea or directly the trichloroacetic acid, followed by at least one method of precipitation for the concentration of the proteins and the elimination of the interferents (Natarajan et al. 2005).

The information about the proteome of a sample may be derived from the analysis of intact proteins (proteomics top-down) or the digestion of proteins in a complex mixture (bottom-up proteomics), the bottom-up analysis of complex mixtures, can be used denaturing gels or liquid chromatography (Schluter et al. 2009).

Proteomic analysis based on the separation by two-dimensional polyacrylamide gel electrophoresis (2-DE) is done through the digestion of proteins with a protease and identification by MS, while the gel free method is performed by digestion of a mixture of proteins with a protease, separation of the peptides by LC and identification by MS (shotgun) (Champagne and Boutry 2013). One of the major challenges for the use of the proteomic approach is the analysis of thousands of proteins present in a single sample, with mass spectrometry (MS) being the most powerful tool to achieve this goal (Colas et al. 2010).

One of the approaches used in proteomics is comparative proteomics. In this approach the objective is not to identify the total set of proteins in a particular sample, but to characterize differences between different protein populations (Rose et al. 2004), which allows the study of protein profileunder different conditions. Moreover is ideal for the identification of stage-specific proteins whose presencecan be used as markers of development (Dias et al. 2010).

3.1.2.4. Proteomics applied to plant breeding

Plant breeding began in prehistory and has been used for the selection of viable phenotypes and since then has brought great advances in food production (Moose and Mumm 2008), however, the growing need for food and energy demand increasingly larger and more accurate information about the genotypes used in breeding programs.

In this context, proteomics, a science that studies the whole or part of the protein complements of a complex biological system, at any moment in time, which includes the characterization and quantification of protein expression, function and structure (Shi et al. 2004), can provide important information for the development of new cultivars.

After sequencing the genome of model plant species, much has been done to understand plant responses to biotic and abiotic stresses, however, these studies

remain essential for understanding phenotypic specificity and diversity (Vanderschuren et al. 2013).

Although proteomics is already widely used in the animal area, its progress in the plant area has been slower, with most studies being carried out on arabidopsis, rice, wheat and maize (Foster et al. 2006). Only in recent years has proteomics emerged as a tool in the plant breeding of agronomic importance (Cramer et al. 2013).

Proteomics can be used as an important tool in plant breeding because it provides molecular level information on the genetic variability that is effectively expressed in the genome (Pennington and Dunn 2001), and it has a greatest impact on phenotype improvement (Boggess et al. 2013).

The proteomic study applied to plant breeding can be used in the most diverse approaches: response of resistant and susceptible cultivars to pathogens (Gonzalez-Fernandez and Jorrin-Novo 2011) and abiotic stress studies (Kosová et al. 2011; Barkla et al. 2013).Proteomic analysis of wild relatives that can provide useful information on the regulationof the protein associated with beneficial characteristics or even serve as a resource of genes and proteins to restore traces lost through domestication and selective breeding (Vanderschuren et al. 2013) andpost-harvest (Pedreschi et al. 2013). However, according toVanderschuren et al. (2013) there is still great difficulty in interpreting large sets of quantitative protein data and translating them into practical use information for plant breeding. Moreover, one of the most promising applications refers to the development of new markers mainly those related to loci of quantitative characteristics also denominated "quantitative locus of protein" PQLs (Vanderschuren et al. 2013) as noted by Bourgeois et al. (2011) in *P. sativum* seeds

Similar to some achievements already made in animals, in plants there is more use of proteomics to create and understand the development of hybrids, however, the use of this tool depends on the complete sequencing of the genome of the plant of interest (Vanderschuren et al. 2013). Since currently one of the major obstacles to the use of plant proteomics, is the difficulty in identifying protein species whose genome has not been sequenced.

Plant proteomics although a relatively young science, has already made several advances in understanding the most diverse aspects of plant biology models,

however, there is still a lack of information and strategies on how to integrate this tool into plant breeding programs.

3.1.3. MATERIALS AND METHODS

3.1.3.1. Plant Material

Seeds of the F_1 hybrid (UC04) (♀ JS12 x ♂ São Mateus), obtained from the crossing of different heterotic groups, and its parental lines (JS12, Formosa group and São Mateus, Solo group), were collected under similar conditions and came from the Agricultural Caliman company S/A, located in Linhares, Espírito Santo (ES), Brazil (19° 23'S and 40° 4'W). The seeds were disinfected for 1 min in 70% ethanol and for 10 min in 50% commercial bleach (2-2.5% sodium hypochlorite), followed by three washes with distilled and autoclaved water. Subsequently, the seeds were germinated in accordance with the standard protocols established by the Rules for Seed Analysis (Brasil 1992), using a BOD-type germination chamber, regulated at 30°C/20°C (16 h light/8 h dark). Six replicates were performed, with each replicate composed of 8 petri dishes, each containing 10 seeds, in a completely randomized design. The number of roots per seed (RN) was evaluated. The primary roots, 3 cm in length on average, were collected from all replicates of the three genotypes. Samples of 300 mg fresh matter (FM) were dried in an oven at 70°C for 48 h for root dry matter (RDM) determination. Samples of 300 mg FM were stored at −20°C for proteomic analysis.

3.1.3.2. Protein extraction and quantification

Protein extracts were prepared in biological triplicate (300 mg FM each) for each evaluated genotype. Proteins were extracted using the trichloroacetic acid (TCA) (Sigma Chemical Co., St. Louis, MO)/acetone precipitation method developed by Damerval et al. (1986), with modifications. Root tissue was frozen in liquid N_2 and ground to a fine powder using a ceramic mortar and pestle. The resulting powder was suspended in 1 mL of chilled extraction buffer containing 10%

(w/v) TCA in acetone with 20 mM dithiothreitol (DTT) (GE Healthcare, Freiburg, Germany) and kept at -20°C for 1 h before centrifugation at 16,000 g for 30 min at 4°C. The resulting pellets were washed three times for 10 min in cold acetone with 20 mM DTT. The pellets were air dried, resuspended in buffer containing 7 M urea, 2 M thiourea, 2% Triton X-100, 1% DTT, 1 mM phenylmethylsulfonyl fluoride (PMSF) (Sigma-Aldrich), and 5 µM pepstatin; incubated for 30 min on ice; vortexed; and centrifuged for 20 min. The supernatants were collected, and the protein concentration was determined using the 2-D Quant Kit (GE Healthcare, Piscataway, NJ, USA).

3.1.3.3. Protein digestion

Samples of 100 µg of total protein were prepared according to Reis et al. (2015). Initially, samples were desalted on 5000 MWCO Vivaspin 500 membranes (GE Healthcare, Little Chalfont, UK). The membranes were filled until maximum capacity with 50 mM ammonium bicarbonate (Sigma-Aldrich) at pH 8.5 and centrifuged at 15,000 g for 20 min at 8°C. This procedure was repeated at least 3 times, with approximately 50 µL of sample remaining.

For protein digestion, we used the methodology described by Calderan-Rodrigues et al. (2014). Briefly, 25 µL of 0.2% (v/v) RapiGest® (Waters, Milford, CT, USA) was added and the samples were briefly vortexed and incubated in an Eppendorf Thermomixer® at 80°C for 15 min. Then, 2.5 µL of 100 mM DTT was added and the tubes were vortexed and incubated at 60°C for 30 min under agitation. Next, 2.5 µL of 300 mM iodoacetamide (GE Healthcare) was added and the samples were vortexed and then incubated in the dark for 30 min at room temperature. The digestion was performed by adding 20 µL of trypsin solution (50 ng μL^{-1}) (V5111, Promega, Madison, WI, USA) and incubating the samples at 37°C overnight. For RapiGest® precipitation, 10 µL of 5% (v/v) trifluoroacetic acid (TFA, Sigma-Aldrich) was added and incubated at 37°C for 90 min, followed by centrifugation for 30 min at 15,000 g. The samples were then transferred to Total Recovery Vials (Waters, USA).

3.1.3.4. Mass spectrometry analysis

A nanoAcquity UPLC connected to a Synapt G2-Si HDMS mass spectrometer (Waters) was used for ESI-LC-MS/MS analysis. Peptide mixtures were separated by liquid chromatography using 1 µL of digested samples in scouting runs. Normalization among samples was based on total ion counts consisting of three replicates per pooled sample. During separation, the samples were loaded onto the nanoAcquity UPLC 5 µm C18 trap column (180 µm x 20 mm) and then onto the nanoAcquity HSS T3 1.8 µm analytical reversed-phase column (100 µm x 100 mm) at 600 nL/min, with a column temperature of 60°C. For peptide elution, the binary gradient consisted of water (Tedia, Fairfield, Ohio, USA) and 0.1% formic acid (Sigma-Aldrich, St. Louis, MO, USA) as mobile phase A, and acetonitrile (Sigma-Aldrich) and 0.1% formic acid as mobile phase B. Gradient elution started at 7% B up to 40% B in 90.09 min and from 40% B to 85% B until 94.09 min, maintained at 85% until 98.09 min, then decreased to 7% B until 100.09 min and maintained at 7% B to the end at 108.09 min. Mass spectrometry was performed in positive and resolution mode, 35,000 FWMH, and the transfer collision energy ramped from 19 v to 45 v in high-energy mode; cone and capillary voltages of 30 v and 2,800 v, respectively; and a source temperature of 70°C. In TOF parameters, the scan time was set to 0.5 s in continuum mode with a mass range of 50 to 2,000 Da. The human [Glu1]-fibrinopeptide B (Sigma-Aldrich) was used as an external calibrant. Data-independent acquisition (DIA) scanning with added specificity and selectivity of a non-linear ‘T-wave’ ion mobility device was performed ($HDMS^E$) (Heringer et al. 2015).

3.1.3.5. Protein identification and quantification

Spectrum processing and database searching were performed using Progenesis QI for Proteomics Software V.2.0 (Nonlinear Dynamics, Newcastle, UK). The analysis used the following parameters: one missed cleavage, minimum fragment ions per peptide equal to 1, minimum fragment ions per protein equal to three, minimum peptides per protein equal to 1, fixed modifications of carbamidomethyl (C) and variable modifications of oxidation (M) and phosphoryl (STY) groups, a default false discovery rate (FDR) value of 4% maximum, a score greater than five, and maximum mass errors of 10 ppm. The generated peptide masses were searched against the Uniprot Brassicales protein sequences database

(2014/11). Label-free relative quantitative analyses were performed by the ratio of protein ion counts among contrasting samples. After the Progenesis analysis and to ensure the quality of results, only proteins present in 3 of 3 runs and with coefficients of variation less than 0.3 were selected. Functional classification of the identified proteins was performed using the program Blast2go (www.blast2go.com).

3.1.3.6. Statistical analysis

The RN and RDM values were used to determine mid-parent heterosis (MPH) and best-parent heterosis (BPH) using the following formulas: MPH = (mean F_1 - mean P)/mean P in % and BPH = (mean F_1 - mean best P)/mean best P in %.

To identify proteins exhibiting non-additive characteristics, data on total ion counts (TIC) were submitted to analysis of variance (ANOVA) using the following linear model:

$$y_{ij}=\mu+ T_i+ e_{ij}\ i\in 1\ldots 3;\ j\in 1\ldots 3,$$

where y_{ij} is the j-th repetition of genotype i; μ is the general average, which is present in every y_i; Ti is the effect of treatment i; and e_{ij} is the random error.

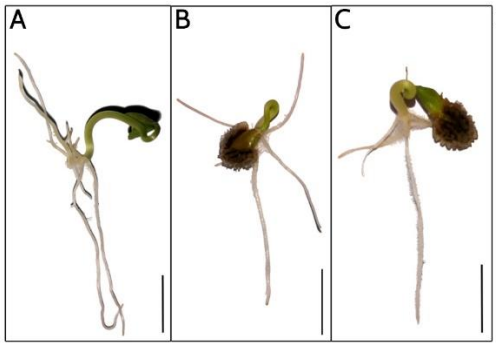
Based on the fit of this model, the contrast between the average hybrid and its parental lines was measured using a t-test assuming the following null hypothesis (H_0): H = (P_1 + P_2) / 2, in which H is the hybrid mean and P_1 and P_2 are the means of the two corresponding parents, respectively. The non-additive protein relative abundances were then subjected to the SNK test using the statistical software R (R Core Team 2014) and the easyanova package (Arnhold 2013).

Proteins with abundances that were significantly higher in the hybrid offspring than in the higher parental line were classified as “above high parent” (++). Proteins with abundances that were significantly lower than the lower parent were classified as “below low-parent abundance” (−−). Proteins in the hybrid that displayed significant differences from the lower parent but no significant differences from the higher parent were labeled “high-parent abundance” (+). Proteins in the hybrid with significantly lower abundances than the high-parent but no significant difference from the low-parent were classified as “low-parent abundance” (−).

3.1.4. RESULTS

3.1.4.1. Heterosis effects in root growth

MPH and BPH were calculated to determine the RDM percentage, the RN, and the protein content (µg/g DM). The heterosis analysis identified positive effects for RDM; MPH was 27.1%, and BPH was 15.5%. For RN, MPH was 28.2%; BPH was 15.9%. Negative effects were observed for the root protein content; MPH was -25.9%, and BPH was -37.5% (Fig. 1).



D

	Root traits		
	RDM(%)	RN	Protein (µg/g DM)
Hybrid	8.9a*	5.7a	43.5a
JS12	7.7b	4.9b	47.9a
São Mateus	6.3c	3.9c	69.6a
MPH(%)	27.1	28.2	-25.9
BPH(%)	15.5	15.9	-37.5

Fig. 1 Seminal root development in *Carica papaya* seedlings in (a) hybrid UC04, (b) parental JS12, (c) parental São Mateus and (d) average values and heterosis for root traits in a papaya hybrid. *Means followed by the same letters are not significantly different according to the SNK test (P < 0.01). RDM: root dry matter; RN: root number; C.V.: coefficient of variation; MPH: mid-parent heterosis, calculated using the following formula: MPH = (mean F 1 - mean P)/mean P in %; and BPH: best-parent heterosis, calculated using the following formula: BPH = (mean F 1 - mean best P)/mean best P in %. (C.V. RDM = 7.2% - n = 3; C.V. RN = 9.6% - n = 3, C.V. Protein = 38.7% - n = 3).

3.1.4.2. Proteomic profiles of the heterosis response

A total of 955 proteins were identified. Among these proteins, 938 were common to the three genotypes, six were exclusive to the hybrid genotype and the female parent JS12, seven were exclusive to the hybrid genotype and the male parent São Mateus, one was exclusive to the parents, two were unique to the progenitor JS12 and one was unique to the hybrid (Fig. 2). Among the 955 identified proteins, 662 displayed significantly different abundances between genotypes (ANOVA p-value < 0.05) and 261 (27.3%) displayed non-additive accumulation in the seminal roots of the hybrid JS12 × São Mateus (t-test p-value < 0.05).

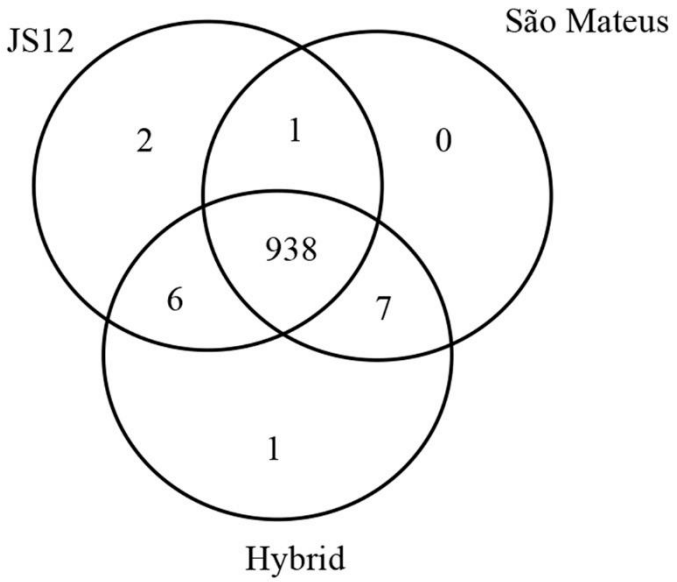


Fig. 2 Venn diagram of proteins identified in primary roots of the *Carica papaya* hybrid UC04, parentalJS12 and parental São Mateus.

3.1.4.3. Non-additively accumulated proteins in seminal roots

Non-additive proteins, when their abundances were significantly different from the parental average, were used for the following analyses. T-tests were performed on the total ion count (TIC) values of these proteins to identify significant differences in protein abundance between the hybrid offspring and the parental inbred lines. A total of 261 proteins exhibited non-additive abundance. Among the different heterotic classes, 58 proteins (22.2%) were classified as “above high-parent abundance,” 17 proteins (6.5%) were classified as “high-parent abundance,” 42 proteins (16.1%) were classified as “low-parent abundance” and 144 proteins (55.2%) were classified as “below low-parent abundance” (Fig. 3). These results demonstrated that the “below low-parent abundance” class was the most frequent heterosis response.

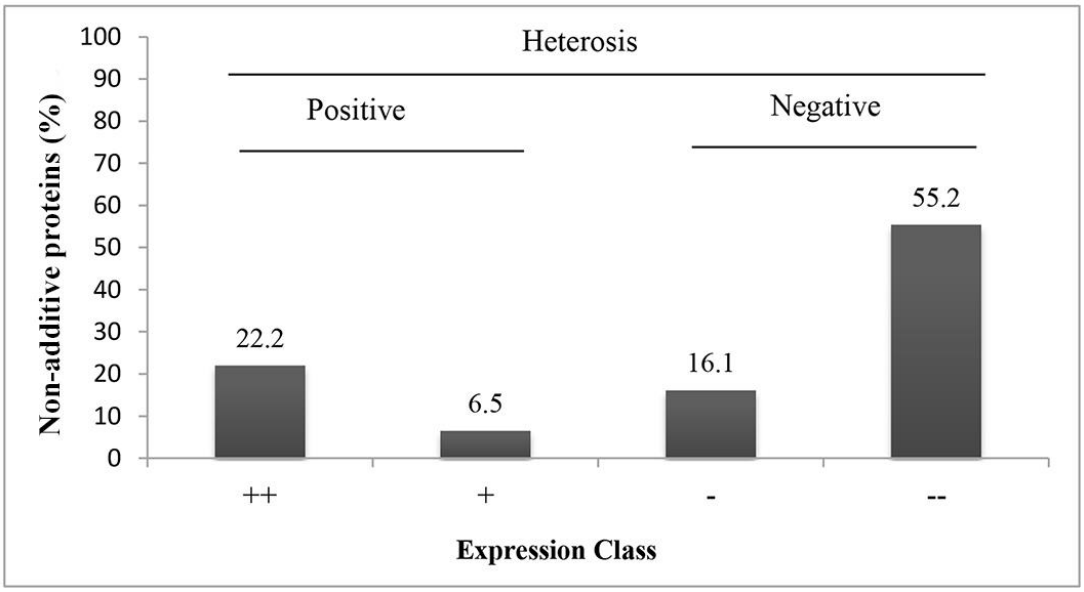


Fig. 3 Number of non-additive identified proteins expressed in *Carica papaya* hybrid UC04 primary roots compared with parental inbred lines, JS12 and São Mateus. “Above high-parent abundance” (++), “Below low-parent abundance” (−−), “High-parent abundance” (+), and “Low-parent abundance” (−).

3.1.4.4. Functional classification of all non-additively accumulated proteins in hybrid roots

Proteins with non-additive abundance were separated into two groups: positive heterosis when the abundance of the hybrid was higher than the average parents (75 proteins) and negative heterosis with abundances lower than the parental average (186 proteins). These proteins were functionally classified according to their biological functions, and the gene ontology of the biological processes of the proteins resulted in 13 functional classes. The most representative classes of proteins for both groups were associated with metabolic, cellular, and single-organism processes. Proteins associated with auxin polar transport, regulation of signaling, root system development, and *reactive oxygen species (ROS) metabolic processes* were more abundant in the positive heterosis group (i.e., most abundant in the hybrid) (Fig. 4). The proteins associated with energy, protein metabolism, root development and *ROS metabolic processes* are listed in Table 1.

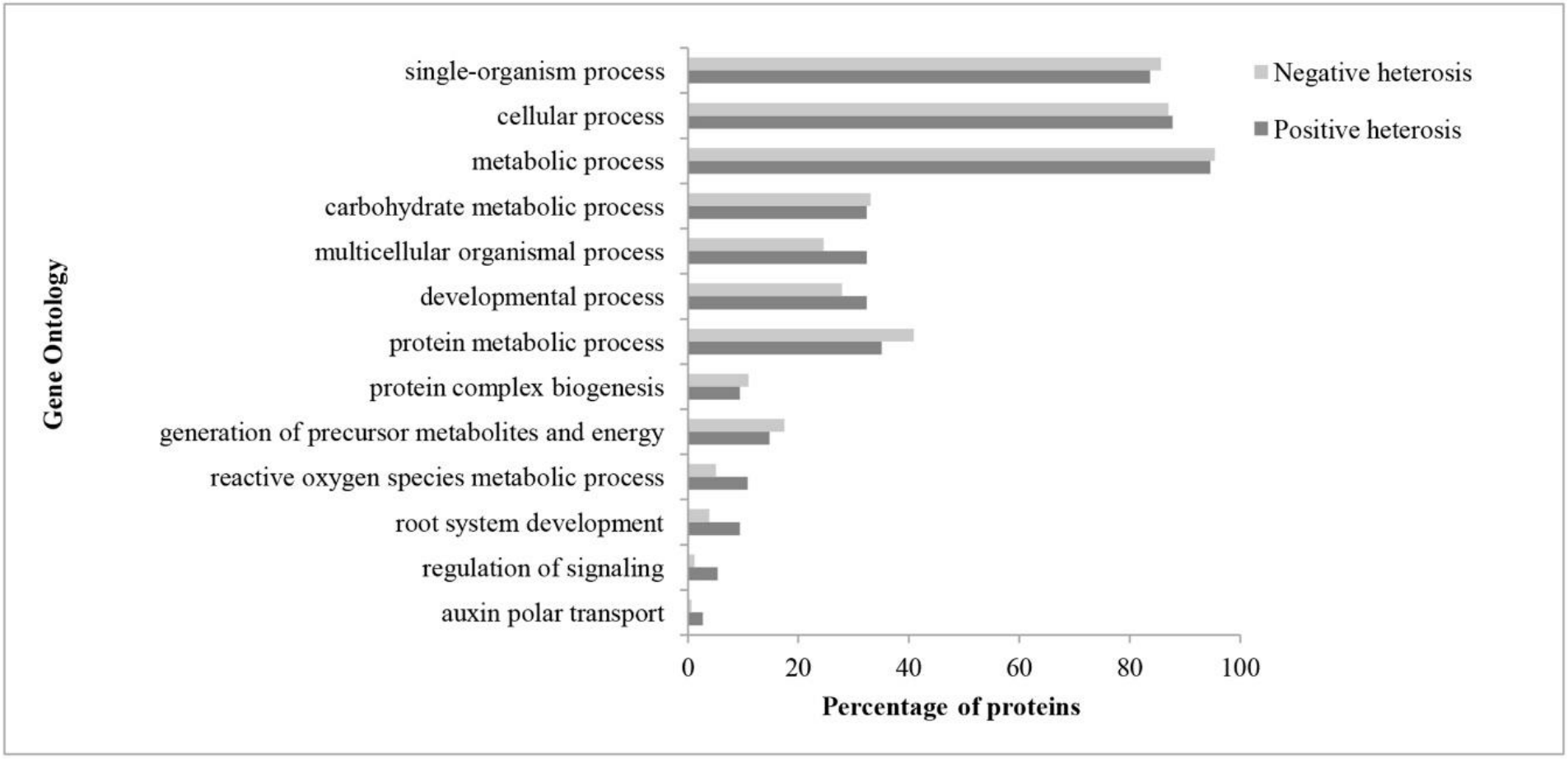


Fig. 4 Functional classification of non-additive identified proteins in *Carica papaya* primary roots.

Table 1 Abundance pattern of proteins important to hybrid vigor identified from the primary roots of the hybrid and its parental lines in papaya.

Protein name	Acession	Organism	Peptide count	Unique peptides	Score	Abundance Class	Class Functional
Oxalate-CoA ligase	4CLLA_ARATH	*Arabidopsis thaliana*	2	2	16.97	(++)	Energy
ABC transporter B family member 15	AB15B_ARATH	*Arabidopsis thaliana*	2	2	15.41	(++)	Root development
ABC transporter B family member 16	AB16B_ARATH	*Arabidopsis thaliana*	3	2	15.50	(-)	Root development
ABC transporter B family member 2	AB2B_ARATH	*Arabidopsis thaliana*	1	1	10.80	(++)	Root development
Aconitate hydratase 1	ACO1_ARATH	*Arabidopsis thaliana*	20	1	183.58	(--)	Energy
Aconitate hydratase 2, mitochondrial	ACO2M_ARATH	*Arabidopsis thaliana*	25	3	241.29	(--)	Energy
ATP synthase subunit beta, chloroplastic	ATPB_AETGR	*Aethionema grandiflorum*	8	2	71.90	(--)	Energy
NADH dehydrogenase subunit F (Fragment)	B3V9A2_9BRAS	*Cremolobus subscandens*	2	1	11.56	(--)	Energy
ATP synthase subunit alpha	B9U3K8_CARPA	*Carica papaya*	26	10	332,75	(-)	Energy
SUR1	B5KJ87_BRARP	*Brassica rapa subsp. pekinensis*	1	1	5.39	(++)	Root development
Chaperonin CPN60, mitochondrial	CH60_BRANA	*Brassica napus*	20	2	182.76	(--)	Proteins metabolism
60S acidic ribosomal protein P0	D7LEE5_ARAL	*Arabidopsis lyrata subsp. lyrata*	7	2	74.29	(-)	Proteins metabolism
Protein binding protein	D7LLN3_ARALL	*Arabidopsis lyrata subsp. lyrata*	2	2	10.63	(--)	Proteins metabolism
Acetyl-coenzyme A synthetase	D7MI40_ARALL	*Arabidopsis lyrata subsp. lyrata*	5	1	28.36	(--)	Energy
Succinate dehydrogenase 1-1	D7MLY6_ARALL	*Arabidopsis lyrata subsp. lyrata*	10	3	73,38	(--)	Energy
Pyruvate kinase	D7MSA3_ARALL	*Arabidopsis lyrata subsp. lyrata*	12	1	76.96	(--)	Energy
Acyl-coenzyme A oxidase	D7KG21_ARALL	*Arabidopsis lyrata subsp. lyrata*	1	1	5,57	(++)	Energy
Proteasome subunit alpha type	D9IVA6_CARPA	*Carica papaya*	10	6	72.41	(-)	Proteins metabolism
Glutamate decarboxylase 1	DCE1_ARATH	*Arabidopsis thaliana*	10	3	114.64	(--)	Energy
Glutamate decarboxylase 2	DCE2_ARATH	*Arabidopsis thaliana*	8	1	79.40	(--)	Energy
Dirigent protein 1	DIR1_ARATH	*Arabidopsis thaliana*	4	1	24.30	(-)	Root development
Ribosomal protein S5/Elongation factor G/III/V family protein	F4JB05_ARATH	*Arabidopsis thaliana*	2	1	15.67	(--)	Proteins metabolism
ATP-dependent zinc metalloprotease FTSH 10, mitochondrial	FTSHA_ARATH	*Arabidopsis thaliana*	2	1	10.64	(--)	Energy
Isocitrate dehydrogenase [NAD] regulatory subunit 1. Mitochondrial	IDH1_ARATH	*Arabidopsis thaliana*	6	1	51.80	(--)	Energy
Isocitrate dehydrogenase [NAD] catalytic subunit 5. Mitochondrial	IDH5_ARATH	*Arabidopsis thaliana*	6	2	48.02	(++)	Energy
Horseradish peroxidase isoenzyme HRP_2021	K7ZW58_ARMRU	*Armoracia rusticana*	3	1	29.00	(++)	ROS metabolic process

Table 1 – Cont.

Protein name	Acession	Organism	Peptide count	Unique peptides	Score	Abundance Class	Class Functional
Probable pyruvate kinase, cytosolic isozyme	KPYC_ARATH	*Arabidopsis thaliana*	7	1	49.35	(--)	Energy
NADP-dependent malic enzyme 2	MAOP2_ARATH	*Arabidopsis thaliana*	10	1	101,70	(++)	Energy
Ribosome biogenesis protein BOP1 homolog	M4C843_BRARP	*Brassica rapa subsp. pekinensis*	4	2	31.42	(--)	Proteins metabolism
17.6 kDa class I heat shock protein 2	M4D253_BRARP	*Arabidopsis thaliana*	2	2	10.98	(+)	Proteins metabolism
Fructose-bisphosphate aldolase	M4DL52_BRARP	*Brassica rapa subsp. pekinensis*	20	5	208.85	(--)	Energy
Pyruvate dehydrogenase E1 component subunit alpha	M4E788_BRARP	*Brassica rapa subsp. pekinensis*	3	1	20.83	(--)	Energy
Ubiquitin carboxyl-terminal hydrolase	M4EEX9_BRARP	*Brassica rapa subsp. pekinensis*	2	1	10.67	(++)	Proteins metabolism
40S ribosomal protein S27	M4F8V6_BRARP	*Brassica rapa subsp. pekinensis*	5	2	36.41	(--)	Proteins metabolism
Malate dehydrogenase. cytoplasmic 2	MDHC2_ARATH	*Arabidopsis thaliana*	10	2	130.66	(++)	Energy
Malate dehydrogenase. chloroplastic	MDHP_ARATH	*Arabidopsis thaliana*	6	2	59.35	(--)	Energy
Nudix hydrolase	NUDT6_ARATH	*Arabidopsis thaliana*	1	1	5.26	(++)	ROS metabolic process
Peroxidase 68	PER68_ARATH	*Arabidopsis thaliana*	5	1	30.63	(+)	ROS metabolic process
ATP-dependent 6-phosphofructokinase 5, chloroplastic	PFKA5_ARATH	*Arabidopsis thaliana*	3	2	16.83	(--)	Energy
Phosphoglycerate kinase 2	PGKH2_ARATH	*Arabidopsis thaliana*	10	2	99.85	(+)	Energy
Phosphoglycerate kinase 2. chloroplastic	PGKH2_ARATH	*Arabidopsis thaliana*	10	2	99.85	(+)	Energy
26S protease regulatory subunit 6A homolog B	PS6AB_ARATH	*Arabidopsis thaliana*	13	1	103.25	(--)	Proteins metabolism
Proteasome subunit alpha type-3	PSA3_ARATH	*Arabidopsis thaliana*	4	3	29.11	(--)	Proteins metabolism
Proteasome subunit beta type-3-A	PSB3A_ARATH	*Arabidopsis thaliana*	4	1	23.46	(-)	Proteins metabolism
26S proteasome non-ATPase regulatory subunit 1 homolog A	PSD1A_ARATH	*Arabidopsis thaliana*	10	7	74.09	(--)	Proteins metabolism
26S proteasome non-ATPase regulatory subunit 14 homolog	PSDE_ARATH	*Arabidopsis thaliana*	3	2	33.02	(--)	Proteins metabolism
Glutamate dehydrogenase	Q1H5A3_ARATH	*Arabidopsis thaliana*	11	1	71.52	(++)	Energy
Acyl-CoA-binding protein	Q7XJJ8_TROMA	*Tropaeolum majus*	4	4	31.36	(++)	Energy
17.7 kDa heat shock protein	Q6XBS2_CARPA	*Carica papaya*	6	3	45.24	(--)	Proteins metabolism
ATP synthase subunit beta	R0F3C6_9BRAS	*Capsella rubella*	24	15	303.28	(--)	Energy
Isocitrate dehydrogenase [NADP]	R0GKG3_9BRAS	*Capsella rubella*	11	1	101.57	(++)	Energy
Lipoxygenase	R0HW05_9BRAS	*Capsella rubella*	3	2	16.32	(++)	Root development

Table 1 – Cont.

Protein name	Acession	Organism	Peptide count	Unique peptides	Score	Abundance Class	Class Functional
Mitogen-activated protein kinase	R0IHQ3_9BRAS	*Capsella rubella*	2	1	10.70	(++)	Root development/ ROS metabolic process
Protein RALF-like 19	RLF19_ARATH	*Arabidopsis thaliana*	1	1	5.33	(++)	Root development
40S ribosomal protein S27-1	RS271_ARATH	*Arabidopsis thaliana*	2	1	14.46	(--)	Proteins metabolism
40S ribosomal protein S28-2	RS282_ARATH	*Arabidopsis thaliana*	2	1	13.41	(--)	Proteins metabolism
40S ribosomal protein S3-2	RS32_ARATH	*Arabidopsis thaliana*	3	3	21.72	(--)	Proteins metabolism
Superoxide dismutase [Cu-Zn]	SODC_CARPA	*Carica Papaya*	7	6	117.10	(+)	ROS metabolic process
Succinate-semialdehyde dehydrogenase, mitochondrial	SSDH_ARATH	*Arabidopsis thaliana*	9	3	61.71	(--)	ROS metabolic process
Transcription factor TCP21	TCP21_ARATH	*Arabidopsis thaliana*	1	1	5.53	(++)	ROS metabolic process
Ubiquitin carboxyl-terminal hydrolase 14	UBP14_ARATH	*Arabidopsis thaliana*	2	1	11.03	(++)	Proteins metabolism
Salt overly sensitive 1 (Fragment)	W8P3P3_TURGL	*Turritis glabra*	2	1	11.31	(++)	ROS metabolic process

3.1.5. Discussion

Heterosis is widely studied because of its importance in increasing vigor, particularly with respect to quantitative characteristics of economic importance. However, the molecular basis of heterosis remains unclear because the expression of these characteristics depends on many genes from the beginning of plant development. The early identification of heterosis in important characteristics for adult plant vigor is of great relevance. In this context, the roots seem to be an important model for understanding hybrid vigor because their efficient development is fundamental for the extraction of mineral nutrients and, consequently, for plant nutrition, which affects vigor directly (Yao et al. 2005).

Although hybrid vigor is more easily observed during adulthood, it was possible to identify hybrid vigor in phenotypic traits using RDM and RN analyses (Fig. 1). An important characteristic for the manifestation of heterosis is the genetic distance between the parents (Hallauer et al. 2010). Thus, the use of parents from two different heterotic groups in this study was important for the manifestation of heterosis. Macedo et al. (2013) reported an early manifestation of heterosis for the germination rate and FM of papaya seedlings. Song et al. (2007) identified heterosis in the early development of *Triticum aestivum* roots, and young roots are being used as a model for the study of the early stages of heterosis manifestation in *Zea mays* (Hoecker et al. 2006).

In the present study, we identified a higher number of proteins (27.5%) displaying non-additive accumulation in the primary roots of the hybrid JS12 × São Mateus (Table 1). The number of non-additive proteins in this study is higher than those observed in other studies that used label-free methods to identify non-additive proteins in Z*ea mays* (Marcon et al. 2013) and *Helianthus annuus* (Mohayeji et al. 2014). In the present study, advanced technology was used to identify proteins. Data independent acquisition was performed using an ion mobility mode (HDMS^{E}) technique. This method allows for the identification of many proteins from complex samples and can identify protein isomers and isoforms (Giles et al. 2011; Heringer et al. 2015).

The allelic complementation of two genomes belonging to different heterotic groups results in cis-trans and chromatin changes that result in differential gene expression. These expression patterns primarily affect a few major regulatory

pathways, inducing signal transduction pathways that may either individually or cumulatively affect several downstream metabolic pathways, including protein metabolism, in either a positive or a negative manner (Baranwal et al. 2012). Therefore, the functional classification of proteins identified here was divided into positive and negative heterosis to identify which classes were more abundant in the hybrid UC04 and which were more abundant in the parents JS12 and São Mateus. This relationship is important because heterosis reflects not only a higher protein abundance but also a decreased abundance of proteins that cause unnecessary energy expenditure, thus increasing efficiency during development. In this sense, the analysis of the functional classification of non-additively accumulated proteins revealed 13 functional classes (Fig. 4). The largest class of proteins was related to several categories, such as metabolic, cellular, and single-organism processes. Other classes of proteins, such as those involved in the regulation of signaling, protein metabolic processes, root system development, and polar auxin transport, were more highly expressed in the hybrid, demonstrating their important roles in heterosis in papaya roots.

Several proteins that exhibit non-additive patterns are involved in root development, signaling regulation and auxin polar transport (Table 1). In maize roots, heterosis was evidenced by an increase in the expression of ribosomal proteins (Marcon et al. 2013). In papaya roots, these proteins were more abundant in the parents. These results demonstrate that heterosis occurs differently depending on the species, type of tissue and stage of development (Melchinger 1999; Hochholdinger and Hoecker 2007). The most important classes of proteins (Table 1) that exhibited heterosis patterns in our study are discussed according to their main functional categories.

3.1.5.1. Energy-related proteins

The production and use of energy are important factors influencing plant development. Our results demonstrated that several non-additive proteins identified in papaya roots were included in this class (Fig. 4). Among these proteins, aconitate hydratase 1 and 2, enzymes containing an Fe-S cluster that catalyzes the reversible isomerization of citrate to isocitrate via cis-aconitate in the tricarboxylic acid cycle (Peyret et al. 1995). Also, acetyl-coenzyme A synthetase, which activates acetate

to acetyl-coenzyme A, provide two carbon metabolites used in many anabolic and energy generation processes in the cell (Starai and Escalante-Semerena 2004). Pyruvate dehydrogenase E1, which participates in a multienzyme complex, catalyzing the oxidative decarboxylation of pyruvate to yield an acetyl, is also related to energy production (Tovar-Méndez et al. 2003). In addition, pyruvate kinase and ATP-dependent 6-phosphofructokinase 5, important proteins in the final steps of the glycolytic pathway (Jurica et al. 1998), exhibited underdominance, demonstrating that energy consumption is higher in the parents than in the hybrid in papaya.

However, some proteins were more abundant in the roots of the hybrid compared with the inbred lines, such as acyl-coenzyme A oxidase, a key enzyme of the peroxisomal fatty acid β-oxidation system (Froman et al. 2000). Acyl-CoA-binding protein is in a family of proteins that facilitate the binding of long-chain acyl-CoA esters at the conserved acyl-CoA-binding domain and has been implicated in acyl-CoA transport, in the maintenance of intracellular acyl-CoA pools, and in the protection of cytosolic acyl-CoAs from hydrolysis by cellular acyl-CoA hydrolases (Xiao et al. 2008).

The small number of proteins related to energy production in the hybrid papaya roots may result from the rapid use of these proteins by the hybrid; moreover, many studies report that hybrids use less metabolic energy per unit growth (Ginn 2010; Goff 2011). One possible explanation for hybrid vigor is its higher efficiency in energy use via selective protein synthesis and metabolism (Goff 2011). Goff suggests that allelic variants often encode unstable or inefficient proteins. The production of these proteins is an energy-intensive process. Selective protein synthesis would be possible by selective regulation of allelic transcription so that the preferred allele encodes most of the protein, causing substantial improvements in cellular energy efficiency and phenotypic performance (Kaeppler 2011; Goff and Zhang 2013; Schnable and Springer 2013). In this sense, the hybrid can optimize protein synthesis, resulting in substantial improvements in energy efficiency and cellular phenotypic performance, as was observed in papaya roots.

3.1.5.2. Protein metabolism-related proteins

Most proteins related to synthesis, processing and, protein degradation were classified as “below low parent” (Table 1). Among them, the 60S acidic ribosomal

protein P0, 40S ribosomal protein S27, 40S ribosomal protein S27-1, 40S ribosomal protein S28-2, 40S ribosomal protein S3-2, protein-binding protein, ribosomal protein S5/elongation factor G/III/V, and ribosome biogenesis protein homolog BOP1 were observed. The lower abundance of this class of proteins in the papaya root hybrid reflects lower energy production, which works to reduce energy-consuming processes such as protein metabolism. This reduction has a positive effect on hybrid development. According to Goff (2011), the hybrid has a greater capacity to produce stable proteins, reducing the need for protein refolding and degradation.

Heat shock proteins (HSPs), which are responsible for protein folding, degradation, and the stabilization of damaged proteins (Rodziewicz et al. 2014), were more abundant in the parental lines. Proteins responsible for the protein degradation process, such as the 26S regulatory subunit 6A protease homolog B, 26S proteasome non-ATPase regulatory subunit 1 homolog A, 26S proteasome non-ATPase regulatory subunit 14 homolog and proteasome subunit alpha type-3, were also more abundant in the parental lines. The rate of decrease in protein metabolism in the hybrid could be due to the existence of fewer protein substrates to degrade. Most of the energy that the parental lines consume must be available for the maintenance of cellular and developmental processes, reducing the amount of energy available for the synthesis of additional biomass (Ginn 2010).

3.1.5.3. Root development-related proteins

Root development interferes directly with plant performance, as roots are responsible for the uptake of nutrients. Thus, an increase in the abundance of this class of proteins in the hybrid can greatly impact plant growth and development. SUPERROOT 1 (SUR1), which was up-regulated in the hybrid, is involved in indolic glucosinolate biosynthesis; this promotes higher-than-normal levels of endogenous free and conjugated auxin (Boerjan et al. 1995; Mikkelsen et al. 2004) and results in adventitious and lateral root development (Suzuki et al. 2008). Thus, we believe that SUR1 is a key protein involved in the superior performance of the hybrid UC04 over the parental lines JS12 and São Mateus.

Lipoxygenases, which also exhibited overdominance abundance, catalyze the formation of hydroperoxy derivates by oxygenating polyunsaturated fatty acids.

These molecules act as signaling molecules responsible for various developmental processes and defense under conditions of stress (Alemayehu et al. 2013).Another protein exhibiting overdominance expression was Mitogen-activated protein kinase (MAPK). MAPK is involved in the signaling of various biotic and abiotic stresses and has been implicated in the regulation of cell cycle and developmental processes (Nishihama et al. 2001). According to Pagnussat et al. (2004), the MAPK signaling cascade is activated during the adventitious rooting process induced by indole-3-acetic acid. In addition, MAPK has been associated with the positive regulation of cell division and growth in Arabidopsis thaliana (Krysan et al. 2002).

Protein dirigents, which demonstrated an overdominance pattern of abundance in papaya roots, can act as guides in lignin synthesis and are thus of great importance as mechanical supports and in the development of tracheal elements that allow for long-distance water conduction (Kwon et al. 1999; Davin and Lewis 2000; Burlat et al. 2001; Hosmani et al. 2013). Overdominance accumulation of this protein was observed in our study. The increase in the abundance of this protein in the hybrid may be related to the improved development of hybrid roots compared with the parental lines.

Auxins are key hormones involved in root development. In addition to an increase in the abundance of auxin-synthesis proteins in the hybrid roots, we also identified an increase in the abundance of ABC transporter family proteins. ABC transporters have been described as auxin carriers (Noh et al. 2001; Luschnig 2002). In Arabidopsis, several studies have demonstrated that ABC transporter family proteins regulate auxin homeostasis to control root development (Gaedeke et al. 2001; Santelia et al. 2005; Larsen et al. 2007). We believe that the presence of these transporters in greater abundance in the papaya hybrid UC04 compared to its parents suggests the optimization of the use of auxin in the development of papaya roots, as shown in Fig. 1.

3.1.5.4. Reactive oxygen species (ROS) metabolic process-*related proteins*

The development of the root requires high plasticity because of various biotic and abiotic stresses that occur during plant development. According to De Tullio et al. (2010), this plasticity is strongly linked to ROS, antioxidants, and related

enzymes, which are part of a signaling module that acts in regulating the operation of the apical root meristem. Thus, the greater abundance of these proteins in the papaya hybrid *UC04* demonstrates that hybrid vigor is directly related to the ability of the hybrid to better regulate its development.

In Arabidopsis, Salt overly sensitive 1 (SOS) is important in maintaining homeostasis through its role in sodium extrusion and controlling the long-distance transport of Na^+ from the root to shoot (Shi et al. 2000). NADP-ICDH catalyzes the production of NADPH, which is an essential component of cellular homeostasis. NADP-ICDH is involved in carbon and nitrogen metabolism, redox regulation, and responses to oxidative stress (Leterrier et al. 2011; Begara-Morales et al. 2013). These proteins were categorized as above high-parent abundance in papaya roots (Table 1), suggesting that the hybrid would have the greatest homeostatic capacity throughout its internal, structural and functional organization. Other proteins were categorized as above high-parent abundance, such as G-type lectin S-receptor-like serine/threonine protein kinase, *horseradish peroxidase isoenzyme and* the transcription factor TCP21. G-type lectin S-receptor-like serine/threonine protein kinase, which acts on tolerance to salt stress, confers greater resistance to salt stress to the hybrid (Sun et al. 2013). Horseradish *peroxidase isoenzyme is an* isoenzyme that participates in a variety of processes, such as the synthesis of cell wall components, indole 3-acetic acid regulation and defense mechanisms (Filizola and Loew 2000). The transcription factor TCP21 is part of a plant-specific family of transcription factors involved in growth, cell proliferation, and organ identity in plants (Giraud et al. 2010). These results demonstrate the ability of the hybrid to reduce the negative effects of environmental conditions. Another important protein in cellular homeostasis that displayed an overdominance pattern in papaya roots is superoxide dismutase (SOD). This protein acts in the first line of defense against ROS, catalyzing the conversion of superoxide radicals into molecular oxygen and hydrogen peroxide (Bowler et al. 1994). In addition, some proteins, such as Nudix hydrolases, are associated with detoxification processes in plants under abiotic stress (Huang et al. 2012). In Arabidopsis, this protein family confers tolerance to oxidative stress (Ogawa et al. 2009). In papaya, Nudix hydrolasesalso exhibited overdominance abundance, providing greater antioxidant activity, which can decrease the deleterious effects of ROS to the hybrid compared with the parents, JS12 and São Mateus.

3.1.6. CONCLUSION

The present study provides a high-resolution quantitative comparison of the proteome complexities of the hybrid UC04 and its parental inbred lines in the young primary roots of papaya. The results suggests that the strategy for hybrid vigor in young papaya roots involves the optimization of energy production and energetic expenditure by reducing protein metabolism, leaving more energy available to synthesize additional biomass (DM) through the expression of proteins, such as SUR1, MAPK and ABC transporters, that increase root number and biomass. Our results contribute to a better understanding of the molecular events that result in the superior performance of the F_1 hybrid.

3.2. DIFFERENTIALLY ABUNDANT PROTEINS ASSOCIATED WITH THE ACQUISITION OF THE EMBRYOGENIC COMPETENCE OF PAPAYA CALLUS

3.2.1. INTRODUCTION

Somatic embryogenesis is an *in vitro* developmental process, in which single cells or a small group of somatic cells are precursors of the embryos (Tautorus et al. 1991). Apart from its biotechnology and economic importance, somatic embryogenesis is great model for investigating the events that regulate morphogenesis (Yang and Zhang 2010; Zimmerman 1993), eliminated some difficulties found in zygotic embryogenesis, which occurs inside maternal tissues and the embryos are difficult to access (Fehér et al. 2003).

Somatic embryogenesis has been applied in studies with several genotypes of *Carica papaya* L., an important tropical fruit crop with outstanding nutritional and medicinal values, considered an excellent model for tree fruit species (Ming et al. 2008; Aryal and Ming 2014; Fang et al. 2016). Although several somatic embryogenesis protocols for *C. papaya*have been published, there is little

information at the molecular level about somatic embryogenesis, being restricted only to the stage of maturation (Vale et al. 2014).

The transition from differentiated somatic cells to embryogenic cells requires rigorous biochemical coordination that includes several metabolic and signal transduction pathways (Fehér et al. 2003; Karami and Saidi 2010; Fehér 2015). As a result from the induction phase occurs the production of callus with different embryogenic competences (de Jong et al. 1993). Callus that are capable of differentiating and originating somatic embryos are called embryogenic, whereas those that are not able to differentiate, even undergoing maturation stimulus, are called non-embryogenic.In general, embryogenic cells are described as small, highly cytoplasmic, and mostly containing starch while in non-embryogenic callus it is common to observe cells vacuolated, and translucent in appearance (Pasternak et al. 2002). Understanding the fate and dynamics of cells during callus formation is essential to understanding cell totipotency and the somatic embryogenesis mechanisms.

A number of genes have been identified that play a significant role during cell division and cell wall formation at various stages of embryo differentiation including genes related to hormone response, homeobox genes, ABA-inducible genes, and housekeeping genes such as actin and tubulin (Chugh and Khurana 2002). Although there are several morphological and genetic markers for embryogenic cells, these markers are not universal, and most of the genes identified do not play a direct role in the induction of somatic embryogenesis (Namasivayam 2007; Karami et al. 2009).

An approach that may provide important information about the biochemical and physiological changes that occur during the acquisition of embryogenic capacity is proteomics, through the identification of differentially abundant proteins and their potential interactions (Rosas et al. 2016). Some studies using comparative proteomics approach have broadened our knowledge about the acquisition of embryogenic competence in several species, such as *A. angustifolia* (dos Santos et al. 2016); *E. guineensis* (de Carvalho Silva et al. 2014); *L. principis-rupprechtii* (Zhao et al. 2015) ; *Musa* spp (Kumaravel et al. 2017); *Saccharum* spp (Heringer et al. 2015); and *Z. mays* (Sun et al. 2013; Varhaníková et al. 2014).

However, proteomic characterization of subcellular compartments such as nucleus, mitochondria, chloroplast, and extracellular space (culture media) has not

yet been carried out. In addition, studies on post-transcriptional modifications and protein-protein interaction have not yet been addressed. That way, proteomics studies in somatic embryogenesis are still in their infancy (Rosas et al. 2016).

The present study used a comparative proteomic to identify and quantify differentially abundant proteins in embryogenic and non-embryogenic callus of papaya during multiplication phase, besides identifying the molecular mechanisms involved in the competence acquisition during early somatic embryogenesis development. Differentially abundant proteins provide an insight into embryogenic competence of papaya, and the development of competence molecular markers, which will facilitate the practical applications of somatic embryogenesis.

3.2.2. REVIEW

3.2.2.1. Papaya somatic embryogenesis

In somatic embryogenesis occurs the formation of somatic embryos from somatic cells. It was first described by three researchers independently, working with *Oenanthe aquatica* seedlings (Waris 1957) and carrot (Reinert 1958; Steward et al. 1958). Since then it has been widely used in several species.

Somatic embryogenesis is similar to zygotic, presenting the same stages of development in both dicotyledons and monocotyledons (Dodeman et al. 1997). Somatic embryogenesis represents the basis of cell totipotency, and unlike zygotic embryogenesis it can be controlled and more easily manipulated, making it an excellent model for studying the biochemical and molecular events that occur during the onset and development of embryogenesis in plants (Quiroz-Figueroa et al. 2006).

There are two ways to induce somatic embryogenesis, directly or indirectly. In direct somatic embryogenesis the formation of somatic embryos occurs directly in the explant, without the formation of callus, whereas in the indirect embryogenesis the dedifferentiation of the explant results in the callus that gives origin to the somatic embryo (Yang and Zhang 2010). Plant growth regulators are one of the main factors responsible for the induction of somatic embryogenesis (Jimenez

2005), especially auxin, which is considered the most important plant growth regulators in the regulation of somatic embryogenesis (Cooke et al. 1993).

In papaya the first work using somatic embryogenesis was published by Bruijne et al. (1974) using petiole as explant, however the regeneration of the plants originating from somatic embryos occurred only in 1977 using internodes as explant (Yie and Liaw 1977). From this, several studies were carried out using different genotypes with different types of explant (Yie and Liaw 1977; Litz and Conover 1980, 1982; Fitch 1993; Lin and Yang 2001; Anandan et al. 2012; Heringer et al. 2013).

Papaya somatic embryogenesis occurs with the use of a auxin such as 2,4-D, after induction and multiplication of the callus, which are submitted to maturation, which uses promoters of differentiation of somatic cells such as abscisic acid and polyethylene glycol (Dhekney et al. 2016). During this period the embryogenic cells differentiate into globular, cordiform, torpedo and cotyledonary, are germinated and can develop a complete plant (Fig. 1).

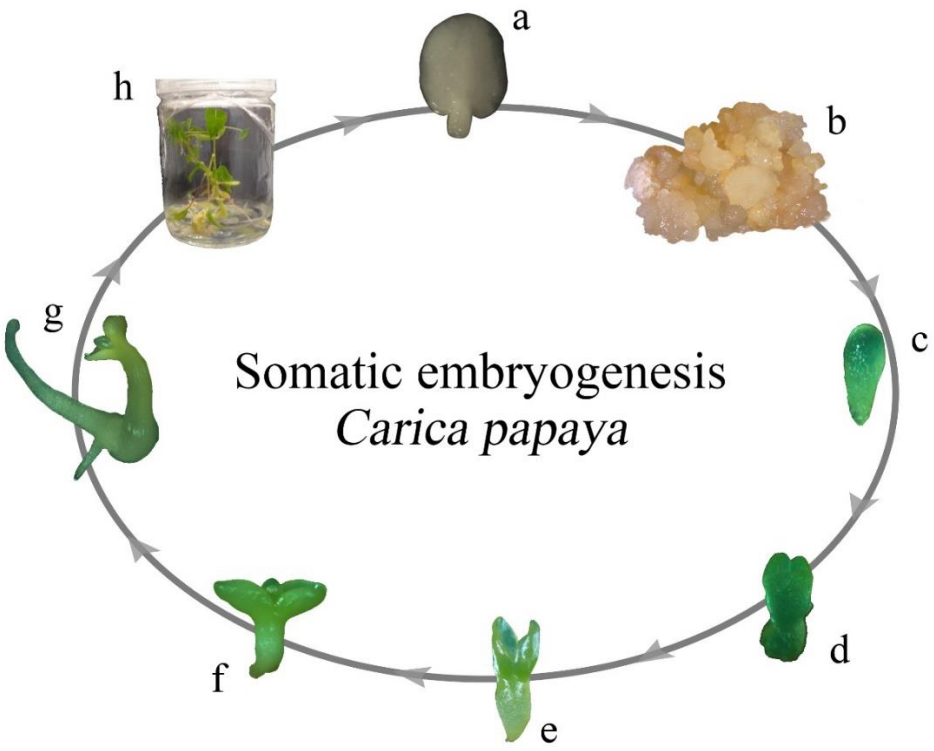


Fig. 1 Scheme of somatic embryogenesis of *Carica papaya* from explant zygotic embryos. Zygotic embryo (a), embryogenic callus (b), and the somatic embryo developmental stages: globular (c), cordiform (d), torpedo (e), cotyledonary (f), germinating embryo (g) and regenerated plantlets (h).

Somatic embryogenesis presents several applications in plant breeding and can contribute to the elimination of barriers found in conventional breeding (Raemakers et al. 1995). Some applications in papaya are the large-scale production of pathogen free plants, efficient conservation of germplasm (Castillo et al. 1998; Lu and Takagi 2000) and the amplification of the geometric variability through the somaclonal variants (Moore and Litz 1984; Clarindo et al. 2008; Homhuan et al. 2008). Moreover, this regeneration routeprovide asupport to other techniques such as introgression of genes (Chen et al. 1991), triploid production (Sun et al. 2011) and a genetic transformation (Gonsalves et al. 1997; Lines et al. 2002; Kung et al. 2010).

Despite the numerous studies proposing different protocols for the somatic embryogenesis of papaya, little is known about the biochemical and molecular processes that occur throughout the process. The first studies address the endogenous protein content in maturation (Heringer et al. 2013) and the proteomic profile of embryogenic callus in different maturation treatments (Vale et al. 2014). From this study important proteins were identified in the maturation of somatic embryos of papaya, such as ADH3, enolase and esterase.

Studies on the biochemical and molecular changes that occur throughout somatic embryogenesis, in addition to providing tools to optimize the conditions of cultivation and production of quality plants can also provide important information about one of the great questions of biology, the totipotency. In the process of somatic embryogenesis, the stage of callus induction and multiplication is a critical step for the somatic cells to acquire embryogenic competence. In papaya, there are no studies that point out the mechanisms responsible for the acquisition of embryogenic competence or markers that help to identify early embryogenic cells.

3.2.2.2. Acquisition of embryogenic competence

Totipotency is an important characteristic of plant cells, however, under specific conditions not all cells express this condition (Quiroz-Figueroa et al. 2006). Despite the six decades using the concept of totipotency in the induction and study of somatic embryogenesis, the mechanisms that ensure the acquisition of the embryogenic competence are still poorly understood. Studies on somatic

embryogenesis provide important information about the molecular and genetic mechanisms that govern the plasticity of plant development (Szyrajew et al. 2017).

In many species the somatic embryogenesiswascharacterized by two types of callus, which present differences in the development and can be differentiated through their cell morphology (Pasternak et al. 2002). The embryogenic callus are characterized by the small cells, highly cytoplasmic, and mostly containing starch while in non-embryogenic callus it is common to observe vacuolatedcells, and translucent in appearance (Silveira et al. 2013).

The acquisition of embryogenic competence depends on many factors, such as growth regulators, nutritional components of the culture medium, physical factors such as illumination, explant type, among others (Karami and Saidi 2010). This factors has been studied over the years, and provided empirical information on key points in the acquisition of this competence, how it can be variable among species, as well as a microscopic profile for embryogenic cultures of several species (Rosas et al. 2016).

The change in somatic cell destination is a dramatic event, requiring reprogramming for dedifferentiation and is dependent of a signal transduction program for the reorganization of a new cellular fate (Grafi and Barak 2015). Stress has been recognized for several years as an important factor in the change in cellular fate that leads to the differentiation and development of somatic embryos, but the mechanisms are still poorly understood (Fehér 2015). One of the hypotheses would be that the acquisition of embryogenic competence could be induced by stress through auxin interaction and stress signaling resulting in reprogramming of cells (Fehér et al. 2003; Karami and Saidi 2010). This process requires intense regulation of gene expression (Leljak-Levanić et al. 2004) and has been showed that plant response to stress promotes the activation of genes encoding transcription factors responsive to reprogramming these cells (Grafi and Barak 2015).

Reactive oxygen species have also shown their importance in plant growth and development, they act as signaling molecules and have crucial effects on cell fate (Kocsy et al. 2013). ROS homeostasis is maintained through enzymes such as superoxide dismutase, catalases (CATs), glutathione transferase (Pasternak et al. 2005). These enzymes are often associated with hormonal reactions and responses during plant development (Bartoli et al. 2013), and have also been considered key points in the acquisition of cell competence (Zhou et al. 2016).

In the last years, several genes even with differential expression has been correlated to somatic embryogenesis, or hasbeen related to the molecular events that occur along the somatic embryogenesis (Hu et al. 2005; Su et al. 2009; Yang et al. 2012). Among them, the WUSCHEL gene, an important regulator of the cell fate in theshoot meristem (Mayer et al. 1998), AGAMOUS-LIKE 15 and LEAFY COTYLEDON 1 and 2, transcription factors associated with seed development and maturation (Santos-Mendoza et al. 2008), BABY BOOM, DNA binding transcription factor activity, regulation of transcription (Kulinska-Lukaszek et al. 2012), and the SOMATIC EMBRYOGENESIS RECEPTOR KINASE (SERK) associated with brassinosteroid signaling and morphogenesis (Albrecht et al. 2008). These studies have expanded our understanding of cell totipotency and embryogenesis route; however, many points remain unclear.

Embryogenesis involves a multiplicity of molecular events that is not just the expression of differential genes but several signal transduction pathways to activate or repress numerous gene sets, many of which have not yet been identified and characterized (Chugh and Khurana 2002). Transducer proteins play a key role in somatic embryogenesis. Calcium-binding proteins such as calmodulin have demonstrated their potential role in hormone and stress-activated reprogramming of developmental pathways during somatic embryogenesis (Roberts and Harmon 1992). SERK proteins are the most studied signal transducers in somatic embryogenesis and have been used as markers of the embryogenic competence of several species (Schmidt et al. 1997; Steiner et al. 2012; aan den Toorn et al. 2015).

Taken together, alterations in chromatin by DNA methylation are fundamental in genomic reprogramming in somatic cells especially required to acquire embryogenic competence(Karami and Saidi 2010). There is extensive post-transcriptional regulation associated with the production of functional microRNAs that control the reprogramming of somatic cells into embryogenic cells (Szyrajew et al. 2017). MicroRNAs have been reported to modulate zygotic embryogenesis, however the influence of these molecules on somatic embryogenesis has been little studied (Wu et al. 2015).

Some extracellular proteins have also demonstrated an important role in the early stages of somatic embryogenesis and consequently in the acquisition of embryogenic competence (Chugh and Khurana 2002). Proteins such as arabinogalactans involved in cell proliferation and regulation of somatic embryo

develop (Pereira et al. 2014) and the lipid transfer proteins associated with the first differentiated tissue of somatic embryos(Dodeman et al. 1997).

In this sense, considereing that proteomics is a science that studies the proteins expressed by an organism, cell or tissue at a given time, under a specific condition (Wasinger et al. 1995), can provide important information about what in fact was expressed (Pennington and Dunn 2001). This finds contribute to a better understanding of cell biology, especially in embryogenesis route. Some studies have used proteomic analysis to look for differences that may help in understanding the factors that characterize embryogenic competence in different species (Rosas et al. 2016). These studies have shown increased abundance of stress-related proteins and reactive oxygen species in embryogenic cultures (Zhang et al. 2009; Zi et al. 2013; Zhou et al. 2016) which has been identified as markers for cell competence in several species (Teyssier et al. 2014). In addition, this tool includes many possibilities of analysis, still little used for this purpose, such as the proteomics of post-translational modifications, glycoproteomics, phosphoproteomics, lipidoproteomics, among others, capable of revealing more deeply the molecular mechanisms that occur during embryogenesis.

3.2.3. MATERIALS AND METHODS

3.2.3.1. Somatic embryogenesis induction

Based on different responses under maturation conditions embryogenic and non-embryogenic callus of *Carica papaya* genotype Golden were isolated from papaya zygotic embryos. Immature fruits of hermaphrodite plants were kindly provided by the Caliman Agricola S/A company, which is located in the city of Linhares, Espírito Santo (ES), Brazil (19° 23'S and 40° 4'W). The culture medium for callus induction and multiplication phase was performed according to Heringer et al. (2013). The MS culture medium (Murashige and Skoog 1962) (Phytotechnology Lab, Shawnee Mission, KS, USA) was supplemented with 3% sucrose (Sigma-Aldrich, St. Louis, MO, USA), 20 µM 2,4-dichlorophenoxyacetic acid (2,4-D) (Sigma-Aldrich) and 2.0 g/L phytagel (Sigma-Aldrich). The pH of the culture medium was adjusted to 5.8

before the phytagel was added. The culture medium was sterilized via autoclaving at 121°C for 15 min. For callus induction and multiplication, cultures were transferred to fresh medium every 21 days and grown at 25 °C with continuous darkness.

3.2.3.2. Multiplication and callus screening

During the multiplication perioda callus screening based on morphological differences allow us isolated and separated into: 1) compact and yellowish embryogenic callus and; 2) a mucilaginous and translucent non-embryogenic callus. To ensure true the callus response to somatic embryogenesis, we submitted and evaluated them through a maturation experiment according to Vale et al. (2014). During the test, embryogenic and non-embryogenic callus were inoculated in MS culture medium supplemented with 3% of sucrose (Sigma-Aldrich), 2.0 g/L phytagel (Sigma-Aldrich) and 6% of PEG 3350 (Sigma-Aldrich). The pH of theculture medium was adjusted to 5.8 before the Phytagelwas added. The culture medium was sterilized by autoclavingat 121°C for 15 min. The cultures were incubatedin a growth chamber at 25 ± 1°C in the dark for the first7 days, after which they were subjected to a 16 h light (60 µmol/m^2 s^1) photoperiod. The experiment was completely randomized, usingten repetitions per type of callus. Each replicate consisted of a Petri dish containing three callus, with a total of 30 colonies per treatment (The initial fresh matter (FM) was 300 mg per callus). After 28 days, cultureswere evaluatedfor the percentage of maturation and the increase of fresh matter. Callus maturation was considered embryogenic when the callus compact and yellowish formed at least one somatic embryo. Callus maturation was considered non-embryogenic when the callus a mucilaginous and translucent did not form any somatic embryo. After the screening the embryogenic and non-embryogenic callus kept in a medium of multiplication in the dark were collected for proteomic analysis (300 mg FM) and microscopy. The callus were collected on the 10th day of cultivation of the third cycle of subculture. Each replicate was composed of a petri dish containing five callus.

3.2.3.3. Microscopy analysis

To study the cell morphology of the embryogenic and non-embryogenic callus, samples were collected callus from the multiplicationmedium were fixed in

an aqueous solution containing glutaraldehyde (2.5%) and paraformaldehyde (4.0%) diluted in sodium cacodylate buffer (0.1 M), pH 7.3, at room temperature for 24 h. After fixation, the samples were dehydrated in a graded ethanol series and embedded in historesin (Leica, Wetzlar, Germany). Sections (approximately 5 µm thick) were stained with 1% toluidine blue (Sigma-Aldrich). Samples were examined under a Zeiss Axioplan light microscope (Carl Zeiss, Jena, Germany) equipped with an Axiocam MRC5 digital camera (Carl Zeiss) interfaced with the AxioVisionLE 4.8 software (Carl Zeiss) for image analysis.

3.2.3.4. Protein extraction

Protein extraction was carried out according toBalbuena et al. (2011). Each biological sample was grounded in a cold mortar using 1 mL of extraction buffer (7 M urea, 2 M thiourea, 2% triton X-100, 1% dithiothreitol - DTT, 1 mM phenylmethanesulfonyl fluoride - PMSF, and 5 µM pepstatin). Then, in microtubes samples were vortexed and incubated on ice for 30 min, followed by centrifugation at 16,000 g for 20min at 4 °C. The supernatants were collected and protein concentration was measured using a 2-D Quant Kit (GE Healthcare, Piscataway, NJ, USA).

3.2.3.5. Protein Digestion

For protein digestion, three biological replicates of 100 µg of proteins were used. Before the trypsin digestion step, protein samples were precipitated using the methanol/chloroform methodology to remove any detergent from samples (Nanjo et al. 2011). Then, samples were resuspended in Urea 7 M and Thiourea 2 M buffer, and desalted on Amicon Ultra-0.5 3 kDa centrifugal filters (Merck Millipore, Germany). Filters were filled to maximum capacity with buffers and centrifuged at 15,000 g for 20 min at 20 °C. The washes were performed twice with Urea 8 M and then twice with 50 mM ammonium bicarbonate (Sigma-Aldrich) pH 8.5, remaining approximately 50 µL per sample after the last wash.

The methodology used for protein digestion was as previously described (Calderan-Rodrigues et al. 2014). For each sample, 25 µL of 0.2% (v/v) RapiGest® (Waters, Milford, CT, USA) was added, and samples were briefly vortexed and

incubated in an Eppendorf Thermomixer® at 80 °C for 15 min. Then, 2.5 µL of 100 mM DTT (Bio-Rad Laboratories, Hercules, CA, USA) was added, and the tubes were vortexed and incubated at 60 °C for 30 min under agitation. Next, 2.5 µL of 300 mM iodoacetamide (GE Healthcare) was added, and the samples were vortexed and then incubated in the dark for 30 min at room temperature. The digestion was performed by adding 20 µL of trypsin solution (50 ng/µL; V5111, Promega, Madison, WI, USA) prepared in 50 mM ammonium bicarbonate, and samples were incubated at 37 °C overnight. For RapiGest® precipitation, 10 µL of 5% (v/v) trifluoroacetic acid (TFA, Sigma-Aldrich) was added and incubated at 37 °C for 90 min, followed by a centrifugation step of 30 min at 16,000 g. Samples were transferred to Total Recovery Vials (Waters).

3.2.3.6. Mass spectrometry analysis

A nanoAcquity UPLC connected to a Synapt G2-Si HDMS mass spectrometer (Waters, Manchester, UK) was used for ESI-LC-MS/MS analysis. The chromatography step was performed by injecting 1 µL of digested samples (500 ng/µL) to normalize them before the relative quantification of proteins. The proteins identified had their standardized molar values for all conditions, normalization among samples was based on stoichiometric measurements of total ion counts of scouting runs prior to analyses. Runs consisted of three biological replicates per callus type. During separation, samples were loaded onto the nanoAcquity UPLC 5 µm C18 trap column (180 µm × 20 mm) at 5 µL/min during 3 min and then onto the nanoAcquity HSS T3 1.8 µm analytical reversed phase column (75 µm × 150 mm) at 400 nL/min, with a column temperature of 45 °C. For peptide elution, a binary gradient was used, with mobile phase A consisting of water (Tedia, Fairfield, Ohio, USA) and 0.1% formic acid (Sigma-Aldrich) and mobile phase B consisting of acetonitrile (SigmaAldrich) and 0.1% formic acid. Gradient elution started at 7% B, then ramped from 7% B to 40% B up to 91.12 min, and from 40% B to 99.9% B until 92.72 min, being maintained at 99.9% until 106.00 min, then decreasing to 7% B until 106.1 min and kept 7% B until the end of experiment at 120.00 min. Mass spectrometry was performed in positive and resolution mode (V mode), 35,000 FWHM, with ion mobility, and in data-independent acquisition (DIA) mode; IMS wave velocity was set to 600 m/s; the transfer collision energy ramped from 19 V to

55 V in high-energy mode; cone and capillary voltages of 30 V and 2750 V, respectively; and a source temperature of 70 °C. In TOF parameters, the scan time was set to 0.5 s in continuum mode with a mass range of 50 to 2000 Da. The human [Glu1]-fibrinopeptide B (Sigma-Aldrich) at 100 fmol/µL was used as an external calibrant and lock mass acquisition was performed every 30 s.

3.2.3.7. Bioinformatics

Spectra processing and database searching conditions were performed by Progenesis QI for Proteomics Software V.2.0 (Nonlinear Dynamics, Newcastle, UK). The analysis used the following parameters: Apex3D of 150 counts for low energy threshold, 50 counts for elevated energy threshold, and 750 counts for intensity threshold; one missed cleavage, minimum fragment ion per peptide equal to two, minimum fragment ion per protein equal to five, minimum peptide per protein equal to two, fixed modifications of carbamidomethyl (C) and variable modifications of oxidation (M) and phosphoryl (STY), and a default false discovery rate (FDR) value at a 4% maximum, score greater than five, and maximum mass errors of 10 ppm. The analysis used the *Carica papaya* protein databank from PhytozomeV11.0 (https://phytozome.jgi.doe.gov/pz/portal.html) (27,793 sequences, Aug 2016). Label-free relative quantitative analyses were performed based on the ratio of protein ion counts among contrasting samples. After data processing and to ensure the quality of results, the following protein refinement parameters were used: only proteins present or absent (for unique proteins) in 3 of 3 runs. Furthermore, differentially abundant proteins were selected based on a max fold change of at least 2.0 and ANOVA $P < 0.05$. Functional annotation was performed using Blast2Go software v. 3.4 (Conesa, 2005). BLAST was performed using the following settings: blast DB, Plants/Arabidopsis_thaliana_protein_sequences; number of blast hits, 20; blast E cut-off, 1.0×10^{-6}; blast program, BLASTP; high-scoring segment pair length cut-off, 33; low complexity filter.

3.2.4. RESULTS

3.2.4.1. Morphological and cellular characterization

The morphological evaluation of *C. papaya* genotype Golden showed non-embryogenic callus translucent, mucilaginous and watery, in contrast to the embryogenic callus smooth and compact (Fig. 1). Moreover, embryogenic callus were to somatic embryogenesis responsiveness while non-embryogenic callus showed only an unorganized cell growth (Table1, Fig. 1). Both of them exhibit actively proliferation (Fig. 1) especially non-embryogenic callus, which are incapable to form somatic embryos, which had a greater increase in the FM (Table 1). By the histological analyses it was identified that embryogenic callus were mainly compost by a high proportion of small isodiametric cells with dense cytoplasm and large nuclei, organized in meristematic aggregates. In the other hand, non-embryogenic callus presented elongated and dispersed cells, with large vacuoles and large intercellular spaces (Fig. 1). In addition, no nucleous was observed in the most part of elongated cells. In embryogenic callus, toluidine blue treated cells showed an orthochromatic reaction in the cytoplasm and a metachromatic reaction in the cell wall, indicating the presence of acidic polysaccharides.

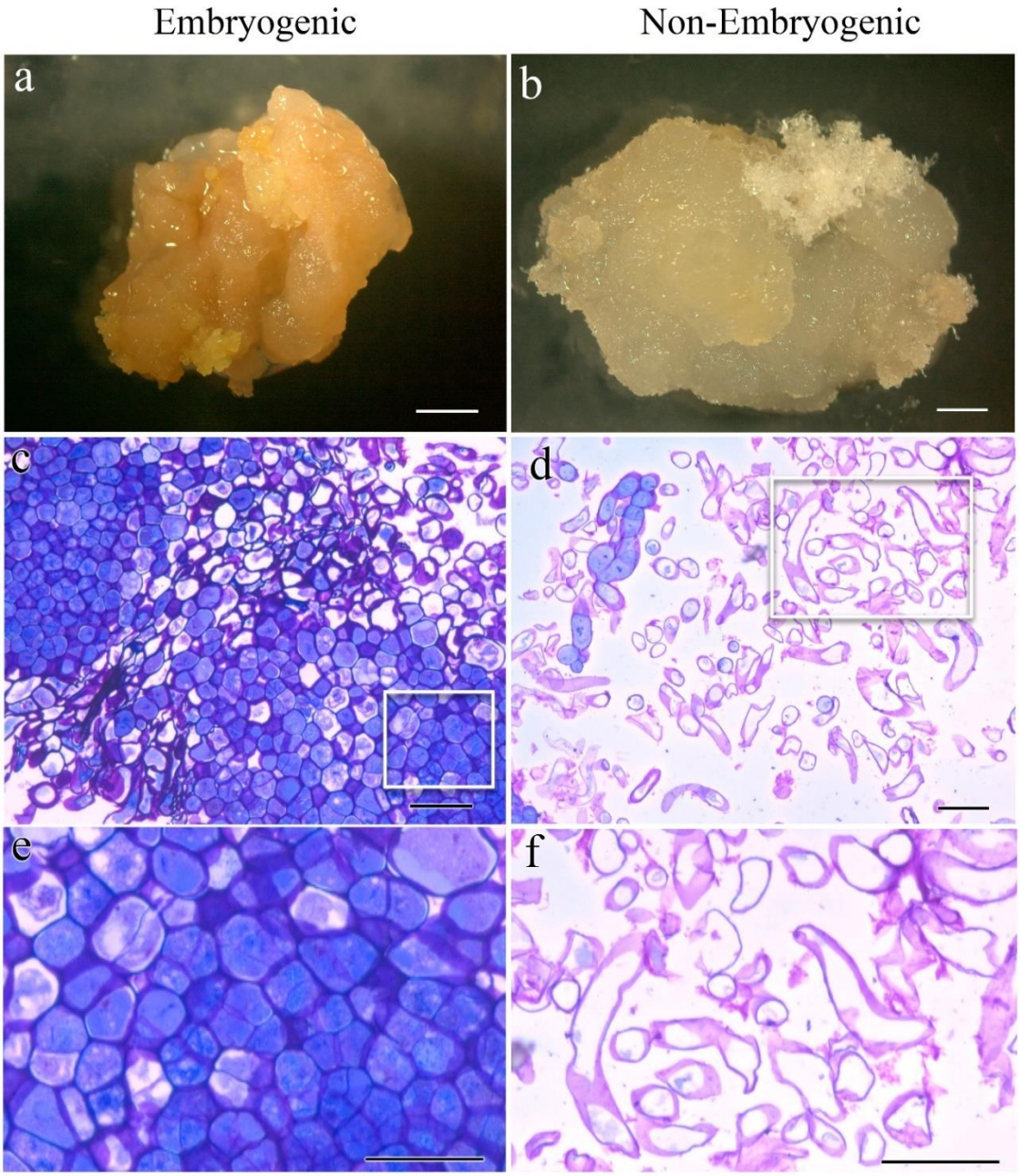


Fig. 1 Morphological and cellular charactherization of embryogenic and non-embryogenic callus of *C. papaya* Golden. (a, c and e) embryogenic callus, (b, d and f) non-embryogenic callus. Bars: a andb = 0.5 mm; c-f = 25 µm.

Table 1. Percentage of responsiveness callus to somatic embryos development in maturation medium and fresh matter increase of *C. papaya* Golden after 28 days incubated.

Callus type	Maturation (%)	Fresh matter (mg)	Protein (µg/g)
Embryogenic	100.0 a*	380 b	84.2 a
Non-embryogenic	0.0 b	560 a	35.1 b

*Means followed by the same letters are not significantly different according to the Kruskal-Wallis ($P < 0.01$) for % callus responsiveness to somatic embryos development maturation and SNK test ($P < 0.05$) for FM and protein. n = 10; FM: C.V. = 2.5%; Protein C.V. =13.6%, n= 3.

3.2.4.2. Protein identification by LC–MS/MS

A total of 668 proteins were identified in embryogenic and non-embryogenic callus of papaya. Subsequent comparative analysis of samples from the embryogenic and non-embryogenic showed that 155 proteins were differentially abundant, with a significant change ($P < 0.05$) in quantity of more than 2.0-fold (Table 2). Comparing the abundance of the embryogenic callus proteins in relation to the non-embryogenic callus, 82 proteins were up-regulated and 73 were down-regulated.

To better understanding the differences between embryogenic and non-embryogenic cells, the identified proteins were categorized using Blast2GO analysis. Functional annotation was divided between the two groups, up- and down-regulated proteins, from embryogenic in relation to the non-embryogenic. These proteins were functionally classified according to their biological functions and the gene ontology of the biological process. The most representative classes of proteins for both groups were associated with response to abiotic stimulus, stress, developmental process, hormone, regulation of biological and catabolic process, lipid, cell morphogenesis and reactive oxygen species (Fig. 2).

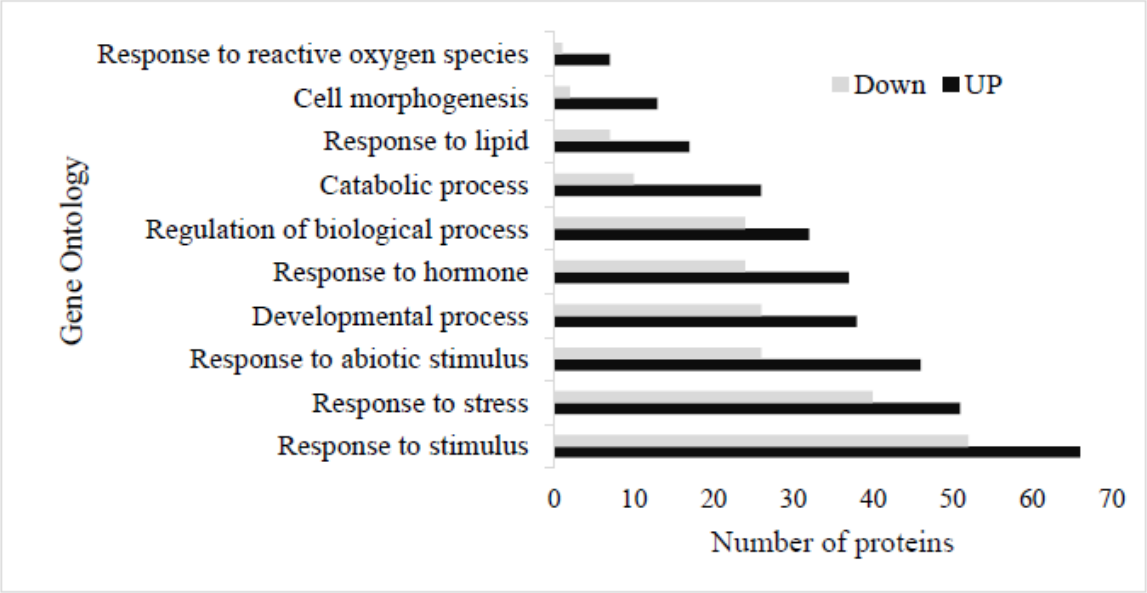


Fig. 2 Functional classification of up and down regulated proteins in embryogenic and non-embryogenic callus of *C. papaya*

Table 2. Differentially abundant proteins identified in embryogenic versus non-embryogenic callus of *C. papaya* Golden.

Accession[1]	Protein name[2]	Biological function[2]	Peptide count	Score	Anova[4]	Max fold change[3]	Regulation
evm.TU.contig_26338.1	26S proteasome regulatory non-ATPase Rpn2 Psmd1 subunit	Regulation of biological process	16	99,08	0.0006	2,1	UP
evm.model.supercontig_113.54	60S Ribosomal protein L13	Gene expression	2	12,50	0.0001	31,4	Down
evm.model.supercontig_56.23	AAA- CDC48	Response to stimulus	15	114,79	0.0169	2,4	Down
evm.model.supercontig_1.364	actin depolymerizing factor 1	Response to stimulus	4	39,26	0.0001	3,6	Down
evm.model.supercontig_139.79	acyl- oxidase 2	Regulation of biological process	3	14,97	0.0491	2,0	UP
evm.model.supercontig_85.4	Adenine nucleotide alpha hydrolases-like superfamily	Response to stress	2	10,94	0.0000	10,7	UP
evm.model.supercontig_505.5	AGAMOUS-like 62	Developmental process	3	15,14	0.0178	6,6	Down
evm.model.supercontig_567.1	Agglutinin	Cell division	11	59,70	0.0000	158,3	UP
evm.model.supercontig_83.14	Agglutinin	Cell division	12	64,80	0.0001	6,5	UP
evm.model.supercontig_83.17	Agglutinin	Cell division	32	278,43	0.0000	26,5	UP
evm.model.supercontig_42.139	Alcohol dehydrogenase	Developmental process	7	39,15	0.0000	33,5	Down
evm.model.supercontig_6.184	Alpha beta-Hydrolases superfamily	Response to abiotic stimulus	2	16,26	0.0003	14,7	Down
evm.model.supercontig_179.10	Alpha beta-Hydrolases superfamily	Response to abiotic stimulus	3	15,50	0.0002	6,0	UP
evm.model.supercontig_18.158	Alpha beta-Hydrolases superfamily	Response to abiotic stimulus	2	11,10	0.0049	14,4	UP
evm.model.supercontig_1.156	Alpha-1,4-glucan- synthase family	Response to abiotic stimulus	17	112,79	0.0169	5,6	UP

Table 2 – Cont.

Accession[1]	Protein name[2]	Biological function[2]	Peptide count	Score	Anova[4]	Max fold change[3]	Regulation
evm.model.supercontig_79.38	AMP-dependent synthetase and ligase family	Developmental process	2	10,95	0.0001	3,2	Down
evm.model.supercontig_40.66	ARM repeat superfamily	Response to stimulus	2	13,71	0.0000	16,8	UP
evm.model.supercontig_20.227	Aspartate aminotransferase	Developmental process	6	42,17	0,0010	3,1	Down
evm.model.supercontig_165.35	ATP citrate lyase (ACL) family	Response to stimulus	12	54,52	0.0000	3,0	Down
evm.model.supercontig_292.1	Auxin-responsive GH3 family	Response to hormone	9	41,39	0.0007	133,4	UP
evm.model.supercontig_6.74	Auxin-responsive GH3 family	Response to hormone	34	225,20	0.0000	12,4	UP
evm.model.supercontig_0.14	beta-D-xylosidase 4	Regulation of biological process	12	66,81	0.0000	3,8	Down
evm.model.supercontig_12.222	Bifunctional inhibitor lipid-transfer seed storage 2S albumin superfamily	Response to lipid	3	24,47	0.0035	10,2	Down
evm.model.supercontig_157.25	Bifunctional inhibitor lipid-transfer seed storage 2S albumin superfamily	Response to lipid	10	67,43	0.0000	17,9	UP
evm.model.supercontig_157.26	Bifunctional inhibitor lipid-transfer seed storage 2S albumin superfamily	Response to lipid	8	40,21	0.0005	12,2	UP
evm.model.supercontig_157.28	Bifunctional inhibitor lipid-transfer seed storage 2S albumin superfamily	Response to lipid	9	47,57	0.0008	7,9	UP
evm.model.supercontig_20.157	Bifunctional inhibitor lipid-transfer seed storage 2S albumin superfamily	Response to lipid	4	28,79	0.0000	99,5	UP
evm.model.supercontig_2639.1	Bifunctional inhibitor lipid-transfer seed storage 2S albumin superfamily	Response to lipid	7	37,33	0.0001	6,0	UP

Table 2 – Cont.

Accession[1]	Protein name[2]	Biological function[2]	Peptide count	Score	Anova[4]	Max fold change[3]	Regulation
evm.model.supercontig_808.3	Bifunctional inhibitor lipid-transfer seed storage 2S albumin superfamily	Response to lipid	7	39,69	0.0000	5,7	UP
evm.model.supercontig_129.58	Branched-chain amino acid transaminase 2	Biosynthetic process	3	19,10	0.0023	2,0	Down
evm.model.supercontig_32.51	Breast basic conserved 1	Response to stimulus	2	10,97	0.0060	9,8	Down
evm.TU.contig_27001.2	CAP (Cysteine-rich secretory Antigen and Pathogenesis-related 1) superfamily	Response to stimulus	4	24,45	0.0003	2,9	Down
evm.model.supercontig_770.2	Carbonic anhydrase 1	Response to stimulus	28	154,17	0.0000	8,8	Down
evm.model.supercontig_23.10	Cation transporter E1-E2 ATPase family	Response to stimulus	4	19,93	0.0290	2,9	UP
evm.model.supercontig_107.112	Chaperone htpG family	Response to stresse	25	140,19	0.0002	2,8	UP
evm.model.supercontig_184.10	Chitinase A	Response to stresse	24	153,46	0.0002	4,8	Down
evm.model.supercontig_29.122	Cupredoxin superfamily	Cell morphogenesis	6	32,51	0.0001	3,1	Down
evm.model.supercontig_37.186	Cyclophilin-like peptidyl-prolyl cis-trans isomerase family	Response to stresse	12	48,98	0.0034	2,5	UP
evm.model.supercontig_228.9	Cystatin B	Response to stresse	6	41,96	0.0004	2,4	UP
evm.model.supercontig_55.102	Cysteine ases superfamily	Developmental process	2	11,13	0.0005	3,9	UP
evm.model.supercontig_11.22	Cysteine synthase D2	Cell morphogenesis	2	11,57	0.0001	6,3	UP
evm.model.supercontig_9.380	Cysteine synthase D2	Cell morphogenesis	8	45,62	0.0073	2,9	UP
evm.model.supercontig_33.52	Cytochrome family subfamily polypeptide 35	Response to stimulus	11	60,62	0.0194	2,8	UP

Table 2 – Cont.

Accession[1]	Protein name[2]	Biological function[2]	Peptide count	Score	Anova[4]	Max fold change[3]	Regulation
evm.model.supercontig_259.6	DEA(D H)-box RNA helicase family	Gene expression	7	32,56	0.0005	4,2	UP
evm.model.supercontig_59.101	Dehydratase family	Response to stimulus	7	30,73	0.0131	2,9	UP
evm.model.supercontig_16.196	Delta(3,5),delta(2,4)-dienoyl- isomerase 1	Developmental process	17	83,04	0,0043	2,9	UP
evm.model.supercontig_1199.1	Dihydrolipoamide succinyltransferase	Response to stimulus	8	38,18	0.0336	2,1	Down
evm.model.supercontig_97.106	D-mannose binding lectin with Apple-like carbohydrate-binding domain-containing	Response to hormone	4	12,90	0.0005	2,3	Down
evm.model.supercontig_97.109	D-mannose binding lectin with Apple-like carbohydrate-binding domain-containing	Response to hormone	9	58,78	0.0002	5,8	Down
evm.model.supercontig_3.274	Enolase 1	Carbohydrate metabolic process	22	181,41	0.0008	2,5	Down
evm.model.supercontig_72.54	Essential meiotic endonuclease 1A	Gene expression	2	10,86	0.0012	2,3	UP
evm.model.supercontig_40.16	Eukaryotic aspartyl protease family	Response to stimulus	6	33,47	0.0014	2,3	Down
evm.model.supercontig_13.23	Eukaryotic aspartyl protease family	Response to stimulus	10	70,93	0.0335	2,3	Down
evm.TU.contig_31827.1	Eukaryotic aspartyl protease family	Response to stimulus	12	83,20	0.0004	6,8	Down
evm.model.supercontig_152.35	Eukaryotic aspartyl protease family	Response to stimulus	14	84,83	0.0000	4,9	UP
evm.model.supercontig_6.305	Eukaryotic aspartyl protease family	Response to stimulus	5	31,77	0.0043	4,1	UP
evm.model.supercontig_5.123	Eukaryotic translation initiation factor 4A1	Gene expression	7	36,83	0.0024	2,3	Down
evm.model.supercontig_677.7	FAD-binding Berberine family	Response to stimulus	4	21,29	0.0184	2,1	Down

Table 2 – Cont.

Accession[1]	Protein name[2]	Biological function[2]	Peptide count	Score	Anova[4]	Max fold change[3]	Regulation
evm.model.supercontig_677.4	FAD-binding Berberine family	Response to stimulus	10	52,43	0.0000	4,0	Down
evm.model.supercontig_677.5	FAD-binding Berberine family	Response to stimulus	27	150,57	0.0000	3,7	Down
evm.model.supercontig_165.4	FASCICLIN-like arabinogalactan 6	Cellular process	3	28,09	0.0072	3,6	Down
evm.model.supercontig_19.67	Fe superoxide dismutase 1	Response to reactive oxygen species	2	11,95	0.0299	5,2	UP
evm.model.supercontig_21.21	FUMARASE 2	Cellular process	4	20,47	0.0000	2,4	UP
evm.model.supercontig_130.78	GDSL-like Lipase Acylhydrolase superfamily	Metabolic process	3	26,84	0.0001	3,5	Down
evm.model.supercontig_39.80	GDSL-like Lipase Acylhydrolase superfamily	Metabolic process	8	60,82	0.0001	3,7	Down
evm.model.supercontig_17.160	GDSL-like Lipase Acylhydrolase superfamily	Metabolic process	8	53,27	0.0001	3,7	UP
evm.model.supercontig_25.179	GDSL-like Lipase Acylhydrolase superfamily	Metabolic process	18	128,91	0.0001	12,1	UP
evm.model.supercontig_25.180	GDSL-like lipase acylhydrolase superfamily	Metabolic process	53	182,94	0.0000	14,7	UP
evm.model.supercontig_179.17	General regulatory factor 3	Regulation of biological process	19	93,81	0.0000	2,2	Down
evm.model.supercontig_55.36	General regulatory factor 9	Regulation of biological process	18	68,07	0.0000	2,3	UP
evm.TU.contig_28829.1	Glucoside glucohydrolase 2	Response to lipid	7	38,74	0.0144	21,7	UP
evm.TU.contig_29746.1	Glucoside glucohydrolase 2	Response to lipid	11	66,59	0.0000	37,2	UP
evm.model.supercontig_74.21	Glutathione peroxidase 6	Response to reactive oxygen species	10	63,64	0.0004	5,2	UP

Table 2 – Cont.

Accession[1]	Protein name[2]	Biological function[2]	Peptide count	Score	Anova[4]	Max fold change[3]	Regulation
evm.model.supercontig_50.116	Glutathione S-transferase F4	Response to reactive oxygen species	5	41,26	0.0004	2,5	UP
evm.model.supercontig_9.201	Glutathione S-transferase F4	Response to reactive oxygen species	25	131,14	0.0010	5,0	UP
evm.model.supercontig_1.329	Glutathione S-transferase family	Response to reactive oxygen species	8	48,10	0.0002	2,5	UP
evm.model.supercontig_1.380	Glutathione S-transferase TAU 18	Response to reactive oxygen species	6	50,70	0.0043	3,3	UP
evm.model.supercontig_73.70	Glutathione S-transferase tau 7	Response to reactive oxygen species	4	29,00	0.0003	2,0	UP
evm.model.supercontig_73.67	Glutathione S-transferase TAU 8	Response to reactive oxygen species	13	109,33	0.0001	3,0	UP
evm.model.supercontig_130.34	Glyceraldehyde 3-phosphate dehydrogenase A subunit 2	Carbohydrate metabolic process	40	212,19	0.0003	2,1	UP
evm.model.supercontig_51.147	Glycosyl hydrolase family	Carbohydrate metabolic process	8	58,42	0.0002	2,4	Down
evm.model.supercontig_200.16	Granulin repeat cysteine protease family	Response to stimulus	6	36,98	0.0000	4,0	Down
evm.model.supercontig_1377.1	H[+]-ATPase 2	Response to stimulus	14	85,32	0.0000	2,6	Down
evm.TU.contig_34651.1	Heat shock 70 (Hsp 70) family	Response to stress	6	34,31	0.0002	2,5	Down
evm.model.supercontig_17.178	Hipl2 precursor	Carbohydrate metabolic process	3	22,86	0.0000	3,8	Down
evm.model.supercontig_92.72	HIS triad family 3	Cellular process	2	12,34	0.0001	207,3	UP

Table 2 - Cont

Accession[1]	Protein name[2]	Biological function[2]	Peptide count	Score	Anova[4]	Max fold change[3]	Regulation
evm.model.supercontig_728.2	Histidine-containing phosphotransmitter 2	Developmental process	3	19,99	0.0001	2,6	Down
evm.model.supercontig_10.53	HXXXD-type acyl-transferase family	Developmental process	2	17,71	0.0416	3,1	Down
evm.model.supercontig_12.233	IAA carboxylmethyltransferase 1	Response to hormone	3	22,41	0.0002	25,8	Down
evm.model.supercontig_118.1	Isocitrate dehydrogenase	Response to stress	6	37,68	0.0030	4,3	UP
evm.model.supercontig_52.151	Isopropylmalate dehydrogenase 1	Response to stimulus	4	20,84	0.0019	6,6	UP
evm.model.supercontig_216.20	Late embryogenesis abundant (LEA)	Unidentified	4	24,66	0.0000	71,3	UP
evm.model.supercontig_209.19	late embryogenesis abundant domain-containing LEA domain-containing	Unidentified	6	30,19	0.0040	8,7	UP
evm.model.supercontig_62.100	like cupins superfamily	Response to stimulus	15	109,14	0.0000	9,4	UP
evm.model.supercontig_283.2	Lipase PLAT LH2 family	Response to stress	14	67,04	0.0004	7,7	Down
evm.model.supercontig_5.334	Matrixin family	Response to stimulus	2	12,70	0.0001	6,7	Down
evm.model.supercontig_1008.1	NAD(P)-binding Rossmann-fold superfamily	Response to stimulus	5	31,20	0.0115	2,2	Down
evm.model.supercontig_540.2	NADPH:quinone oxidoreductase	Oxidation-reduction process	6	37,68	0.0003	2,2	Down
evm.model.supercontig_1815.1	NAP1-related 1	Developmental process	4	29,42	0.0003	2,1	Down
evm.model.supercontig_29.25	Nucleotide-rhamnose synthase epimerase-reductase	Oxidation-reduction process	6	35,57	0.0006	2,9	Down
evm.model.supercontig_120.1	O-Glycosyl hydrolases family 17	Carbohydrate metabolic process	9	79,75	0.0010	2,3	Down

Table 2 – Cont.

Accession[1]	Protein name[2]	Biological function[2]	Peptide count	Score	Anova[4]	Max fold change[3]	Regulation
evm.model.supercontig_1476.3	Pathogenesis-related 4	Response to stress	2	12,85	0.0037	4,2	Down
evm.model.supercontig_1476.2	Pathogenesis-related 4	Response to stress	6	46,74	0.0002	4,4	Down
evm.model.supercontig_12.208	Pathogenesis-related thaumatin superfamily	Response to stress	15	120,29	0.0001	4,1	Down
evm.model.supercontig_190.4	Pathogenesis-related thaumatin superfamily	Response to stress	2	11,72	0.0000	14,3	UP
evm.model.supercontig_143.7	Pectinacetylesterase family	Cellular process	4	19,82	0.0227	2,2	Down
evm.model.supercontig_42.86	Peroxidase superfamily	Response to stress	17	90,94	0.0004	5,7	Down
evm.model.supercontig_42.87	Peroxidase superfamily	Response to stress	15	133,97	0.0002	8,5	Down
evm.model.supercontig_119.100	Peroxidase superfamily	Response to stress	3	35,16	0.0007	2,4	UP
evm.model.supercontig_19.93	Peroxidase superfamily	Response to stress	4	24,75	0.0006	5,2	UP
evm.model.supercontig_116.34	pfkB-like carbohydrate kinase family	Carbohydrate metabolic process	13	49,74	0.0001	3,3	Down
evm.model.supercontig_9.297	pfkB-like carbohydrate kinase family	Carbohydrate metabolic process	17	123,52	0.0066	2,4	Down
evm.model.supercontig_13.7	Phosphatase 2A subunit A2	Response to stimulus	8	41,00	0.0003	5,8	UP
evm.model.supercontig_52.175	Phosphoenolpyruvate carboxylase 1	Response to stimulus	15	90,19	0.0001	2,4	UP
evm.model.supercontig_170.35	Phosphomannomutase		4	23,14	0.0072	7,7	Down
evm.model.supercontig_42.121	Plant invertase pectin methylesterase inhibitor superfamily	Carbohydrate metabolic process	3	17,99	0.0474	42,5	Down

Table 2 – Cont.

Accession[1]	Protein name[2]	Biological function[2]	Peptide count	Score	Anova[4]	Max fold change[3]	Regulation
evm.model.supercontig_151.18	Plant invertase pectin methylesterase inhibitor superfamily	Carbohydrate metabolic process	3	23,27	0.0010	9,9	Down
evm.model.supercontig_48.180	Plant/protein (Protein of unknown function, DUF538)	Response to stress	2	12,80	0.0000	60,5	UP
evm.model.supercontig_30.65	Plasma-membrane associated cation-binding 1	Response to stress	9	70,45	0.0003	4,2	Down
evm.TU.contig_30095.2	Polyketide cyclase dehydrase and lipid transport superfamily	Response to lipid	10	61,73	0.0000	3,4	UP
evm.model.supercontig_8.291	Pyridoxal phosphate (PLP)-dependent transferases superfamily	Carbohydrate metabolic process	6	35,67	0.0018	4,3	UP
evm.model.supercontig_57.76	Pyrophosphorylase 1	Response to stimulus	3	18,58	0.0047	2,3	UP
evm.model.supercontig_188.19	Pyrophosphorylase 4	Response to stimulus	2	11,36	0.0150	7,2	UP
evm.model.supercontig_163.30	RAB GTPase homolog 8	Response to stimulus	5	20,77	0.0499	16,9	Down
evm.model.supercontig_131.82	RAB GTPase homolog A1F	Response to stimulus	5	25,46	0.0056	11,3	UP
evm.model.supercontig_5.161	RAB GTPASE HOMOLOG B18	Response to stimulus	3	11,66	0.0266	9,5	Down
evm.model.supercontig_22.63	Ras-related small GTP-binding family	Cell Morphogenesis	12	57,35	0.0013	8,2	Down
evm.model.supercontig_52.66	Ras-related small GTP-binding family	Cell Morphogenesis	7	34,50	0.0008	3,5	UP
evm.model.supercontig_117.78	Ribose 5-phosphate type A	Developmental process	4	29,14	0.0207	2,1	Down
evm.model.supercontig_13.286	Ribosomal L23 L15e family	Gene expression	4	19,73	0.0019	2,2	Down

Table 2 – Cont.

Accession[1]	Protein name[2]	Biological function[2]	Peptide count	Score	Anova[4]	Max fold change[3]	Regulation
evm.model.supercontig_165.7	RmlC-like cupins superfamily	Response to hormone	15	95,31	0.0000	10,8	UP
evm.model.supercontig_165.8	RmlC-like cupins superfamily	Response to hormone	6	40,52	0.0000	7,4	UP
evm.model.supercontig_54.82	RmlC-like cupins superfamily	Response to hormone	13	70,14	0.0000	7,0	UP
evm.model.supercontig_58.77	RmlC-like cupins superfamily	Response to hormone	2	11,44	0.0001	89,6	UP
evm.model.supercontig_95.9	RNA-binding KH domain-containing	Gene expression	2	11,36	0.0028	3,2	UP
evm.model.supercontig_5.200	S-adenosyl-L-methionine-dependent methyltransferases superfamily	Developmental process	5	27,10	0.0009	16,7	Down
evm.TU.contig_39307.1	Sec14p-like phosphatidylinositol transfer family	Developmental process	12	100,90	0.0014	2,6	Down
evm.model.supercontig_53.140	Serine carboxypeptidase-like 12	Protein metabolic process	3	15,50	0.0003	3,1	UP
evm.model.supercontig_77.85	SOUL heme-binding family	Unidentified	2	16,75	0.0002	6,8	UP
evm.model.supercontig_64.102	Stromal ascorbate peroxidase	Oxidation-reduction process	4	28,53	0.0001	3,6	Down
evm.model.supercontig_6.53	Subtilase family	Developmental process	26	178,00	0.0001	2,4	Down
evm.model.supercontig_108.71	Sucrase ferredoxin-like family		5	28,17	0.0215	2,2	Down
evm.model.supercontig_43.88	Target of Myb 1	Establishment of localization	2	11,74	0.0026	3,8	Down
evm.model.supercontig_8.201	TCP-1 cpn60 chaperonin family	Response to stresse	30	162,79	0.0056	2,7	UP
evm.model.supercontig_97.56	Thiazole biosynthetic chloroplast (ARA6) (THI1) (THI4)	Oxidation-reduction process	3	14,98	0.0012	2,4	UP

Table 2 – Cont.

Accession[1]	Protein name[2]	Biological function[2]	Peptide count	Score	Anova[4]	Max fold change[3]	Regulation
evm.model.supercontig_146.45	Thioredoxin superfamily	Developmental process	13	77,69	0.0000	10,0	UP
evm.model.supercontig_91.78	Transmembrane protein	Developmental process	2	15,17	0.0014	6,1	UP
evm.model.supercontig_2.74	tRNA synthetase class I (M and V) family	Gene expression	9	51,05	0.0007	2,9	Down
evm.model.supercontig_83.83	Tyrosine transaminase family	Regulation of biological process	16	106,18	0.0000	5,9	UP
evm.model.supercontig_13.153	Tyrosyl-tRNA class bacterial mitochondrial	Gene expression	3	14,87	0.0038	4,2	UP
evm.model.supercontig_97.98	Ubiquitin conjugating enzyme 8	Cell differentiation	3	23,05	0.0023	3,4	UP
evm.model.supercontig_81.53	UDP-Glycosyltransferase superfamily	Response to hormone	2	10,45	0.0091	89,3	Down
evm.model.supercontig_216.23	Undecaprenyl pyrophosphate synthetase family	Unidentified	3	15,18	0.0214	2,0	UP
evm.model.supercontig_176.21	Vacuolar H+-pyrophosphatase 2	Developmental process	4	15,06	0.0002	11,1	Down
evm.model.supercontig_98.56	Vacuolar protein sorting-associated protein 13	Unidentified	3	16,03	0.0002	5,9	UP

[1] Phytozome gene identification number.

[2] Protein name and biological function from Blast2GO.

[3] Protein relative abundance between embryogenic and non-embryogenic callus

[4] Analysis of variance (ANOVA) based p value.

3.2.5. DISCUSSION

The differences between embryogenic and non-embryogenic callus of papaya are characterized by large morphological changes and a differential regulation in protein abundance. According to the histomorphological analyzes, the embryogenic callus presented small isodiametric cells and prominent nucleus, whereas the cells of the non-embryogenic callus were elongated, with large vacuoles that were dispersed along the callus (Fig 1). These differences confirm the characterization of embryogenic and non-embryogenic callus reported in other species (Verdeil et al. 2001; Pasternak et al. 2002; Silveira et al. 2013).

These differences in competence between induced cells can be explained by the role of regulators in inducing stress responses and these responses play an essential role in reprogramming patterns of gene expression (Fehér et al. 2003; Karami and Saidi 2010). The maintenance of the embryogenic potential requires the activity of signaling of genetic pathways, which in turn lead to specific responses, such as cytoplasmic remodeling, changes in patterns of cell division and differentiation (Smertenko and Bozhkov 2014).

Stress is widely characterized as fundamental for the induction of ES, and growth regulators such as 2,4-D have an important role in inducing stress for the acquisition of embryogenic competence which is characterized by extensive cellular reprogramming manifested at different levels (Karami and Saidi 2010). The potential biological significance and associated biochemical functions that were differentially abundant between the embryogenic and non-embryogenic callus, is discussed according with five functional classes that were most representative: Stress and detoxification-related proteins, response to lipid, auxin regulation, gene regulation, and proteins not associated with the selected functional groups.

3.2.5.1. Stress and detoxification-related proteins

Optimal balances of ROS are achieved by multiple antioxidative systems in plants. In embryogenic callus of papaya, several proteins with antioxidant function were up-regulated, among them, different classes of Glutathione S-transferase (glutathione S-transferase F4, glutathione S-transferase TAU 8, and glutathione S-

transferase tau 7), Fe superoxide dismutase 1 and glutathione peroxidase 6 (Table 2). The role of stress response proteins during the acquisition of embryogenic competence has been demonstrated in several studies (Takac et al. 2011). In addition, these proteins too were important for the maturation of somatic embryos (Vale et al. 2014), demonstrating its role during all phases of somatic embryogenesis of this species.

The importance of stress in somatic embryogenesis is closely related to the production of ROS in cell homeostasis (Elhiti and Stasolla 2015). The greater abundance of proteins related to response to reactive oxygen species in the embryogenic callus demonstrates that the acquisition of the embryogenic competence of the embryogenic callus of papaya depends on the ability to maintain cellular homeostasis. The Glutathione S-transferase family (GST), have been associated with antioxidative functions in plants (Noctor et al. 2012) and play a role in the defense mechanism to protect plants from cell damage (Marrs 1996) and can also act as auxin-binding proteins (Bilang et al. 1993). In species such as *Cichorium* (Galland et al. 2001) and *D. carota* (Kitamiya et al. 2000) the formation of somatic embryos was associated with the expression of a GST. Isoforms have also been described in the embryogenic callus proteomes of *V. vinifera* (Marsoni et al. 2008), *P. americana*(Guzmán-García et al. 2013), *T. cacao* (Noah et al. 2013) and *A. angustifolia* (dos Santos et al. 2016). Additionally roles of GST genes during acquisition of embryogenic potential are likely to be associated with protecting the cell against the harmful effects of excessive amount of auxin (Karami and Saidi 2010).

The glutathione peroxidase 6 is the important enzyme for protection of cell membranes from peroxidative damage (Gueta-Dahan et al. 1997), and in the present work was up-regulated in embryogenic callus (5.2-fold) demonstrate greater response against oxidative damage this callus. According to Kocsy et al. (2013) low levels of ROS or a short exposure to a certain concentration can promote differentiation, while excessive exposure can be damaging to the cell. The importance the ROS in cell differentiation homeostasis (Lee et al. 2016) reinforces the importance of these proteins being up-regulated in embryogenic callus.

Taken together, chaperone molecules are important in maintaining cellular homeostasis (Rodziewicz et al. 2014), due to their function in folding, assembly, translocation and degradation (Wang et al. 2004). In our study, two chaperones

were up-regulated (TCP-1 cpn60 chaperonin family, and Chaperone htpG family), and only one was down-regulated in embryogenic callus (Heat shock 70 family). The greater abundance of these proteins in the embryogenic callus suggest a greater control in the repair of damages the synthesized proteins or refolding of damaged proteins, that can occur due to the somatic embryogenic stress, avoiding cellular damages.

In addition, we have identified an DUF538 up-regulated in embryogenic callus 60.4-fold (Table 2). Although its function is not well known, it has been demonstrated which it regulates the abundance of other proteins that under stress it elevates redox enzyme activities including catalase, peroxidase, polyphenol oxidase and phenyalanine ammonia lyase (Gholizadeh 2011). We believe that the greatest abundance of this protein in embryogenic callus is critical for the acquisition of embryogenic competence, since it can modify the abundance of proteins important for cellular homeostasis.

3.2.5.2. Response to Lipid-related proteins

In the present study, several proteins related to lipid metabolism were identified. We found six Bifunctional inhibitor lipid-transfer seed storage 2S superfamily albumin up-regulated in embryogenic callus, one of them with 99.4-fold and only one down-regulated, 10.1-fold. These proteins are involved in the inhibition of proteases, lipid transport and lipid binding. They play a role in the development and germination and contribute to the resistance to biotic and abiotic stress(Palmeros-Suárez et al. 2016). It was described its role in the membrane stabilization and cell wall organization (Jin et al. 2015; Liu et al. 2015), and can promote plant cell wall loosening which may play a role in cell expansion and plant growth (Nieuwland et al. 2005; Yeats and Rose 2008). The importance of lipids metabolism has been shown in the maturation of papaya somatic embryos (Vale et al. 2014). In this study, it is clear that the lipid metabolism is essential from the acquisition of competence of papaya callus. We propose that the high levels of Bifunctional inhibitor lipid-transfer seed storage 2S superfamily albumin protein in the embryogenic callus may represent a key factor in embryogenic potential acquisition, enabling embryogenic development under maturation conditions.

Taken together, an acyl- oxidase, important protein in starch degradation, was up-regulated in embryogenic callus. These proteins catalyze the first step of peroxisomal fatty acid β-oxidation during early, postgerminative growth in oilseed species (Rylott et al. 2003). According to these authors, the elimination of the capacity of β-oxidation through the reduction of an acyl oxidase in *A. thaliana* compromised the development of embryos. Polyketide cyclase dehydrase and lipid transport superfamily, up-regulated in embryogenic callus, are involved in hormone abscisic acid (ABA) signaling by inhibiting hormone negative regulators (Santiago et al. 2009; Dorosh et al. 2013). It was also reported that embryonic callus contained higher levels of endogenous ABA than non-embryonic callus(Kiyosue et al. 1991; Nakagawa et al. 2001).

3.2.5.3. Auxin regulation-related proteins

The abundance of proteins related to auxin regulation reinforces the importance of auxin homeostasis in the acquisition of genetic competence in papaya callus. Auxins are thought to regulate or influence diverse responses on a whole-plant level, such as tropisms, apical dominance and root initiation, and responses on a cellular level, such as cell extension, division and differentiation (Hagen and Guilfoyle 2002; Jain et al. 2006; Bajguz and Piotrowska 2009).

IAA carboxyl methyltransferase 1 (IAMT1) is a protein which catalyzes the methylation of the carboxyl group, which inactivates IAA to its methyl ester (MeIAA) , which provides a distinct regulation of IAA activity, results in a modified non-polar auxin that is probably capable of independent movement of the carrier (Qin et al. 2005; Tanaka et al. 2014). In *A. thaliana* the overexpression of IAMT1 results in decreased the IAA response and agravitropic growth, whereas IAMT1 RNAi lines display decreased fertility, leaf epinasty, and decreased stature (Qin et al. 2005). Although, the role of MeIAA in auxin homeostasis is not yet well know (Korasick et al. 2013), we believe that high abundance of this protein in non-embryogenic callus 25.8-fold plays a negative role in auxin homeostasis.

UDP-glycosyltransferases (UGTs) superfamily, which were highly abundant in non-embryogenic callus, 80-fold, are a ubiquitous group of enzymes that catalyze the transfer of a sugar moiety from an activated donor molecule onto saccharide or non-saccharide acceptors, including molecules involved in secondary metabolism

(Caputi et al. 2012), and plant hormones (Poppenberger et al. 2005). Many of the plant hormones are known to be inactivated following glucosylation (Sembdner et al. 1994; Kleczkowski et al. 1995). The presence of high abundance of this protein may be related to its ability to affect homeostasis auxin (Jackson et al. 2001; Douglas Grubb et al. 2004; Jin et al. 2013).

Two Auxin GH3 were found to be more abundant in the embryogenic callus then non-embryogenic (133.4 and 12.4-fold). Auxin GH3 is an auxin-responsive protein, which regulates the expression of auxin-responsive genes (Tiwari et al. 2003) and maintains its homeostasis (Staswick et al. 2005). Despite the intense regulatory activity in both types of callus, we believe that the activity of these regulatory proteins has provided the necessary homeostasis for the progress of somatic embryogenesis in embryogenic callus, which has not occurred in non-embryogenic callus (Fig. 1).

3.2.5.4. Gene regulation-related proteins

The acquisition of embryogenic competence requires the reprogramming of gene expression patterns that will promote the change in cellular fate. These changes are dependent on an intense regulatory process (Fehér et al. 2003; Namasivayam 2007; Karami and Saidi 2010), including proteins associated with DNA replication and repair, mRNA splicing, translation and protein degradation.

Several protein-related with transcriptional and post-transcriptional regulation were up-regulated in the embryogenic callus. Among them, one RNA-binding KH domain-containing which plays a key role in the regulation of splicing and alternative splicing (Reddy et al. 2012); and one DEA(D H)-box RNA helicase family, proteins responsible for unwinding double-stranded RNA in an energy-dependent manner and disrupts RNA–protein interaction, being essential in all aspects of RNA metabolism at the level of expression and at the post-transcriptional level for pre-messenger RNA splicing, translation, and nucleocytoplasmic trafficking (Fung et al. 2006).

The variation in DNA methylation during somatic embryogenesis is variable among species, however, embryogenic cultures generally exhibit lower levels of global methylation than non-embryogenic ones (LoSchiavo et al. 1989; Miguel and Marum 2011; Maury et al. 2012; Teyssier et al. 2014), and therefore can be used as

markers of embryogenic competence. S-Adenosyl-L-methionine-dependent methyltransferase-like protein (SAM-MTases), which is involved in methylation of protein and nucleic acid was down-regulated in embryogenic callus, 16.7-fold, (Table 2). We believe that the greater abundance of this protein in non-embryogenic callus in relation to the embryogenic callus may lead to loss maturation competence.

3.2.5.5. Proteins not associated with the selected functional groups

Proteomic results show a great difference in the abundance of proteins related to differentiation and morphogenesis (Fig. 1 e Table 2). Among them, two Late embryogenesis abundant (LEA) proteins that were up-regulated in embryogenic callus, 8.7 and 71.2-fold. LEA proteins are more common at the end of embryogenesis, however, in papaya it was possible to identify them at an early stage. In papaya the beginning of somatic embryo formation until the globular stage occurs still in the induction stage, in this way, it is a critical stage for the progress of the normal development of the somatic embryo.

In embryogenic callus one abundant HIS triad family 3 (207-fold) were up-regulated (Table 2). These proteins are part of a superfamily of nucleotide-binding and -hydrolyzing enzymes (Brenner 2002). In *A. thaliana* it is associated with peroxisomes, revealing the unique involvement of this class of nucleotide hydroxylases/transferases in plant peroxisome functions (Reumann et al. 2009). These proteins play an important role in tumorigenesis by regulating the function of transcription complexes and possibly other unidentified enzymes (Huber and Weiske 2008).

The agglutinins are lectins with two or more carbohydrate binding sites that bind reversibly to specific free sugars or glycans present on glycoproteins and glycolipids (Peumans and Van Damme 1995). They have been associated with growth and endogenous functions related to the development and adaptation to biotic and abiotic stress (Zhang et al. 2015). Lectins play a critical role in the development of somatic embryos of alfalfa (Zeng et al. 2007) as well as during embryogenic pattern formation (Brill et al. 2001). Higher abundance of these proteins in embryogenic callus indicated that these proteins might be involved in specific endogenous protein–carbohydrate interactions in the regulation of adaptation of the cells to the differentiation.

3.2.6. CONCLUSION

During the somatic embryogenesis of *C. papaya* var. Golden is possible to identify two types of callus, embryogenic and non-embryogenic, that present great histo morphological and proteins abundance differences. In somatic embryogenesis stress plays a fundamental role in cell fate, however, the mechanisms responsible for these changes are misunderstood. In the present study, most of the proteins that regulate auxin were up-regulated in non-embryogenic callus, demonstrating an imbalance of this hormone. Embryogenic callus have demonstrated a greater ability to control ROS and preserve cellular homeostasis, and lipid metabolism also appears to play a key role in the acquisition of the embryogenic competence of papaya. Proteins not yet characterized in early stages of somatic embryogenesis have been identified and appear to play an important role in the acquisition of competence in papaya, such as Bifunctional inhibitor lipid-transfer seed storage 2S superfamily albumin, HIS triad family 3 and DUF538.

REFERENCES

aan den Toorn M, Albrecht C, de Vries S (2015) On the origin of SERKs: bioinformatics analysis of the somatic embryogenesis receptor kinases. Molecular plant 8 (5):762-782

Albrecht C, Russinova E, Kemmerling B, Kwaaitaal M, de Vries SC (2008) Arabidopsis SOMATIC EMBRYOGENESIS RECEPTOR KINASE proteins serve brassinosteroid-dependent and-independent signaling pathways. Plant Physiol 148 (1):611-619

Alemayehu A, Bočová B, Zelinová V, Mistrík I, Tamás L (2013) Enhanced lipoxygenase activity is involved in barley root tip swelling induced by cadmium, auxin or hydrogen peroxide. Environ Exper Bot 93:55-62

Anandan R, Sudhakar D, Balasubramanian P, Gutieirrez-Mora A (2012) In vitro somatic embryogenesis from suspension cultures of *Carica papaya* L. Sci Hortic 136:43-49. doi:10.1016/j.scienta.2012.01.003

Angel Huerta-Ocampo J, Alberto Osuna-Castro J, Jareth Lino-Lopez G, Barrera-Pacheco A, Mendoza-Hernandez G, De Leon-Rodriguez A, Barba de la Rosa AP (2012) Proteomic analysis of differentially accumulated proteins during ripening and in response to 1-MCP in papaya fruit. Journal of Proteomics 75 (7):2160-2169

Arnhold E (2013) Pacote em ambiente R para análise de variância e análises complementares. Braz J Vet Res Anim Sci 50 (6):488-492

Aryal R, Ming R (2014) Sex determination in flowering plants: papaya as a model system. Plant Sci 217:56-62

Badillo VM (1993) Caricaceae: segundo esquema. Rev Fac Agron Univ Cent Venezuela, Alcance 43

Bajguz A, Piotrowska A (2009) Conjugates of auxin and cytokinin. Phytochemistry 70 (8):957-969. doi:http://dx.doi.org/10.1016/j.phytochem.2009.05.006

Balbuena TS, Jo L, Pieruzzi FP, Dias LLC, Silveira V, Santa-Catarina C, Junqueira M, Thelen JJ, Shevchenko A, Floh EIS (2011) Differential proteome analysis of mature and germinated embryos of *Araucaria angustifolia*. Phytochemistry 72 (4-5):302-311. doi:10.1016/j.phytochem.2010.12.007

Banks RE, Dunn MJ, Hochstrasser DF, Sanchez J-C, Blackstock W, Pappin DJ, Selby PJ (2000) Proteomics: new perspectives, new biomedical opportunities. The Lancet 356 (9243):1749-1756

Baranwal VK, Mikkilineni V, Zehr UB, Tyagi AK, Kapoor S (2012) Heterosis: emerging ideas about hybrid vigour. J Exp Bot 63 (18):6309-6314

Barkla BJ, Vera-Estrella R, Pantoja O (2013) Progress and challenges for abiotic stress proteomics of crop plants. Proteomics 13 (12-13):1801-1815

Bartoli CG, Casalongué CA, Simontacchi M, Marquez-Garcia B, Foyer CH (2013) Interactions between hormone and redox signalling pathways in the control of growth and cross tolerance to stress. Environ Exp Bot 94:73-88

Begara-Morales JC, Chaki M, Sánchez-Calvo B, Mata-Pérez C, Leterrier M, Palma JM, Barroso JB, Corpas FJ (2013) Protein tyrosine nitration in pea roots during development and senescence. J Exp Bot 64 (4):1121-1134

Bilang J, Macdonald H, King PJ, Sturm A (1993) A soluble auxin-binding protein from Hyoscyamus muticus is a glutathione S-transferase. Plant Physiology 102 (1):29-34

Birchler JA, Auger DL, Riddle NC (2003) In search of the molecular basis of heterosis. The Plant Cell Online 15 (10):2236-2239

Boerjan W, Cervera M-T, Delarue M, Beeckman T, Dewitte W, Bellini C, Caboche M, Van Onckelen H, Van Montagu M, Inzé D (1995) Superroot, a recessive mutation in Arabidopsis, confers auxin overproduction. Plant Cell 7 (9):1405-1419

Boggess MV, Lippolis JD, Hurkman WJ, Fagerquist CK, Briggs SP, Gomes AV, Righetti PG, Bala K (2013) The need for agriculture phenotyping:“Moving from genotype to phenotype”. Journal of proteomics 93:20-39

Bourgeois M, Jacquin F, Cassecuelle F, Savois V, Belghazi M, Aubert G, Quillien L, Huart M, Marget P, Burstin J (2011) A PQL (protein quantity loci) analysis of mature pea seed proteins identifies loci determining seed protein composition. Proteomics 11 (9):1581-1594

Bowler C, Van Camp W, Van Montagu M, Inzé D, Asada K (1994) Superoxide dismutase in plants. Crit Rev Plant Sci 13 (3):199-218

Brasil M (1992) Regras para análise de sementes. LANARV/SNAD/MA, Brasília

Brenner C (2002) Hint, Fhit, and GalT: function, structure, evolution, and mechanism of three branches of the histidine triad superfamily of nucleotide hydrolases and transferases. Biochemistry 41 (29):9003-9014

Brill LM, Evans CJ, Hirsch AM (2001) Expression of MsLEC1-and MsLEC2-antisense genes in alfalfa plant lines causes severe embryogenic, developmental and reproductive abnormalities. The Plant Journal 25 (4):453-461

Bruce A (1910) The Mendelian theory of heredity and the augmentation of vigor. Science (New York, NY) 32 (827):627

Bruijne E, Langhe E, Rijck RV (1974) Actions of hormones and embryoid formation in callus cultures of *Carica papaya*. International Symposium on Crop Protection, Fytopharmacie en Fytiatrie Rijkslandbouwhoogeskool Medelingen 39:637-645

Burlat V, Kwon M, Davin LB, Lewis NG (2001) Dirigent proteins and dirigent sites in lignifying tissues. Phytochemistry 57 (6):883-897

Calderan-Rodrigues MJ, Jamet E, Bonassi MBCR, Guidetti-Gonzalez S, Begossi AC, Setem LV, Franceschini LM, Fonseca JG, Labate CA (2014) Cell wall proteomics of sugarcane cell suspension cultures. Proteomics 14 (6):738-749

Caputi L, Malnoy M, Goremykin V, Nikiforova S, Martens S (2012) A genome-wide phylogenetic reconstruction of family 1 UDP-glycosyltransferases revealed the expansion of the family during the adaptation of plants to life on land. The Plant Journal 69 (6):1030-1042

Cardoso DL (2012) Análise dialelica das capacidades geral e específica de combinação em mamoeiro. Universidade Estadual do Norte Fluminense

Cardoso DL, Luz LNd, Macêdo CMPd, Azeredo LGS, Pereira MG (2014) Heterosis in papaya: inter and intragroup analysis. Rev Bras Frutic 36 (3):610-619

Carpentier SC, Witters E, Laukens K, Deckers P, Swennen R, Panis B (2005) Preparation of protein extracts from recalcitrant plant tissues: an evaluation of different methods for two-dimensional gel electrophoresis analysis. Proteomics 5 (10):2497-2507. doi:DOI 10.1002/pmic.200401222

Castillo B, Smith M, Yadava U (1998) Plant regeneration from encapsulated somatic embryos of Carica papaya L. Plant Cell Rep 17 (3):172-176

Chakraborty AB, Berger SJ, Gebler JC (2007) Use of an integrated MS–multiplexed MS/MS data acquisition strategy for high-coverage peptide mapping studies. Rapid Commun Mass Spectrom 21 (5):730-744

Champagne A, Boutry M (2013) Proteomics of nonmodel plant species. Proteomics 13 (3-4):663-673. doi:10.1002/pmic.201200312

Chen M, Chen C, Wang D, Chen F (1991) Somatic embryogenesis and plant regeneration from immature embryos of Carica papaya× Carica cauliflora cultured in vitro. Canadian journal of botany 69 (9):1913-1918

Chen ZJ (2013) Genomic and epigenetic insights into the molecular bases of heterosis. Nature Reviews Genetics

Chugh A, Khurana P (2002) Gene expression during somatic embryogenesis - recent advances. Curr Sci 83 (6):715-730

Clarindo WR, de Carvalho CR, Araújo FS, de Abreu IS, Otoni WC (2008) Recovering polyploid papaya in vitro regenerants as screened by flow cytometry. Plant Cell Tiss Org Cult 92 (2):207-214

Colas I, Koroleva O, Shaw P (2010) Mass spectrometry in plant proteomic analysis. Plant Biosystems 144 (3):703-714

Cooke TJ, Racusen RH, Cohen JD (1993) The role of auxin in plant embryogenesis. The Plant Cell 5 (11):1494

Cooper M, Messina CD, Podlich D, Totir LR, Baumgarten A, Hausmann NJ, Wright D, Graham G (2014) Predicting the future of plant breeding: complementing empirical evaluation with genetic prediction. Crop and Pasture Science 65 (4):311-336

COSTA AdFd, Pacova B, MARTINS DdS, COSTA AdFd (2003) Caracterização de cultivares, estratégias e perspectivas do melhoramento genético do mamoeiro. A cultura do mamoeiro: tecnologias de produção:59-102

Cramer R, Bindschedler L, Agrawal G (2013) Plant proteomics in crop improvement. Proteomics 13 (12-13):1771. doi:10.1002/pmic.201370104

Dahal D, Newton KJ, Mooney BP (2016) Quantitative proteomics of *Zea mays* hybrids exhibiting different levels of heterosis. Journal of proteome research 15 (8):2445-2454

Damerval C, De Vienne D, Zivy M, Thiellement H (1986) Technical improvements in two-dimensional electrophoresis increase the level of genetic variation detected in wheat-seedling proteins. Electrophoresis 7 (1):52-54

Dantas JLL, Lima Jd (2001) Seleção e recomendação de variedades de mamoeiro-avaliação de linhagens e híbridos. Revista Brasileira de Fruticultura 23 (3):617-621

Davin LB, Lewis NG (2000) Dirigent proteins and dirigent sites explain the mystery of specificity of radical precursor coupling in lignan and lignin biosynthesis. Plant Physiol 123 (2):453-462

de Carvalho Silva R, Carmo LST, Luis ZG, Silva LP, Scherwinski-Pereira JE, Mehta A (2014) Proteomic identification of differentially expressed proteins during the acquisition of somatic embryogenesis in oil palm (*Elaeis guineensis* Jacq.). Journal of proteomics 104:112-127

de Jong AJ, Schmidt ED, de Vries SC (1993) Early events in higher-plant embryogenesis. Plant Mol Biol 22 (2):367-377

De Tullio MC, Jiang K, Feldman LJ (2010) Redox regulation of root apical meristem organization: connecting root development to its environment. Plant Physiol Biochem 48 (5):328-336

Dhekney S, Kandel R, Bergey D, Sitther V, Soorinathasundaram K, Litz R (2016) Advances in Papaya Biotechnology. Biocatalysis and Agricultural Biotechnology

Dias LLC, Balbuena TS, Silveira V, Santa-Catarina C, Shevchenko A, Floh EIS (2010) Two-dimensional gel electrophoretic protein profile analysis during seed development of *Ocotea catharinensis*: a recalcitrant seed species. Brazilian Journal of Plant Physiology 22:23-33

Dias NLP, Oliveira E, Dantas JLL (2011) Avaliação de genótipos de mamoeiro com uso de descritores agronômicos e estimação de parâmetros genéticos. Pesq agropec bras, Brasília 46 (11):1471-1479

Dodeman VL, Ducreux G, Kreis M (1997) Zygotic embryogenesis versus somatic embryogenesis. J Exp Bot 48 (313):1493-1509. doi:10.1093/jxb/48.8.1493

Domon B, Aebersold R (2006) Mass spectrometry and protein analysis. Science 312 (5771):212-217

Dorosh L, Kharenko OA, Rajagopalan N, Loewen MC, Stepanova M (2013) Molecular mechanisms in the activation of abscisic acid receptor PYR1. PLoS Comput Biol 9 (6):e1003114

dos Santos ALW, Elbl P, Navarro BV, de Oliveira LF, Salvato F, Balbuena TS, Floh EIS (2016) Quantitative proteomic analysis of *Araucaria angustifolia* (Bertol.) Kuntze cell lines with contrasting embryogenic potential. Journal of proteomics 130:180-189

Douglas Grubb C, Zipp BJ, Ludwig-Müller J, Masuno MN, Molinski TF, Abel S (2004) Arabidopsis glucosyltransferase UGT74B1 functions in glucosinolate biosynthesis and auxin homeostasis. The Plant Journal 40 (6):893-908

Duncan DR (2011) Organogenesis and embryogenesis in plant genetic transformation. Historical technology developments in plant transformation Bentham Sci Publ, Oak Park:46-54

Eldakak M, Milad SI, Nawar AI, Rohila JS (2013) Proteomics: A Biotechnology Tool for Crop Improvement. Frontiers in Plant Science 4. doi:10.3389/fpls.2013.00035

Elhiti M, Stasolla C (2015) ROS Signalling in Plant Embryogenesis. In: Reactive Oxygen and Nitrogen Species Signaling and Communication in Plants. Springer, pp 197-214

Falconer D, Mackay T (1996) Introduction to quantitative genetics, vol 4. Longman Sci and Tech, Harlow UK

Fang J, Lin A, Qiu W, Cai H, Umar M, Chen R, Ming R (2016) Transcriptome Profiling Revealed Stress-Induced and Disease Resistance Genes Up-Regulated in PRSV Resistant Transgenic Papaya. Frontiers in Plant Science 7

FAO (2016) FAOSTAT - Crop production. http://faostat.fao.org/. Accessed 10-06-2013 2013

Farias A, Oliveira A, Santos Filho H, Dantas J, Oliveira MdA, Sanches N, Medina V, Cordeiro Z (1994) A cultura do mamão. Embrapa-SPI; Cruz das Almas: Embrapa-CNPMF,

Fehér A (2015) Somatic embryogenesis—stress-induced remodeling of plant cell fate. Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms 1849 (4):385-402

Fehér A, Pasternak TP, Dudits D (2003) Transition of somatic plant cells to an embryogenic state. Plant Cell Tissue and Organ Culture 74 (3):201-228. doi:10.1023/a:1024033216561

Ferreguetti GA Caliman 01- o primeiro híbrido de mamão formosa brasileiro. In: MARTINS DS (ed) Papaya Brasil: qualidade do mamão para o mercado interno, Vitória, 2003. INCAPER, pp 211-218

Filizola M, Loew GH (2000) Role of protein environment in horseradish peroxidase compound I formation: molecular dynamics simulations of horseradish peroxidase-HOOH complex. J Am Chem Soc 122 (1):18-25

Fitch MM (1993) High frequency somatic embryogenesis and plant regeneration from papaya hypocotyl callus. Plant Cell Tiss Org Cult 32 (2):205-212

Foster LJ, de Hoog CL, Zhang Y, Zhang Y, Xie X, Mootha VK, Mann M (2006) A mammalian organelle map by protein correlation profiling. Cell 125 (1):187-199

Fröhlich A, Lindermayr C (2011) Deep insights into the plant proteome by pretreatment with combinatorial hexapeptide ligand libraries. Journal of proteomics 74 (8):1182-1189

Froman BE, Edwards PC, Bursch AG, Dehesh K (2000) ACX3, a novel medium-chain acyl-coenzyme A oxidase from Arabidopsis. Plant Physiol 123 (2):733-742

Fu Z, Jin X, Ding D, Li Y, Fu Z, Tang J (2011) Proteomic analysis of heterosis during maize seed germination. Proteomics 11 (8):1462-1472

Fung RW, Wang CY, Smith DL, Gross KC, Tao Y, Tian M (2006) Characterization of alternative oxidase (AOX) gene expression in response to methyl salicylate and methyl jasmonate pre-treatment and low temperature in tomatoes. Journal of plant physiology 163 (10):1049-1060

Gaedeke N, Klein M, Kolukisaoglu U, Forestier C, Müller A, Ansorge M, Becker D, Mamnun Y, Kuchler K, Schulz B (2001) The *Arabidopsis thaliana* ABC transporter AtMRP5 controls root development and stomata movement. EMBO J 20 (8):1875-1887

Galland R, Randoux B, Vasseur J, Hilbert J-L (2001) A glutathione S-transferase cDNA identified by mRNA differential display is upregulated during somatic embryogenesis in *Cichorium*. Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression 1522 (3):212-216

Gholizadeh A (2011) Heterologous expression of stress-responsive DUF538 domain containing protein and its morpho-biochemical consequences. The protein journal 30 (5):351-358

Giles K, Williams JP, Campuzano I (2011) Enhancements in travelling wave ion mobility resolution. Rapid Commun Mass Spectrom 25 (11):1559-1566

Ginn BR (2010) The implications of gene heterozygosity for protein folding and protein turnover. J Theor Biol 265 (4):554-564

Giraud E, Ng S, Carrie C, Duncan O, Low J, Lee CP, Van Aken O, Millar AH, Murcha M, Whelan J (2010) TCP transcription factors link the regulation of genes encoding mitochondrial proteins with the circadian clock in *Arabidopsis thaliana*. Plant Cell 22 (12):3921-3934

Goff SA (2011) A unifying theory for general multigenic heterosis: energy efficiency, protein metabolism, and implications for molecular breeding. New Phytol 189 (4):923-937

Goff SA, Zhang Q (2013) Heterosis in elite hybrid rice: speculation on the genetic and biochemical mechanisms. Curr Opin Plant Biol 16 (2):221-227

Gonsalves C, Cai W, Tennant P, Gonsalves D Effective development of papaya ringspot virus resistant papaya with untranslatable coat protein gene using a modified microprojectile transformation method. In: International Symposium on Biotechnology of Tropical and Subtropical Species Part 2 461, 1997. pp 311-314

Gonzalez-Fernandez R, Jorrin-Novo JV (2011) Contribution of proteomics to the study of plant pathogenic fungi. Journal of proteome research 11 (1):3-16

Grafi G, Barak S (2015) Stress induces cell dedifferentiation in plants. Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms 1849 (4):378-384

Gueta-Dahan Y, Yaniv Z, Zilinskas BA, Ben-Hayyim G (1997) Salt and oxidative stress: similar and specific responses and their relation to salt tolerance in citrus. Planta 203 (4):460-469

Guzmán-García E, Sánchez-Romero C, Panis B, Carpentier SC (2013) The use of 2D-DIGE to understand the regeneration of somatic embryos in avocado. Proteomics 13 (23-24):3498-3507

Hagen G, Guilfoyle T (2002) Auxin-responsive gene expression: genes, promoters and regulatory factors. Plant molecular biology 49 (3-4):373-385

Hallauer AR, Carena MJ, Miranda Filho J (2010) Heterosis. In: Quantitative Genetics in Maize Breeding. Springer, pp 477-529

Han P, Lu X, Mi F, Dong J, Xue C, Li J, Han B, Zhang X (2016) Proteomic analysis of heterosis in the leaves of sorghum-sudangrass hybrids. Acta biochimica et biophysica Sinica 48 (2):161-173

Heringer AS, Barroso T, Macedo AF, Santa-Catarina C, Souza GHMF, Floh EIS, de Souza-Filho GA, Silveira V (2015) Label-Free Quantitative Proteomics of Embryogenic and Non-Embryogenic Callus during Sugarcane Somatic Embryogenesis. PloS one 10 (6):e0127803

Heringer AS, Vale EM, Barroso T, Santa-Catarina C, Silveira V (2013) Polyethylene glycol effects in Somatic embryogenesis of papaya hybrid UENF/CALIMAN01 seeds. Theor Exp Plant Physiol 25 (2):116-124. doi:10.1590/S2197-00252013000200004

Hochholdinger F, Hoecker N (2007) Towards the molecular basis of heterosis. Trends Plant Sci 12 (9):427-432

Hoecker N, Keller B, Muthreich N, Chollet D, Descombes P, Piepho H-P, Hochholdinger F (2008) Comparison of maize (*Zea mays* L.) F1-hybrid and parental inbred line primary root transcriptomes suggests organ-specific patterns of nonadditive gene expression and conserved expression trends. Genetics 179 (3):1275-1283

Hoecker N, Keller B, Piepho H-P, Hochholdinger F (2006) Manifestation of heterosis during early maize (*Zea mays* L.) root development. Theor Appl Genet 112 (3):421-429

Homhuan S, Kijwijan B, Wangsomnuk P, Bodhipadma K, Leung DW (2008) Variation of plants derived from indirect somatic embryogenesis in cotyledon explants of papaya. Science Asia 34:347-352

Hosmani PS, Kamiya T, Danku J, Naseer S, Geldner N, Guerinot ML, Salt DE (2013) Dirigent domain-containing protein is part of the machinery required for formation of the lignin-based Casparian strip in the root. Proc Natl Acad Sci U S A 110 (35):14498-14503

Hu H, Xiong L, Yang Y (2005) Rice SERK1 gene positively regulates somatic embryogenesis of cultured cell and host defense response against fungal infection. Planta 222 (1):107-117

Huang H, Cao H, Niu Y, Dai S (2012) Expression analysis of Nudix hydrolase genes in Chrysanthemum lavandulifolium. Plant Mol Bio Rep 30 (4):973-982

Huber O, Weiske J (2008) β-Catenin takes a HIT. Cell cycle 7 (10):1326-1331

Ide CD, Pereira MG, Viana AP, Pereira TNS (2009) Use of testers for combining ability and selection of papapya hybrids. Crop Breeding and Applied Biotechnology 9 (1):60-66

Jackson RG, Lim E-K, Li Y, Kowalczyk M, Sandberg G, Hoggett J, Ashford DA, Bowles DJ (2001) Identification and biochemical characterization of anArabidopsis indole-3-acetic acid glucosyltransferase. J Biol Chem 276 (6):4350-4356

Jain M, Kaur N, Tyagi AK, Khurana JP (2006) The auxin-responsive GH3 gene family in rice (*Oryza sativa*). Functional & integrative genomics 6 (1):36

Jimenez VM (2005) Involvement of plant hormones and plant growth regulators on in vitro somatic embryogenesis. Plant Growth Regulation 47 (2-3):91-110. doi:10.1007/s10725-005-3478-x

Jin J, Shi J, Liu B, Liu Y, Huang Y, Yu Y, Dong A (2015) MORF-RELATED GENE702, a reader protein of trimethylated histone H3 lysine 4 and histone H3 lysine 36, is involved in brassinosteroid-regulated growth and flowering time control in rice. Plant physiology 168 (4):1275-1285

Jin S-H, Ma X-M, Han P, Wang B, Sun Y-G, Zhang G-Z, Li Y-J, Hou B-K (2013) UGT74D1 is a novel auxin glycosyltransferase from *Arabidopsis thaliana*. PloS one 8 (4):e61705

Jorrin-Novo JV (2009) Plant Proteomics. Journal of Proteomics 72 (3):283-284. doi:10.1016/j.jprot.2009.03.002

Jurica MS, Mesecar A, Heath PJ, Shi W, Nowak T, Stoddard BL (1998) The allosteric regulation of pyruvate kinase by fructose-1, 6-bisphosphate. Structure 6 (2):195-210

Kaeppler S (2011) Heterosis: one boat at a time, or a rising tide? New Phytol 189 (4):900-902

Kaplan DR, Cooke TJ (1997) Fundamental concepts in the embryogenesis of dicotyledons: a morphological interpretation of embryo mutants. The Plant Cell 9 (11):1903

Karami O, Aghavaisi B, Pour AM (2009) Molecular aspects of somatic-to-embryogenic transition in plants. Journal of chemical biology 2 (4):177-190

Karami O, Saidi A (2010) The molecular basis for stress-induced acquisition of somatic embryogenesis. Mol Biol Rep 37 (5):2493-2507. doi:10.1007/s11033-009-9764-3

Kitamiya E, Suzuki S, Sano T, Nagata T (2000) Isolation of two genes that were induced upon the initiation of somatic embryogenesis on carrot hypocotyls by high concentrations of 2, 4-D. Plant Cell Reports 19 (6):551-557

Kiyosue T, Satoh S, Kamada H, Harada H (1991) Purification and immunohistochemical detection of an embryogenic cell protein in carrot. Plant Physiol 95 (4):1077-1083

Kleczkowski K, Schell J, Bandur R (1995) Phytohormone conjugates: nature and function. Crit Rev Plant Sci 14 (4):283-298

Kocsy G, Tari I, Vanková R, Zechmann B, Gulyás Z, Poór P, Galiba G (2013) Redox control of plant growth and development. Plant Sci 211:77-91

Korasick DA, Enders TA, Strader LC (2013) Auxin biosynthesis and storage forms. J Exp Bot 64 (9):2541-2555

Kosová K, Vítámvás P, Prášil IT, Renaut J (2011) Plant proteome changes under abiotic stress - Contribution of proteomics studies to understanding plant stress response. Journal of Proteomics 74 (8):1301-1322. doi:http://dx.doi.org/10.1016/j.jprot.2011.02.006

Krysan PJ, Jester PJ, Gottwald JR, Sussman MR (2002) An Arabidopsis mitogen-activated protein kinase kinase kinase gene family encodes essential positive regulators of cytokinesis. Plant Cell 14 (5):1109-1120

Kulinska-Lukaszek K, Tobojka M, Adamiok A, Kurczynska E (2012) Expression of the BBM gene during somatic embryogenesis of Arabidopsis thaliana. Biol Plant 56 (2):389-394

Kumaravel M, Uma S, Backiyarani S, Saraswathi MS, Vaganan MM, Muthusamy M, Sajith KP (2017) Differential proteome analysis during early somatic embryogenesis in Musa spp. AAA cv. Grand Naine. Plant Cell Rep 36 (1):163-178

Kung Y-J, Yu T-A, Huang C-H, Wang H-C, Wang S-L, Yeh S-D (2010) Generation of hermaphrodite transgenic papaya lines with virus resistance via transformation of somatic embryos derived from adventitious roots of in vitro shoots. Transgenic Res 19 (4):621-635

Kwon M, Burlat V, Davin LB, Lewis NG (1999) Localization of dirigent protein involved in lignan biosynthesis: implications for lignification at the tissue and subcellular level. In: Gross G, Hemingway R, Yoshida T, Branham S (eds) Plant Polyphenols 2. Chemistry, Biology, Pharmacology, Ecology. Springer, New York, pp 393-411

Larsen PB, Cancel J, Rounds M, Ochoa V (2007) Arabidopsis ALS1 encodes a root tip and stele localized half type ABC transporter required for root growth in an aluminum toxic environment. Planta 225 (6):1447-1458

Lassoudiére A (1968) Le papayer: description e genetique. Fruits 23 (11):585-596

Lee K, Park O-S, Seo PJ (2016) RNA-Seq Analysis of the Arabidopsis Transcriptome in Pluripotent Calli. Molecules and cells 39 (6):484

Leljak-Levanić D, Bauer N, Mihaljević S, Jelaska S (2004) Changes in DNA methylation during somatic embryogenesis in *Cucurbita pepo* L. Plant Cell Rep 23 (3):120-127

Leterrier M, Chaki M, Airaki M, Valderrama R, Palma JM, Barroso JB, Corpas FJ (2011) Function of S-nitrosoglutathione reductase (GSNOR) in plant development and under biotic/abiotic stress. Plant Signal Behav 6 (6):789-793

Lin C-M, Yang J-S (2001) Papaya Somatic Embryo Induction from Fruit-Bearing Field Plants: Effects of Root Supporting Material and Position of the Root Explants. Acta Horticulturae:489-492

Lines RE, Persley D, Dale JL, Drew R, Bateson MF (2002) Genetically engineered immunity to Papaya ringspot virus in Australian papaya cultivars. Mol Breed 10 (3):119-129

Lippman ZB, Zamir D (2007) Heterosis: revisiting the magic. Trends Genet 23 (2):60-66

Litz R, Conover R (1980) Somatic embryogenesis in cell cultures of Carica stipulata [Large-scale production model]. HortScience (USA)

Litz R, Conover R (1982) In vitro somatic embryogenesis and plant regeneration from Carica papaya L. ovular callus. Plant Science Letters 26 (2-3):153-158

Liu F, Zhang X, Lu C, Zeng X, Li Y, Fu D, Wu G (2015) Non-specific lipid transfer proteins in plants: presenting new advances and an integrated functional analysis. Journal of experimental botany:erv313

LoSchiavo F, Pitto L, Giuliano G, Torti G, Nuti-Ronchi V, Marazziti D, Vergara R, Orselli S, Terzi M (1989) DNA methylation of embryogenic carrot cell cultures and its variations as caused by mutation, differentiation, hormones and hypomethylating drugs. Theor Appl Genet 77 (3):325-331

Lu T, Takagi H Cryopreservation of somatic embryos of papaya (*Carica papaya* L.) by vitrification. In: 2000 World Congress on In Vitro Biology, Program Issue, 2000. p 73

Luschnig C (2002) Auxin transport: ABC proteins join the club. Trends Plant Sci 7 (8):329-332

Luz LN, Pereira MG, Barros FR, Barros GB, Ferreguetti GA (2015) New hybrids of papaya evaluated in the traditional region for crop and in the brazilian semiarid. Revista Brasileira de Fruticultura 37 (1):159-171

Macedo CMP, Pereira MG, Cardoso DL, da Silva RF (2013) Avaliação de híbridos elites de mamão, de seus recíprocos e respectivos genitores, quanto a qualidade fisiológica de sementes. J Seed Sci 35 (2):190-197

Marcon C, Lamkemeyer T, Malik WA, Ungrue D, Piepho H-P, Hochholdinger F (2013) Heterosis-associated proteome analyses of maize (*Zea mays* L.) seminal roots by quantitative label-free LC–MS. Journal of proteomics 93:295-302

Marcon C, Schützenmeister A, Schütz W, Madlung J, Piepho H-P, Hochholdinger F (2010) Nonadditive protein accumulation patterns in maize (*Zea mays* L.) hybrids during embryo development. Journal of proteome research 9 (12):6511-6522

Marin SLD, Pereira MG, Amaral Júnior ATd, Martelleto LAP, Ide CD (2006) Partial diallel to evaluate the combining ability for economically important traits of papaya. Sci Agric 63 (6):540-546

Marrs KA (1996) The functions and regulation of glutathione S-transferases in plants. Annual review of plant biology 47 (1):127-158

Marsoni M, Bracale M, Espen L, Prinsi B, Negri AS, Vannini C (2008) Proteomic analysis of somatic embryogenesis in *Vitis vinifera*. Plant Cell Rep 27 (2):347-356. doi:10.1007/s00299-007-0438-0

Martins GN (2007) Qualidade de sementes de mamão: Determinações metodológicas e de componentes genéticos. Universidade Estadual do Norte Fluminense, Campos dos Goytacazes

Maury S, Trap-Gentil MV, Hébrard C, Weyens G, Delaunay A, Barnes S, Lefebvre M, Joseph C (2012) Genic DNA methylation changes during in vitro organogenesis: organ specificity and conservation between parental lines of epialleles. Physiol Plant 146 (3):321-335

Mayer KF, Schoof H, Haecker A, Lenhard M, Jürgens G, Laux T (1998) Role of WUSCHEL in regulating stem cell fate in the Arabidopsis shoot meristem. Cell 95 (6):805-815

Melchinger A (1999) Genetic diversity and heterosis. vol 10. chapter,

Meyer S, Pospisil H, Scholten S (2007) Heterosis associated gene expression in maize embryos 6 days after fertilization exhibits additive, dominant and overdominant pattern. Plant Mol Biol 63 (3):381-391

Miguel C, Marum L (2011) An epigenetic view of plant cells cultured in vitro: somaclonal variation and beyond. J Exp Bot 62 (11):3713-3725

Mikkelsen MD, Naur P, Halkier BA (2004) Arabidopsis mutants in the C–S lyase of glucosinolate biosynthesis establish a critical role for indole-3-acetaldoxime in auxin homeostasis. Plant J 37 (5):770-777

Ming R, Hou S, Feng Y, Yu Q, Dionne-Laporte A, Saw JH, Senin P, Wang W, Ly BV, Lewis KL (2008) The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus). Nature 452 (7190):991-996

Mohayeji M, Capriotti AL, Cavaliere C, Piovesana S, Samperi R, Stampachiacchiere S, Toorchi M, Lagana A (2014) Heterosis profile of sunflower leaves: A label free proteomics approach. Journal of Proteomics 99:101-110

Moore G, Litz R (1984) Biochemical markers for Carica papaya, C. cauliflora, and plants from somatic embryos of their hybrid. J Am Soc Hort Sci 109 (2):213-218

Moose SP, Mumm RH (2008) Molecular plant breeding as the foundation for 21st century crop improvement. Plant Physiol 147 (3):969-977

Murashige T, Skoog F (1962) A revised medium for rapid growth and bio assays with tobacco tissue cultures. Physiol Plant 15 (3):473-497. doi:10.1111/j.1399-3054.1962.tb08052.x

Nakagawa H, Saijyo T, Yamauchi N, Shigyo M, Kako S, Ito A (2001) Effects of sugars and abscisic acid on somatic embryogenesis from melon (Cucumis melo L.) expanded cotyledon. Scientia Horticulturae 90 (1):85-92

Namasivayam P (2007) Acquisition of embryogenic competence during somatic embryogenesis. Plant Cell Tiss Org Cult 90 (1):1-8

Nanjo Y, Skultety L, Uváčková Lu, Klubicová Kn, Hajduch M, Komatsu S (2011) Mass spectrometry-based analysis of proteomic changes in the root tips of flooded soybean seedlings. Journal of proteome research 11 (1):372-385

Natarajan S, Xu CP, Caperna TJ, Garrett WA (2005) Comparison of protein solubilization methods suitable for proteomic analysis of soybean seed proteins. Anal Biochem 342 (2):214-220. doi:10.1016/j.ab.2005.04.046

Nieuwland J, Feron R, Huisman BA, Fasolino A, Hilbers CW, Derksen J, Mariani C (2005) Lipid transfer proteins enhance cell wall extension in tobacco. The Plant Cell 17 (7):2009-2019

Nishihama R, Ishikawa M, Araki S, Soyano T, Asada T, Machida Y (2001) The NPK1 mitogen-activated protein kinase kinase kinase is a regulator of cell-plate formation in plant cytokinesis. Genes Dev 15 (3):352-363

Noah AM, Niemenak N, Sunderhaus S, Haase C, Omokolo DN, Winkelmann T, Braun H-P (2013) Comparative proteomic analysis of early somatic and zygotic embryogenesis in *Theobroma cacao* L. Journal of Proteomics 78:123-133. doi:10.1016/j.jprot.2012.11.007

Noctor G, Mhamdi A, Chaouch S, Han Y, Neukermans J, MARQUEZ-GARCIA B, Queval G, Foyer CH (2012) Glutathione in plants: an integrated overview. Plant, Cell & Environment 35 (2):454-484

Noh B, Murphy AS, Spalding EP (2001) Multidrug resistance–like genes of Arabidopsis required for auxin transport and auxin-mediated development. Plant Cell 13 (11):2441-2454

Ogawa T, Ishikawa K, Harada K, Fukusaki E, Yoshimura K, Shigeoka S (2009) Overexpression of an ADP-ribose pyrophosphatase, AtNUDX2, confers enhanced tolerance to oxidative stress in Arabidopsis plants. Plant J 57 (2):289-301

Oliveira AMG, Farias A, Santos Filho H, Oliveira JRP, Dantas JLL, Santos Ld, Oliveira MdA, Souza Júnior M, Silva MJ, Almeida Od (1994) Mamão para exportação: aspectos técnicos da produção, vol 9. Ministério da Agricultura, do Abastecimento e da Reforma Agrária, Secretaria de Desenvolvimento Rural, SDR, Programa de Apoio à Produção e Exportação de Frutas, Hortaliças, Flores e Plantas Ornamentais--FRUPEX,

Pagnussat GC, Lanteri ML, Lombardo MC, Lamattina L (2004) Nitric oxide mediates the indole acetic acid induction activation of a mitogen-activated protein kinase cascade involved in adventitious root development. Plant Physiol 135 (1):279-286

Palmeros-Suárez PA, Massange-Sánchez JA, Sánchez-Segura L, Martínez-Gallardo NA, Rangel EE, Gómez-Leyva JF, Délano-Frier JP (2016) AhDGR2, an amaranth abiotic stress-induced DUF642 protein gene, modifies cell wall structure and composition and causes salt and ABA hyper-sensibility in transgenic Arabidopsis. Planta:1-18

Panchaud A, Affolter M, Moreillon P, Kussmann M (2008) Experimental and computational approaches to quantitative proteomics: status quo and outlook. J Proteomics 71 (1):19-33

Pasternak T, Potters G, Caubergs R, Jansen MA (2005) Complementary interactions between oxidative stress and auxins control plant growth responses at plant, organ, and cellular level. J Exp Bot 56 (418):1991-2001

Pasternak TP, Prinsen E, Ayaydin F, Miskolczi P, Potters G, Asard H, Van Onckelen HA, Dudits D, Feher A (2002) The role of auxin, pH, and stress in the activation of embryogenic cell division in leaf protoplast-derived cells of alfalfa. Plant Physiol 129 (4):1807-1819. doi:10.1104/pp.000810

Pedreschi R, Lurie S, Hertog M, Nicolaï B, Mes J, Woltering E (2013) Post-harvest proteomics and food security. Proteomics 13 (12-13):1772-1783

Pennington SR, Dunn MJ (2001) Proteomics: from protein sequence to function, vol 1. BIOS Scientific Publishers Ltd, Oxford

Pereira AM, Masiero S, Nobre MS, Costa ML, Solís M-T, Testillano PS, Sprunck S, Coimbra S (2014) Differential expression patterns of arabinogalactan proteins in *Arabidopsis thaliana* reproductive tissues. J Exp Bot:eru300

Peumans WJ, Van Damme E (1995) Lectins as plant defense proteins. Plant Physiol 109 (2):347

Peyret P, Perez P, Alric M (1995) Structure, genomic organization, and expression of the *Arabidopsis thaliana* aconitase gene plant aconitase show significant homology with mammalian iron-responsive element-binding protein. J Biol Chem 270 (14):8131-8137

Poppenberger B, Fujioka S, Soeno K, George GL, Vaistij FE, Hiranuma S, Seto H, Takatsuto S, Adam G, Yoshida S (2005) The UGT73C5 of Arabidopsis thaliana glucosylates brassinosteroids. Proceedings of the National Academy of Sciences of the United States of America 102 (42):15253-15258

Qin G, Gu H, Zhao Y, Ma Z, Shi G, Yang Y, Pichersky E, Chen H, Liu M, Chen Z (2005) An indole-3-acetic acid carboxyl methyltransferase regulates Arabidopsis leaf development. The Plant Cell 17 (10):2693-2704

Quiroz-Figueroa FR, Rojas-Herrera R, Galaz-Avalos RM, Loyola-Vargas VM (2006) Embryo production through somatic embryogenesis can be used to study cell

differentiation in plants. Plant Cell Tissue and Organ Culture 86 (3):285-301. doi:10.1007/s11240-006-9139-6

R Core Team (2014) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna

Raemakers C, Jacobsen E, Visser R (1995) Secondary somatic embryogenesis and applications in plant breeding. Euphytica 81 (1):93-107

Reddy AS, Day IS, Göhring J, Barta A (2012) Localization and dynamics of nuclear speckles in plants. Plant physiology 158 (1):67-77

Reinert J (1958) Morphogenese und ihre Kontrolle an Gewebekulturen aus Carotten. Naturwissenschaften 45 (14):344-345

Reis RS, Vale EM, Heringer AS, Santa-Catarina C, Silveira V (2015) Putrescine induces somatic embryo development and proteomic changes in embryogenic callus of sugarcane. J Proteomics 130:170-179. doi:http://dx.doi.org/10.1016/j.jprot.2015.09.029

Reumann S, Quan S, Aung K, Yang P, Manandhar-Shrestha K, Holbrook D, Linka N, Switzenberg R, Wilkerson CG, Weber AP (2009) In-depth proteome analysis of Arabidopsis leaf peroxisomes combined with in vivo subcellular targeting verification indicates novel metabolic and regulatory functions of peroxisomes. Plant Physiol 150 (1):125-143

Roberts DM, Harmon AC (1992) Calcium-modulated proteins: targets of intracellular calcium signals in higher plants. Annu Rev Plant Biol 43 (1):375-414

Rodrigues SP, Da Cunha M, Ventura JA, Bueno Fernandes PM (2009) Effects of the *Papaya meleira* virus on papaya latex structure and composition. Plant Cell Rep 28 (5):861-871. doi:10.1007/s00299-009-0673-7

Rodrigues SP, Ventura JA, Aguilar C, Nakayasu ES, Almeida IC, Fernandes PMB, Zingali RB (2011) Proteomic analysis of papaya (*Carica papaya* L.) displaying typical sticky disease symptoms. Proteomics 11 (13):2592-2602. doi:10.1002/pmic.201000757

Rodrigues SP, Ventura JA, Aguilar C, Nakayasu ES, Choi H, Sobreira TJ, Nohara LL, Wermelinger LS, Almeida IC, Zingali RB, Fernandes PM (2012) Label-free quantitative proteomics reveals differentially regulated proteins in the latex of sticky diseased *Carica papaya* L. plants. Journal of proteomics 75 (11):3191-3198. doi:10.1016/j.jprot.2012.03.021

Rodziewicz P, Swarcewicz B, Chmielewska K, Wojakowska A, Stobiecki M (2014) Influence of abiotic stresses on plant proteome and metabolome changes. Acta Physiologiae Plantarum 36 (1):1-19

Rosas MM, Quiroz-Figueroa F, Shannon LM, Ruiz-May E (2016) The Current Status of Proteomic Studies in Somatic Embryogenesis. In: Somatic Embryogenesis: Fundamental Aspects and Applications. Springer, pp 103-119

Rose JK, Bashir S, Giovannoni JJ, Jahn MM, Saravanan RS (2004) Tackling the plant proteome: practical approaches, hurdles and experimental tools. The Plant Journal 39 (5):715-733

Ruggiero C, Marin SLD, Durigan JF (2011) Mamão, uma história de sucesso. Revista Brasileira de Fruticultura 33 (SPE1):76-82

Rylott EL, Rogers CA, Gilday AD, Edgell T, Larson TR, Graham IA (2003) Arabidopsis mutants in short-and medium-chain acyl-CoA oxidase activities accumulate acyl-CoAs and reveal that fatty acid β-oxidation is essential for embryo development. J Biol Chem 278 (24):21370-21377

Santelia D, Vincenzetti V, Azzarello E, Bovet L, Fukao Y, Düchtig P, Mancuso S, Martinoia E, Geisler M (2005) MDR-like ABC transporter AtPGP4 is involved in auxin-mediated lateral root and root hair development. FEBS Lett 579 (24):5399-5406

Santiago J, Dupeux F, Round A, Antoni R, Park S-Y, Jamin M, Cutler SR, Rodriguez PL, Márquez JA (2009) The abscisic acid receptor PYR1 in complex with abscisic acid. Nature 462 (7273):665-668

Santos-Mendoza M, Dubreucq B, Baud S, Parcy F, Caboche M, Lepiniec L (2008) Deciphering gene regulatory networks that control seed development and maturation in Arabidopsis. The Plant Journal 54 (4):608-620

Schluter H, Apweiler R, Holzhutter H, Jungblut PR (2009) Finding one's way in proteomics: a protein species nomenclature. Chem Cent J 3 (11):11

Schmidt EDL, Guzzo F, Toonen MAJ, deVries SC (1997) A leucine-rich repeat containing receptor-like kinase marks somatic plant cells competent to form embryos. Development 124 (10):2049-2062

Schmildt ER, Teixeira SL, Schmildt O (2005) Estabelecimento e multiplicação in vitro do mamoeiro 'Sunrise Solo Line 72/12'e 'Tainung 01'. Papaya Brasil: mercado e inovações tecnológicas para o mamão Vitória, ES: INCAPER:221-224

Schnable PS, Springer NM (2013) Progress toward understanding heterosis in crop plants. Annu Rev Plant Biol 64:71-88

Sembdner G, Atzorn R, Schneider G (1994) Plant hormone conjugation. In: Signals and Signal Transduction Pathways in Plants. Springer, pp 223-245

Shi H, Ishitani M, Kim C, Zhu J-K (2000) The Arabidopsis thaliana salt tolerance gene SOS1 encodes a putative Na+/H+ antiporter. Proc Natl Acad Sci U S A 97 (12):6896-6901

Shi Y, Xiang R, Horváth C, Wilkins JA (2004) The role of liquid chromatography in proteomics. J Chromatogr 1053 (1):27-36

Shull GH (1908) The composition of a field of maize. J Hered (1):296-301

Silva JC, Denny R, Dorschel CA, Gorenstein M, Kass IJ, Li G-Z, McKenna T, Nold MJ, Richardson K, Young P (2005) Quantitative proteomic analysis by accurate mass retention time pairs. Anal Chem 77 (7):2187-2200

Silveira V, Vita A, Macedo A, Dias M, Floh E, Santa-Catarina C (2013) Morphological and polyamine content changes in embryogenic and non-embryogenic callus of sugarcane. Plant Cell, Tissue and Organ Culture (PCTOC) 114 (3):351-364. doi:10.1007/s11240-013-0330-2

Smertenko A, Bozhkov P (2014) The life and death signalling underlying cell fate determination during somatic embryogenesis. In: Applied plant cell biology. Springer, pp 131-178

Soares EdA, Werth EG, Madroñero LJ, Ventura JA, Rodrigues SP, Hicks LM, Fernandes PM (2017) Label-free quantitative proteomic analysis of pre-flowering PMeV-infected *Carica papaya* L. Journal of Proteomics 151:275-283

Song X, Ni Z, Yao Y, Xie C, Li Z, Wu H, Zhang Y, Sun Q (2007) Wheat (*Triticum aestivum* L.) root proteome and differentially expressed root proteins between hybrid and parents. Proteomics 7 (19):3538-3557

Springer NM, Stupar RM (2007) Allelic variation and heterosis in maize: how do two halves make more than a whole? Genome Res 17 (3):264-275

Starai V, Escalante-Semerena J (2004) Acetyl-coenzyme A synthetase (AMP forming). Cell Mol Life Sci 61 (16):2020-2030

Staswick PE, Serban B, Rowe M, Tiryaki I, Maldonado MT, Maldonado MC, Suza W (2005) Characterization of an Arabidopsis enzyme family that conjugates amino acids to indole-3-acetic acid. The Plant Cell 17 (2):616-627

Steiner N, Santa-Catarina C, Guerra MP, Cutri L, Dornelas MC, Floh EIS (2012) A gymnosperm homolog of SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASE-1 (SERK1) is expressed during somatic embryogenesis. Plant Cell Tissue and Organ Culture 109 (1):41-50. doi:10.1007/s11240-011-0071-z

Steward FC, Marion OM, Smith J (1958) Growth and organized development of cultured cells. I. growth and division of freely suspended cells. Am J Bot 45 (9):693-703. doi:10.2307/2439507

Su YH, Zhao XY, Liu YB, Zhang CL, O'Neill SD, Zhang XS (2009) Auxin-induced WUS expression is essential for embryonic stem cell renewal during somatic embryogenesis in Arabidopsis. The Plant Journal 59 (3):448-460

Sun D-Q, Lu X-H, Liang G-L, Guo Q-G, Mo Y-W, Xie J-H (2011) Production of triploid plants of papaya by endosperm culture. Plant Cell, Tissue and Organ Culture (PCTOC) 104 (1):23-29

Sun L, Wu Y, Zou H, Su S, Li S, Shan X, Xi J, Yuan Y (2013) Comparative proteomic analysis of the H99 inbred maize (*Zea mays* L.) line in embryogenic and non-embryogenic callus during somatic embryogenesis. Plant Cell Tissue and Organ Culture 113 (1):103-119. doi:10.1007/s11240-012-0255-1

Suzuki M, Kobayashi K, Nagata N, Matsuda O, Iba K, Yoshida S, Muranaka T (2008) Novel root culture system using a recessive mutant with a rooty phenotype. Plant Biotechnol 25 (2):197-200

Szyrajew K, Bielewicz D, Dolata J, Wójcik AM, Nowak K, Szczygieł-Sommer A, Szweykowska-Kulinska Z, Jarmolowski A, Gaj MD (2017) microRNAs are intensively regulated during induction of somatic embryogenesis in Arabidopsis. Frontiers in Plant Science 8:18

Takac T, Pechan T, Samaj J (2011) Differential proteomics of plant development. Journal of Proteomics 74 (5):577-588. doi:10.1016/j.jprot.2011.02.002

Tanaka K, Hayashi K-i, Natsume M, Kamiya Y, Sakakibara H, Kawaide H, Kasahara H (2014) UGT74D1 catalyzes the glucosylation of 2-oxindole-3-acetic acid in the auxin metabolic pathway in Arabidopsis. Plant and Cell Physiology 55 (1):218-228

Tautorus TE, Fowke LC, Dunstan DI (1991) Somatic embryogenesis in conifers. Can J Bot 69 (9):1873-1899. doi:10.1139/b91-237

Teyssier C, Maury S, Beaufour M, Grondin C, Delaunay A, Le Metté C, Ader K, Cadene M, Label P, Lelu-Walter MA (2014) In search of markers for somatic embryo maturation in hybrid larch (Larix× eurolepis): global DNA methylation and proteomic analyses. Physiol Plant 150 (2):271-291

Tiwari SB, Hagen G, Guilfoyle T (2003) The roles of auxin response factor domains in auxin-responsive transcription. The Plant Cell 15 (2):533-543

Tovar-Méndez A, Miernyk JA, Randall DD (2003) Regulation of pyruvate dehydrogenase complex activity in plant cells. Eur J Biochem 270 (6):1043-1049

Vale EM, Heringer AS, Barroso T, da Silva Ferreira AT, da Costa MN, Perales JE, Santa-Catarina C, Silveira V (2014) Comparative proteomic analysis of somatic embryo maturation in *Carica papaya* L. Proteome Sci 12 (1):37. doi:10.1186/1477-5956-12-37

VanBuren R, Zeng F, Chen C, Zhang J, Wai CM, Han J, Aryal R, Gschwend AR, Wang J, Na J-K (2015) Origin and domestication of papaya Yh chromosome. Genome Res 25 (4):524-533

Vanderschuren H, Lentz E, Zainuddin I, Gruissem W (2013) Proteomics of model and crop plant species: status, current limitations and strategic advances for crop improvement. Journal of proteomics 93:5-19

Varhaníková M, Uvackova L, Skultety L, Pretova A, Obert B, Hajduch M (2014) Comparative quantitative proteomic analysis of embryogenic and non-embryogenic calli in maize suggests the role of oxylipins in plant totipotency. Journal of proteomics 104:57-65

Verdeil J-L, Hocher V, Huet C, Grosdemange F, Escoute J, Ferrière N, Nicole M (2001) Ultrastructural changes in coconut calli associated with the acquisition of embryogenic competence. Ann Bot 88 (1):9-18

Vivas M, Silveira S, Cardoso DL, Pereira MG, Santos P, Ferreguetti GA (2012) Capacidade combinatória e heterose para resistência a pinta-preta em mamoeiro por meio de análise dialélica. Trop Plant Pathol 37 (5):326-332

Vivas M, Silveira SFd, Cardoso DL, Amaral Júnior ATd, Pereira MG (2014) Heterosis for resistance to phoma leaf spot in papaya hybrids obtained from crosses between and within heterotic groups. Summa Phytopathol 40 (4):318-322

Wang W, Meng B, Ge X, Song S, Yang Y, Yu X, Wang L, Hu S, Liu S, Yu J (2008) Proteomic profiling of rice embryos from a hybrid rice cultivar and its parental lines. Proteomics 8 (22):4808-4821

Wang WX, Vinocur B, Shoseyov O, Altman A (2004) Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. Trends Plant Sci 9 (5):244-252. doi:10.1016/j.tplants.2004.03.006

Waris H (1957) A striking morphogenetic effect of amino acid in seed plant. Suom Kemistil B 30:121

Wasinger VC, Cordwell SJ, Cerpapoljak A, Yan JX, Gooley AA, Wilkins MR, Duncan MW, Harris R, Williams KL, Humpherysmith I (1995) Progress with gene-product mapping of the mollicutes - *Mycoplasma genitalium*. Electrophoresis 16 (7):1090-1094. doi:10.1002/elps.11501601185

Wilkins MR, Sanchez JC, Williams KL, Hochstrasser DF (1996) Current challenges and future applications for protein maps and post-translational vector maps in proteome projects. Electrophoresis 17 (5):830-838

Wu XM, Kou SJ, Liu YL, Fang YN, Xu Q, Guo WW (2015) Genomewide analysis of small RNAs in nonembryogenic and embryogenic tissues of citrus: microRNA- and siRNA-mediated transcript cleavage involved in somatic embryogenesis. Plant Biotechnol J 13 (3):383-394

Xiao S, Gao W, Chen QF, Ramalingam S, Chye ML (2008) Overexpression of membrane-associated acyl-CoA-binding protein ACBP1 enhances lead tolerance in Arabidopsis. Plant J 54 (1):141-151

Yang X, Zhang X (2010) Regulation of somatic embryogenesis in higher plants. Critical Reviews in Plant Science 29 (1):36-57

Yang X, Zhang X, Yuan D, Jin F, Zhang Y, Xu J (2012) Transcript profiling reveals complex auxin signalling pathway and transcription regulation involved in dedifferentiation and redifferentiation during somatic embryogenesis in cotton. BMC Plant Biol 12 (1):110

Yao Y, Ni Z, Zhang Y, Chen Y, Ding Y, Han Z, Liu Z, Sun Q (2005) Identification of differentially expressed genes in leaf and root between wheat hybrid and its parental inbreds using PCR-based cDNA subtraction. Plant Mol Bio 58 (3):367-384

Yeats TH, Rose JK (2008) The biochemistry and biology of extracellular plant lipid-transfer proteins (LTPs). Protein Science 17 (2):191-198

Yie S-T, Liaw S (1977) Plant regeneration from shoot tips and callus of papaya. In vitro 13 (9):564-568. doi:10.1007/BF02627852

Zeng F, Zhang X, Cheng L, Hu L, Zhu L, Cao J, Guo X (2007) A draft gene regulatory network for cellular totipotency reprogramming during plant somatic embryogenesis. Genomics 90 (5):620-628. doi:10.1016/j.ygeno.2007.07.007

Zhang J, Ma H, Chen S, Ji M, Perl A, Kovacs L, Chen S (2009) Stress response proteins' differential expression in embryogenic and non-embryogenic callus of *Vitis vinifera* L. cv. Cabernet Sauvignon - a proteomic approach. Plant Sci 177 (2):103-113. doi:http://dx.doi.org/10.1016/j.plantsci.2009.04.003

Zhang W, Wang X, Yu Q, Ming R, Jiang J (2008) DNA methylation and heterochromatinization in the male-specific region of the primitive Y chromosome of papaya. Genome Res 18 (12):1938-1943

Zhang Y, Wang L, Xie J, Zheng H (2015) Differential protein expression profiling of Arabidopsis thaliana callus under microgravity on board the Chinese SZ-8 spacecraft. Planta 241 (2):475-488

Zhao J, Wang B, Wang X, Zhang Y, Dong M, Zhang J (2015a) iTRAQ-based comparative proteomic analysis of embryogenic and non-embryogenic tissues of Prince Rupprecht's larch (*Larix principis-rupprechtii* Mayr). Plant Cell, Tissue and Organ Culture (PCTOC) 120 (2):655-669

Zhao Y, Li Z, Liu G, Jiang Y, Maurer HP, Würschum T, Mock H-P, Matros A, Ebmeyer E, Schachschneider R (2015b) Genome-based establishment of a high-yielding heterotic pattern for hybrid wheat breeding. Proc Natl Acad Sci 112 (51):15624-15629

Zhou T, Yang X, Guo K, Deng J, Xu J, Gao W, Lindsey K, Zhang X (2016) ROS homeostasis regulates somatic embryogenesis via the regulation of auxin signaling in cotton. Molecular & Cellular Proteomics 15 (6):2108-2124

Zi J, Zhang J, Wang Q, Zhou B, Zhong J, Zhang C, Qiu X, Wen B, Zhang S, Fu X (2013) Stress responsive proteins are actively regulated during rice (Oryza sativa) embryogenesis as indicated by quantitative proteomics analysis. PLoS One 8 (9):e74229