Collège Doctoral du Languedoc-Roussillon

THÈSE

Pour obtenir le grade de

Docteur

Délivré par **Montpellier SupAgro**

Préparée au sein de l'école doctorale SIBAGHE
Et de l'unité de recherche AGAP

Spécialité: **Biologie Intégrative des Plantes**

Présentée par **Michelle GUITTON COTTA**

MOLECULAR MECHANISMS IN THE FIRST STEP OF ABA-MEDIATED RESPONSE IN *COFFEA* ssp.

Soutenue le 27 Janvier 2017 devant le jury composé de

Dra. Annie Marion-Poll, Directeur de Recherche, INRA	Rapporteur
Dr. Michel G. A. Vincentz, Professeur, UNICAMP	Rapporteur
Dr. Pierre Marraccini, Chargé de Recherche, CIRAD	Co-directeur
Dr. Alan C. Andrade, Chargé de Recherche, EMBRAPA	Co-directeur
Dr. Pascal Gantet, Professeur, UM2	Examinateur
Dr. Mário Lúcio V. Resende, Professeur, UFLA	Examinateur

Thèse préparée en cotutelle avec
l'Université Fédérale de Lavras, Brésil

Acknowledgements

Foremost, I would like to express my special appreciation and sincere gratitude to my advisors and Professors Dr. Alan Andrade and Dr. Pierre Marraccini for the support of my PhD study and research. First of all, I would like to thanks Dr. Alan for the PhD project idea and for the opportunity given to me to go to Montpellier to work in this project. In special, I am very grateful to Dr. Pierre not only for also providing the means for this experience abroad, making all effort to translate documents and collect information, but as well for his personal contribution to my scientific career since undergraduation. I'm thankful to you for all the academic lessons, thesis correction, sleepless nights working together, suggestions, patience, immense knowledge and precious advises. Your guidance helped me in all the time of research and mainly during the writing of this thesis. Finally, many thanks for encouraging me to be always better and for contributing to my growth as a scientist.

I would like to give special thanks to Dominique for your friendly welcome in France, the kindly insertion in CIRAD and for all intellectual support, expertise and patience with my basic English and poor French. I wish to thanks Luciano Paiva for all the lessons as my professor, coordinator and for all help and suggestions during the stressful ending. I'm also grateful to Jeff Leung and Hervé Etienne for the insights, advises and recommendations during my PhD committees. I would also like to thank my PhD jury to attend as committee members even at hardship.

I'm really grateful to all members of the SouthGreen platform and also to my French colleagues from UMR AGAP ID team, especially Steph, Jeff, Felix and Fredo for all the attention, patience, insights and knowledge shared in Bioinformatics which was a complete new area for me. I thank as well my fellow labmates in Brazil, for the stimulating discussions, for the hard work in experiments, for the support during writing, for all the fun and encouragement given to achieve my goals.

Last but not the least, I would like to give a special thanks to my family, my parents and brothers that give me unconditionally love and support. I would like to give also a particular thanks to my husband Gabriel Alves for encouraging my scientific and personal goals and to be with me in the hard moments in the hospital, working together during weekends, holydays and endless nights.

I should like most sincerely thank all!

“Everybody is a genius. But if you judge a fish by its ability to climb a tree it will live its whole life believing that is stupid”

Albert Einstein

Abstract

Abscisic acid (ABA) is a phytohormone universally conserved in land plants which coordinates several aspects of the plant response to water deficit such as root architecture, seed dormancy and regulation of stomatal closure. A mechanism of ABA signal transduction has been proposed, involving intracellular ABA receptors (PYR/PYL/RCARs) interacting with PP2Cs phosphatases and SnRK2 protein kinases regulating this tripartite protein system. The goal of this study was to identify and characterize for the first time the orthologs genes of this tripartite system in *Coffea*. For this purpose, protein sequences from *Arabidopsis*, citrus, rice, grape, tomato and potato were chosen as query to search orthologous genes in the Coffee Genome Hub (http://coffee-genome.org/). Differential expression in tissues as leaves, seeds, roots and floral organs was checked through *in silico* analyses. *In vivo* gene expression analyses were also performed by RT-qPCR in leaves and roots of drought-tolerant (D^T 14, 73 and 120) and drought-susceptible (D^S 22) *C. canephora* Conilon clones submitted (or not) to drought. The expression profiles of the tripartite system *CcPYL-PP2C-SnRK2* genes were also analyzed in leaves of *C. arabica* (*Ca*) and *C. canephora* (*Cc*) plants grown under hydroponic condition and submitted to exogenous ABA treatment (500 µM). This approach allowed the identification and characterization of 24 candidate genes (9 *PYL/RCARs*, 6 *PP2Cs* and 9 *SnRK2s*) in *Cc* genome. The protein motifs identified in the predict coffee sequences enabled characterize these genes as family's members of PYL/RCARs receptors, PP2Cs phosphatases or SnRK2 kinases of the ABA-dependent response pathway. These families were functionally annotated in the *Cc* genome. *In vivo* analyses revealed that eight genes were up-regulated under drought conditions in both leaves and roots tissues. On the other hand, *CcPYL4* was down-regulated under water deficit in both tissues for all clones. Among them, three genes coding phosphatases were expressed in all (D^T and D^S) clones therefore suggesting that they were activated as a general response to cope with drought stress. However, two other phosphatase coding genes were up-regulated only in the D^T clones, suggesting that they constitute key-genes for drought tolerance in these clones. The D^T clones also showed differential gene expression profiles for five other genes thus reinforcing the idea that multiple biological mechanisms are involved in drought tolerance in *Cc*. In response to exogenous ABA, 17 genes were expressed in leaves of *Cc* and *Ca* plants. Several genes were differentially expressed in the D^T clone 14 either in control condition or after 24h with ABA treatment. Under control condition, five genes were higher expressed in *Cc* as in *Ca* D^T plants. The kinase *CcSnRK2.6* was highlighted as a gene specifically expressed in the *Cc* plants (D^T and D^S) after 72h of ABA treatment. Overall, it was observed that ABA signaling pathway is delayed in the D^S *C. arabica* Rubi. Those molecular evidences corroborated with microscopies analyses which showed that the D^T clone 14 was more efficient to control the stomatal closure than other coffee plants in response to ABA treatment. All these evidences will help us to identify the genetic determinism of drought tolerance through ABA pathway essential to obtain molecular markers that could be used in coffee breeding programs.

Résumé de la thèse en français

Introduction

Le genre *Coffea*, membre de la famille des rubiacées qui comprend plus de 124 espèces, constitue une matière première agricole parmi les plus échangées au niveau du commerce mondial. Ce genre comprend des espèces vivaces toutes originaires du continent africain, que l'on rencontre aussi bien sur les hauts plateaux d'Ethiopie, dans les savanes du grand ouest, les forêts tropicales et équatoriales du bassin du Congo, jusqu'à Madagascar et ses îles avoisinantes (Mascareignes et Comores). Parmi toutes les espèces de ce genre, seules *Coffea arabica* et *C. canephora* ont une importance économique. La plus cultivée est *C. arabica* qui est aussi la connue et la plus appréciée car elle fournit une boisson de qualité, riche en aromes et flaveurs avec des teneurs limités en caféine. *C. arabica* est une plante allotétraploïde (2n = 4 x = 44) issue d'une hybridation naturelle survenue il y a environ 1 million d'années entre les deux espèces diploïdes *C. canephora* et *C. eugenioides* qui constituent ces deux génomes ancestraux. En raison de son mode de reproduction par autopollinisation, cette espèce est caractérisée par une faible diversité génétique. A l'inverse, *C. canephora* est une espèce diploïde (2n = 2 x = 22) allogame qui présente une forte variabilité génétique et la capacité à s'adapter à différentes conditions climatiques. Le café issu de ses grains est par contre considéré comme de qualité inférieure, car riche en caféine et en acides chlorogéniques, et essentiellement commercialisé sous la forme café lyophilisé utilisé dans les boissons instantanées. Ces deux espèces sont cultivées dans plus de 80 pays et recouvrent une surface totale d'environ 11 millions d'hectares. Cette filière caféicole emploie plusieurs millions de personnes, ce qui souligne son importance économique et sociale dans les pays de la zone intertropicale ou les caféiers sont cultivés. Parmi ceux-ci, le Brésil est le premier pays producteur avec environ un tiers de la production mondiale (soit 45 millions de sacs de 60kgs par an).

Comme de nombreuses grandes productions végétales, le caféier est une plante sensible aux changements climatiques, particulièrement aux épisodes de sécheresse et fortes températures. Ces facteurs affectent ainsi le développement des plantes et leur floraison mais également leur production en quantité (rendement) et en qualité (composition biochimique). Le dernier rapport du groupe d'experts intergouvernemental sur l'évolution du climat (GIEC), mentionne une augmentation des périodes de sécheresse et des températures (de 2 à 3°C au cours des 40 prochaines années). Des études montrent déjà que ces changements modifieront la répartition mondiale des principales zones de production de café, engendrant ainsi des problèmes tant environnementaux, qu'économiques et sociaux. Dans ce contexte, la création de nouvelles variétés de caféiers plus tolérantes à la sécheresse est devenue l'une des priorités des institutions de recherche travaillant sur l'amélioration génétique de cette plante.

Au cours des deux dernières décennies, plusieurs clones tolérants à la sécheresse de *C. canephora* Conilon (population cultivée au Brésil), caractérisés par leur vigueur et par leur capacité de production en condition de limitation en eau, ont été identifiés. Ces clones ont fait l'objet de plusieurs études de physiologie et de biologie moléculaire notamment pour analyser leurs réponses face stress hydrique. Les

analyses réalisées au sein de notre groupe ont ainsi permis à identifier une quarantaine de gènes candidats (GCs) potentiellement impliqués dans le déterminisme génétique de la tolérance à la sécheresse au sein cette espèce, et pour lesquels l'expression dans les feuilles augmente en condition de sécheresse notamment chez clones tolérants. Parmi ces gènes, plusieurs sont connus pour coder des protéines de régulation (facteurs de transcription de type DREB, NAC) essentielles dans les réponses des plantes aux stress biotiques et abiotiques par exemple en réponse à l'acide abscissique (ABA). Plusieurs études de diversité génétique et d'analyse de la régulation (promoteurs) de ces GCs (comme par exemple *CcDREB1D*) sont en actuellement cours au laboratoire. Elles sont facilitées par la mise à disposition récente du séquençage complet du génome de *C. canephora*.

Le travail présenté dans cette thèse consiste à profiter de cette information pour étudier les gènes codant pour les protéines impliquées dans les premières étapes de perception et de transduction du signal ABA chez le caféier.

Le système «tripartite» de perception et de transduction du signal ABA

L'ABA est une phytohormone très conservée au sein du règne végétal, impliquée dans les réponses des plantes aux stress abiotiques (notamment à la sécheresse) mais également dans l'architecture racinaire, la dormance des graines et la régulation de la fermeture des stomates. Récemment, un mécanisme de perception et de transduction de signal ABA a été proposé. Celui-ci fait intervenir des récepteurs intracellulaires de ABA (dénommés PYR/PYL/RCARs), des phosphatases (dénommées PP2Cs) et des protéines kinases (dénommées SnRK2), l'ensemble constituant un système tripartite de protéines. Dans ce système, les protéines SnRK2 sont donc les régulateurs «positifs» alors que les phosphatases PP2Cs sont des régulateurs «négatifs».

Le modèle actuel de transduction du signal ABA peut ainsi être décrit de la manière suivante:

- en absence d'ABA, les récepteurs intracellulaires PYR/PYL/RCAR sont libres et inactifs, alors que les protéines kinases SnRK2 sont inactivées par fixation des phosphatases PP2Cs qui, en les déphosphorylant, inhibe leur activité.
- en présence d'ABA, l'ABA se lie aux récepteurs PYR/PYL/RCAR ce qui engendre leur changement de conformation et augmente leur affinité aux phosphatases PP2Cs qui ne sont plus liées aux kinases SnRK2. Sous leur forme libre et phosphorylée, les kinases sont alors actives.

Objectifs de ce travail

Cette thèse consiste donc à identifier et caractériser les gènes orthologues de ce système tripartite chez *C. canephora*. Ce travail, qui est le premier à utiliser les données du séquençage du génome complet de cette plante pour analyser plusieurs familles de gènes, vise notamment à répondre aux questions scientifiques suivantes:

- combien de gènes composent le système tripartite *PYR/PYL/RCAR-SnRK2-PP2C* chez *C. canephora*?
- comment ces gènes sont-ils organisés au sein du génome de cette espèce?
- ces gènes sont-ils exprimés de la même manière dans:
  - les différents tissus et organes de *C. canephora* et de *C. arabica*?
  - les racines et les feuilles des clones tolérants et sensibles à la sécheresse de *C. canephora*?
- ces gènes sont-ils directement régulés par l'ABA?

Ainsi, les principaux objectifs de ce travail étaient :

1. d'identifier les gènes de *C. canephora* codant pour chacune des protéines du système «tripartite» de perception de l'ABA,
2. de caractériser ces gènes en comparant leur structure et leurs familles par rapport à ceux déjà connus dans plusieurs plantes modèles,
3. d'identifier les gènes fonctionnels du système tripartite PYR/PYL/RCAR-SnRK2-PP2C de *C. canephora* en étudiant leur expression dans les feuilles et les racines,
4. d'étudier l'expression des gènes fonctionnels dans ces mêmes tissus chez des clones tolérants et sensibles à la sécheresse de *C. canephora* cultivés en serre en condition de stress hydrique,
5. de comparer les profils d'expression obtenus *in vivo* à ceux obtenus *in silico* pour cette même espèce,
6. d'étudier les effets d'un apport exogène d'ABA sur l'expression de ces gènes dans les feuilles de plantules de *C. canephora et* de *C. arabica* cultivées en hydroponie.

Principaux résultats

Identification des gènes *PYR/PYL/RCAR-SnRK2-PP2C* du système tripartite chez *C. canephora*

Les données génomiques de plantes modèles ont été utilisées pour initier les études de génomique comparative et de génomique fonctionnelle des gènes de *C. canephora* codant pour les protéines du système tripartite PYR/PYL/RCAR-SnRK2-PP2C. Ainsi, les séquences de ces protéines préalablement identifiées chez *Arabidopsis thaliana*, mais également chez l'oranger (*Citrus sinensis*), le riz asiatique (*Oryza sativa*), la vigne (*Vitis vinifera*), la tomate (*Solanum lycopersicum*) et la pomme de terre (*Solanum tuberosum*) ont servi de séquences de référence pour rechercher les séquences orthologues de *C. canephora* (http://coffee-genome.org/). Ces analyses, menées avec le programme BLASTP, ont permis d'identifier 24 gènes candidats répartis comme suit :

- 9 codant pour les protéines PYR/PYL/RCAR,
- 6 codant pour les phosphatases de type PP2C et,
- 9 codant pour les kinases de type SnRK2s.

Cette classification a été réalisée en utilisant les motifs protéiques spécifiques de chacune de ces familles identifiés dans les protéines putatives déduites des gènes de *C. canephora*.

L'expression de ces gènes a ensuite été analysée par RT-qPCR dans les feuilles et les racines des clones tolérants (D^T: 14, 73 et 120) et sensible (D^S: 22) de *C. canephora* Conilon cultivés en condition d'irrigation (contrôle non stressé) ou condition de sécheresse après (suspension de l'irrigation) après les plantes aient atteint un potentiel de nuit en eau des feuilles (Ψ_{pd}: *pre-dawn leaf water potential*) de -3,0 MPa.

Analyses phylogénétiques et profils d'expression de ces gènes

- Famille des gènes *PYR/PYL/RCAR*

Concernant cette famille, neuf protéines ont été identifiées dans le génome de *C. canephora*. Ce nombre est similaire aux protéines PYL identifiées chez *C. sinensis* et *V. vinifera*, mais est inférieur aux protéines PYL présentes chez Arabidopsis, la tomate et le riz. Les analyses phylogénétiques et de structure des gènes *PYL* de *C. canephora* ont monté la présence de duplications, comme c'est le cas pour les gènes *CcPYL7* (gènes dupliqués *CcPYL7a* et *CcPYL7b*) et *CcPYL8* (gènes dupliqués *CcPYL8a* et *CcPYL8b*). Les gènes *CcPYL7a* et *CcPYL7b* sont tous les deux localisés sur le chromosome 0 qui correspond à un pseudo-chromosome formé par le montage arbitraire de séquences génomiques non encore apparentées aux onze autres chromosomes de *C. canephora*. Contrairement aux gènes dupliqués *CcPYL8a* et *CcPYL8b*, les analyses d'expression par qPCR ont montré que les gènes *CcPYL7a* et *CcPYL7b* ne s'exprimaient pas dans les feuilles et les racines, aussi bien clones tolérants ou sensibles à la sécheresse, et ceci quel que soient leurs conditions de culture. Ces résultats sont en accord avec les analyses d'expression *in silico* déduites à partir la base de données du génome de café. Cependant, comme *CcPYL7a* et *CcPYL7b* s'expriment durant le développement des grains de *C. arabica* (données non présentées), on peut en conclure que ces gènes dupliqués sont fonctionnels et proposer qu'ils jouent de fonctions différentes puisqu'ils sont régulés différemment dans les tissus de *C. canephora*.

Nos travaux ont montré que les gènes paralogues *CcPYL8a* et *CcPYL8b* présentaient des profils d'expression différents dans les racines des clones *C. canephora*, l'expression du gène *CcPYL8b* augmentant en condition de sécheresse notamment chez les clones tolérants 73 et 120 alors que celle du gène *CcPYL8a* était peut affectée. Ces différences d'expression entre ces deux gènes pourraient s'expliquer par la présence d'un intron de 316pb dans la région 5' UTR du gène *CcPYL8a* ou de séquences retrotransposons de type *copia* dont des régions LTRs (*long terminal repeats*) sont trouvées par exemple dans le promoteur de ce gène. Cette observation renforcerait le rôle déjà rapporté dans la littérature, que pourraient jouer les éléments transposables dans la régulation de l'expression des gènes de caféiers soumis à des périodes de sécheresse.

Les analyses d'expression ont également montré une forte activation du gène *CcPYL9* par la sécheresse dans les feuilles et les racines des clones D^T 14 et 73 de *C. canephora*. A l'inverse, les

conditions de sécheresse répriment fortement l'expression du gène *CcPYL4* dans ces deux organes et quelques soient les clones.

- Famille des gènes *PP2C*

C. canephora possède six protéines de type phosphatase-PP2Cs très similaires à celles présentent chez les *Solanaceae* comme CcABI1, CcABI2, CcHAB et CcHAI similaires aux protéines de la pomme de terre alors que CcABI2, CcAHG3 et CcHAB sont plus proches des protéines de tomate.

Nos résultats montrent très clairement une augmentation de l'expression des gènes *CcABI2*, *CcAHG3* et *CcHAI* en condition de sécheresse dans les feuilles et les racines des clones D^T et D^S clones de *C. canephora*. La sécheresse induit également l'expression des gènes *CcAHG2* et *CcHAB* mais seulement dans les feuilles des clones tolérants D^T 14, 73 et 120. Concernant le gène *CcAHG2*, il est intéressant de noter que celui-ci ne s'exprime que dans les feuilles des clones tolérants mais pas dans les racines. Au sein des clones tolérants, on note également le comportement singulier du clone 120 pour lequel l'expression racinaire des gènes *CcABI1*, *CcABI2* et *CcAHG3* augmente spécifiquement en condition de sécheresse. Enfin, d'un point de vue quantitatif, *CcHAI* est le gène le plus fortement surexprimé en condition de sécheresse dans les feuilles et dans les racines.

- Famille des gènes SnRK2

Neuf protéines kinase de type SnRK2 ont été identifiées *C. canephora*. Parmi celles-ci, les protéines déduites des gènes *CcSnRK2.12* and *CcSnRK2.13* sont considérées comme incomplètes et n'ont pas été étudiées plus en détail. Par comparaison avec les autres gènes *SnRK2* végétaux, les autres gènes de *C. canephora* se divisent en trois sous-groupes qui se différencient en fonction de leur réponse vis-à-vis de l'ABA. Ainsi, *CcSnRK2.1* et *CcSnRK2.10* constituent le sous-groupe I des gènes *SnRK2* non activés par l'ABA. Les gènes *CcSnRK2.7* et *CcSnRK2.8* appartiennent quant à eux au sous-groupe II des gènes faiblement activés par l'ABA. Enfin, le sous-groupe III est composé de *CcSnRK2.2* et *CcSnRK2.6* qui sont fortement activés par l'ABA. De manière intéressante, le gène de caféier *CcSnRK2.11* ne présente aucune homologie avec les gènes des trois autres sous-groupes.

Dans les feuilles, le gène *CcSnRK2.2* est le seul qui présente une augmentation significative de son expression en condition de sécheresse pour les clones tolérants (D^T) 14 et 73. Dans les racines, l'expression de ce gène augmente également en condition de sécheresse cette fois-ci chez les trois clones tolérants mais pas chez le clone sensible (D^S) 22. On note à nouveau le comportement singulier du clone 120 qui présente une augmentation significative en condition de stress de l'expression des gènes *CcSnRK2.2*, *CcSnRK2.6* et *CcSnRK2.7*. A l'inverse, la sécheresse diminue l'expression du gène *CcSnRK2.10* dans les racines des clones tolérants 14 et 120 mais pas dans celles des clones 73 et 22. On note par ailleurs que le gène *CcSnRK2.11* s'exprime dans les feuilles mais pas dans les racines. Enfin, aucune expression des gènes *CcSnRK2.1*, *CcSnRK2.12* et *CcSnRK2.13* n'a été observée avec les

amorces utilisées lors des expériences de PCR quantitative dans les feuilles et les racines, ceci quelques soient les clones et leurs conditions de culture.

Effets de l'ABA exogène sur l'expression des gènes du système tripartite PYR/PYL/RCAR-SnRK2-PP2C

Afin d'analyser les effets de l'ABA sur l'expression des gènes du système tripartite caractérisés chez *C. canephora*, des jeunes plantes de *C. arabica* (plantules âgées de 3 mois issues de graines des cultivars IAPAR59 and Rubi, respectivement considérés comme tolérant et sensible à la sécheresse [Moffato *et al.*, 2016]) et de *C. canephora* (boutures âgées de 6 mois des clones précédemment décrits) ont été cultivées en hydroponie (Hoagland) et soumises à un traitement exogène ABA (500 µM). Les études d'expression ont été réalisées dans les feuilles de ces plantes prélevées après 24 et 72h de traitement ABA. Parmi les 24 gènes testés (9 *PYLs*, 6 *PP2Cs* et 9 *SnRK2s*), 17 d'entre eux s'expriment dans les feuilles de *C. canephora* et de *C. arabica*. Plusieurs présentent des profils d'expression différents entre les deux espèces de caféier, les génotypes de ces espèces et les temps d'exposition à l'ABA.

Après 24h d'exposition, on observe par exemple une très forte augmentation de l'expression des gènes *CcPYR1*, *CcPYL8b*, *CcSnRK2.7* et *CcSnRK2.11* et une nette diminution de l'expression des gènes *CcAHG2* chez le clone tolérant 14 de *C. canephora*, suggérant ainsi une activation (mode « on ») rapide du système tripartite chez le clone 14 en réponse à l'application d'ABA. A l'inverse, aucune variation d'expression n'est observé pour ces mêmes gènes en réponse à ABA (24h et 72h) chez le clone 22, ce qui semble traduire son incapacité à néo-synthétiser de nouveaux récepteurs ABA et les protéines kinases SnRK2 (régulateurs « positifs » du système tripartite). Ces résultats semblent en accord avec ceux des analyses de microscopie qui montrent un contrôle plus efficace la fermeture des stomates chez le clone 14 que chez le clone 22 de *C. canephora*.

Globalement, les profils d'expression des gènes *PYR1, PYL8b, SnRK2.7* et *SnRK2.7* semblent également montrer que les clones de *C. canephora* répondent à l'ABA plus rapidement que les cultivars de *C. arabica*. Par ailleurs, même si le gène *PYL8a* ne semble pas jouer un rôle prépondérant dans la réponse des clones de *C. canephora* à la sécheresse (cf. Chapitre I), on note chez *C. arabica* une expression plus précoce de ce gène dans les feuilles du IAPAR59 que dans celle du Rubi. Par contre, aucun de ces cultivars n'exprime le gène *SnRK2.6* pour lequel l'expression est détectée chez *C. canephora*. Ce résultat met en évidence la nécessité de tester chez l'espèce *C. arabica*, l'expression de ces gènes par qPCR avec des amorces spécifiques de chacun de ses sous-génomes.

Discussion

Chez les plantes supérieures, l'ABA augmente en réponse à la sécheresse aussi bien dans les racines que dans les feuilles dans lesquelles les quantités d'ABA traduisent un équilibre entre la

biosynthèse et la dégradation de cette hormone, mais dépendent aussi de sa localisation (séquestration) cellulaire et de son transport. Les résultats présentés dans ce travail montrent que seul le clone tolérant 120 présente une augmentation significative de la quantité d'ABA dans ses feuilles en réponse à la sécheresse. Par ailleurs, et quel que soit les clones de *C. canephora*, aucune différence significative des teneurs en ABA n'est observée dans les racines en fonction des conditions de stress. Ces résultats suggèrent que les phénotypes de tolérance et de sensibilité à la sécheresse des clones étudiés de *C. canephora* ne sont probablement pas dus à des altérations de la voie de biosynthèse et de dégradation de l'ABA, mais pourraient plutôt provenir d'altérations des mécanismes de perception et de transduction de ce signal hormonal.

Des études antérieures publiées au sein de notre laboratoire (Marraccini *et al.*, 2012; Vieira *et al.*, 2013) ont montré que le clone tolérant 73 présentait un rapport significativement plus élevé du taux transport des électrons/taux assimilation nette de CO_2 (ETR/A : *electron transport rate/CO2 assimilation rate*) en condition de sécheresse ratio que les autres clones tolérants (D^T 14 et 120). Ceci suggère l'existence au sein du clone 73 de mécanismes spécifiques lui permettant de protéger son appareil photosynthétique contre la photoinhibition par exemple en réduisant la formation des dérivés réactifs de l'oxygène (ROS, en anglais pour *reactive oxygen species*). Ce clone présente d'ailleurs en condition de sécheresse une augmentation de l'expression de plusieurs gènes (par exemple *CcAPX1*, *CcPDH1* et *CcNSH1*) codant pour des protéines impliquées dans les systèmes antioxydants et d'osmoprotection. La surexpression observée dans les feuilles de ce clone en condition de sécheresse des gènes *SnRK2.2*, *SnRK2.7* et *SnRK2.8* pourrait favoriser l'activation de ces voies de protection et de détoxification.

Un autre résultat intéressant concerne le gène *CcAHG2* (*ABA-hypersensitive germination*) de la famille des phosphatases PP2C, pour lequel l'expression augmente en condition de sécheresse spécifiquement dans les feuilles des clones tolérants. Même si ce gène n'a pas d'orthologue chez *A. thaliana* et n'a pas de fonctions connues, il pourrait être intéressant de poursuivre son étude (cf. Conclusion générale).

D'autres travaux ont aussi montré que le clone 120 tolérant à la sécheresse possédait un système racinaire plus profond qui pourrait lui permettre un meilleur accès à l'eau du sol et par conséquent expliquer (en partie) son phénotype (Pinheiro *et al.*, 2005). Les résultats d'expression présentés dans ce travail montrent que le clone 120 se distingue très clairement des autres en surexprimant dans ses racines en condition de stress notamment le gène *CcPYL8b*, mais aussi les gènes des kinases *CcSnRK2.2*, *CcSnRK2.6* (activés par l'ABA) et *CcSnRK2.7* ainsi que les gènes des phosphatases PP2C *CcABI1*, *CcABI2* et *CcAHG3*. Ces résultats suggèrent que le système racinaire joue un rôle clé dans la réponse à la sécheresse dans ce clone qui fait probablement directement intervenir l'ABA au moins pour le clone 120.

Conclusion générale et perspectives

Les résultats présentés dans ce travail sont les premiers à utiliser les données de séquençage du génome complet de *C. canephora* récemment publiées pour analyser plusieurs familles de gènes, comme ceux codant pour les protéines du système « tripartite » PYR/PYL/RCAR-SnRK2-PP2C de perception et de transduction du signal ABA. Par comparaison avec les études similaires déjà réalisées sur ces gènes dans d'autres espèces, nos résultats montrent pour la première fois l'existence de duplication de gènes *PYL,* notamment de *CcPYL8.*

Les analyses d'expression ont permis de confirmer ma fonctionnalité de la plupart d ces gènes dans les feuilles et dans les racines. Comme plusieurs travaux montrent également l'importance des gènes du système « tripartite » au cours de la maturation des fruits, des études d'expression plus approfondies devront être réalisées pour analyser l'expression de ces gènes au cours du développement du grain de caféier.

Même si nos résultats ne semblent pas montrer de différences importantes entre les clones des quantités d'ABA, il pourrait être intéressant d'analyser ses quantités à différents temps durant l'établissement du stress. En effet, les quantifications d'ABA réalisées chez les plantes présentant une valeur de Ψ_{pd} de -3,0 MPa en condition de stress (soit après 6 jours de suspension de l'irrigation pour le clone sensible 22, et entre 12 et 15 jours pour les clones tolérants, Marraccini *et al.*, 2011), ne permettent pas de savoir si des variations de teneur en ABA ont eu lieu dans les feuilles et les racines précocement après l'application du stress. Afin de vérifier que le métabolisme de l'ABA n'est pas altéré au sein des différents clones de *C. canephora*, il serait aussi intéressant de tester l'expression des gènes *CcNCED3* et *CcCYP707A1*, respectivement impliqués dans la synthèse et la dégradation de cette phytohormone. Ce travail est d'ailleurs en cours au laboratoire (Costa *et al.,* manuscrit en préparation).

Les résultats présentés dans ce travail confirment ceux précédemment obtenus (Vieira *et al.*, 2013) qui montrent qu'il n'existe pas un mais plusieurs mécanismes impliqués responsables de la tolérance à la sécheresse chez *C. canephora*. En dépit de ce constat, il serait toutefois intéressant de rechercher les polymorphismes nucléotidiques (SNPs : *single-nucleotide polymorphism* et INDELS : *INsertion/DELetion*) au sein des gènes identifiés dans ce travail, par exemple dans les génomes des clones tolérants (14, 73 et 120) et (22) sensibles de *C. canephora* puisque ceux-ci sont maintenant séquencés (A.C. Andrade, communication personnelle). Ces recherches pourraient être menées aussi bien dans les séquences codantes, afin de savoir si les protéines du système « tripartite » sont modifiées au sein des clones tolérants et sensibles à la sécheresse utilisés lors de ce travail, qu'au sein de leurs séquences de régulation (promoteurs), afin de savoir si les différences d'expression observées entre les clones peuvent s'expliquer par l'existence de mutations dans ces séquences, comme cela a récemment été observé pour le gène *CcDREB1D* de *C. canephora* (Alves *et al.*, 2017).

Enfin, l'expression des gènes dont les profils d'expression particuliers ont été mis en évidence lors de ce travail (comme par exemple *CcAHG2* et *CcSnRK2.2*), pourrait être testée dans d'autres clones

tolérants et sensibles à la sécheresse de *C. canephora* (Carneiro *et al.*, 2015) afin de savoir si leurs profils d'expression sont conservés. Si tel devait être le cas, on pourrait alors envisager de les utiliser comme marqueurs moléculaires qui qui pourraient être utilisées dans les programmes de sélection des caféiers pour la création de nouvelles variétés plus tolérantes à la sécheresse.

Figures List

Review of related literature

Chapter 1

Tables List

Review of related literature

Chapter 1

Summary

Preface

The *Coffea* genus belongs to *Rubiacea* family and contains more than 124 species (Davis *et al.*, 2006, 2011) that represents a major agricultural commodity in world trade (ICO, 2016). This genus comprises perennial species, all native to Madagascar, Africa, the Mascarene Island and the Comoros Island. Among all species, *C. arabica* and *C. canephora* are the two economically important species. As provider of a higher quality beverage *C. arabica* is the most cultivated specie (Poncet *et al.*, 2007). *C. arabica* is an allotetraploid (2n = 4x = 44) that was originated 1 million years from the natural hybridization of two ancestral diploid genomes, *C. canephora* and *C. eugenioides* (Lashermes *et al.*, 1999). Due to the self-pollination of the flowers, the species is characterized by a low genetic diversity (Hatanaka *et al.*, 1999). Conversely, *C. canephora* is a diploid species (2n = 2x = 22), it has high genetic variability and ability to adapt to various climatic conditions (Bertrand *et al.*, 2003). However, produces a lower quality coffee, more suitable for the production of instant coffee (Hatanaka *et al.*, 1999).

Currently, the annual world production is around 143.3 million bags (60 Kg) coffee beans (ICO, 2016), being Brazil the largest producer (30,2%). Nowadays, drought and unfavorable temperatures are the major climatic limitations for coffee production, in some marginal regions with no irrigation coffee yields may decrease as much as 80% in very dry years (Damatta & Ramalho, 2006). As a consequence of global warming, coffee-growing geographical regions could also suffer delocalization (Assad *et al.*, 2004). Variatiations in rainfall and temperature also influences biochemical composition of beans (Mazzafera, 2007) affecting directly the final cup quality. There is genetic variability within the *Coffea* genus that could be used to increase drought tolerance and generate coffee varieties better adapted to climatic variations which has been turned into one of the priorities of many coffee research institutes (Marraccini *et al.*, 2012). Elucidate the genetic and molecular mechanisms of drought tolerance is essential to identificate molecular markers that could be used to speed up coffee breeding programmes (Leroy *et al.*, 2011).

Abscisic Acid (ABA), discovered in the 1960s (Ohkuma *et al.*, 1963; Cornforth *et al.*, 1965) is a vital plant hormone synthesized in roots and leaves (Zhang & Davies, 1989; Thompson *et al.*, 2007) which act as central regulator that protects plants against abiotic stresses such drought (Wasilewska *et al.*, 2008; Soon *et al.*, 2012). ABA can accumulate up to 10 to 30-fold in plants under drought stress relative to unstressed conditions (Leung *et al.*, 2012). ABA has been characterized as important endogenous small molecule that mediates stress-responsive gene expression, stomatal closure, and vegetative growth modulation (Rodríguez-Gacio *et al.*, 2009). A great deal of effort has been focused on elucidating the molecular mechanisms underlying ABA sensing and signalling over the past few decades (Umezawa *et al.*, 2010). Recently, two independent research groups discovered novel intracelular ABA receptors, PYL/RCARs, that are involved in ABA sensing and signaling via their direct interaction with clade A PP2Cs in *Arabidopsis thaliana* (Ma *et al.*, 2009; Park *et al.*, 2009). With

the looming prospect of global water crisis, these recent laudable success in deciphering the early steps in the signal transduction of the "stress hormone" ABA has ignited hopes that crops can be engineered with the capacity to maintain productivity while requiring less water input (Leung *et al.*, 2012).

The core of the ABA signaling network comprises a subfamily of type 2C proteins phosphatases (PP2Cs) and three Snfl-related kinases, SnRK2.2, 2.3 and 2.6 (Umezawa *et al.*, 2009; Fujii *et al.*, 2009) whose activities are controlled by ABA. The current ABA signal transduction model can be described as follow: In the absence of ABA, SnRK2 kinases are inactivated by PP2Cs which physically interact with SnRK2 and dephosphorylate a serine residue in the kinase activation loop, a phosphorylation essential for kinase activity (Belin *et al.*, 2006). ABA binds to the ABA receptors family PYR/PYL/RCAR allowing the bounds of the receptors and the catalytic site of PP2Cs to inhibit their enzymatic activity. In turn, ABA-induced inhibition of PP2Cs leads to SnRK2 activation by activation loop autophosphorylation (Boudsocq *et al.*, 2007; Soon *et al.*, 2012).

In the last years, great efforts have been implemented in genomics to attempt to understand the genetic determinism of tolerance to environmental stresses, biotic and abiotic, especially in species models (Umezawa *et al.*, 2006; Ashraf, 2010). The same applies to the case of coffee on which the recent progress in genome sequencing resulted in thousands of EST sequences (Lin *et al.*, 2005; Poncet *et al.*, 2006; Vieira *et al.*, 2006; Mondego *et al.*, 2011), for the construction of genetic maps (Lefebvre-Pautigny *et al.*, 2010; Leroy *et al.*, 2011), improvement of genetic transformation techniques (Ribas *et al.*, 2011) and complete genome sequencing of coffee (De Kochko *et al.*, 2010). These scientific advances have paved the way for studies of genetic determinism and molecular drought tolerance in this plant.

Reference

Alves GSC, Torres LF, Déchamp E. Breitler JC, Joët T, Gatineau F, Andrade AC, Bertrand B, Marraccini P, Etienne H. 2017. Differential fine tune regulation of gene expression in coffee leaves by *CcDREB1D* promoter haplotypes under drought. Accepted for publication in Journal of Experimental Botany.

Ashraf M. **2010**. Inducing drought tolerance in plants: recent advances. *Biotechnology Advances* **28**: 169–183.

Assad ED, Pinto HS, Zullo Junior J, Ávila AMH. **2004**. Impacto das mudanças climáticas no zoneamento agroclimático do café no Brasil. *Pesquisa Agropecuária Brasileira* **39**: 1057–1064.

Belin C, de Franco P-O, Bourbousse C, Chaignepain S, Schmitter J-M, Vavasseur A, Giraudat J, Barbier-Brygoo H, Thomine S. **2006**. Identification of features regulating OST1 kinase activity and OST1 function in guard cells. *Plant Physiology* **141**: 1316–27.

Bertrand B, Guyot B, Anthony F, Lashermes P. **2003**. Impact of the Coffea canephora gene introgression on beverage quality of *C. arabica*. *Theoretical and Applied Genetics* **107**: 387–394.

Boudsocq M, Droillard M-J, Barbier-Brygoo H, Laurière C. **2007**. Different phosphorylation mechanisms are involved in the activation of sucrose non-fermenting 1 related protein kinases 2 by osmotic stresses and abscisic acid. *Plant Molecular Biology* **63**: 491–503.

Cornforth JW, Milborrow BV, Ryback G, Wareing PF. **1965**. Chemistry and Physiology of 'Dormins' In Sycamore: Identity of Sycamore 'Dormin' with Abscisin II. *Nature* **205**: 1269–1270.

DaMatta FM, Ramalho JDC. **2006**. Impacts of drought and temperature stress on coffee physiology and production : a review. *Brazilian Journal of Plant Physiology* **18**: 55–81.

Davis AP, Govaerts R, Bridson DM, Stoffelen P. **2006**. An annotated taxonomic conspectus of the genus *Coffea* (Rubiaceae). *Botanical Journal of the Linnean Society* **152**: 465–512.

Davis AP, Tosh J, Ruch N, Fay MF. **2011**. Growing coffee: Psilanthus (Rubiaceae) subsumed on the basis of molecular and morphological data; implications for the size, morphology, distribution and evolutionary history of *Coffea*. *Botanical Journal of the Linnean Society* **167**: 357–377.

Fujii H, Chinnusamy V, Rodrigues A, Rubio S, Antoni R, Park S-Y, Cutler SR, Sheen J, Rodriguez PL, Zhu J-K. **2009**. In vitro reconstitution of an abscisic acid signalling pathway. *Nature* **462**: 660–664.

Hatanaka T, Choi YE, Kusano T, Sano H. **1999**. Transgenic plants of coffee Coffea canephora from embryogenic callus via Agrobacterium tumefaciens-mediated transformation. *Plant Cell Reports* **19**: 106–110.

De Kochko A, Akaffou S, Andrade AC, Campa C, Crouzillat D, Guyot R, Hamon P, Ming R, Mueller LA, Poncet V, *et al.* 2010. Advances in Coffea Genomics. *Advances in Botanical Research* **53**: 23–63.

Lashermes P, Combes M-C, Robert J, Trouslot P, D'Hont A, Anthony F, Charrier A. **1999**. Molecular characterisation and origin of the *Coffea arabica* L. genome. *Molecular and General Genetics* **261**: 259–266.

Lefebvre-Pautigny F, Wu F, Philippot M, Rigoreau M, Priyono, Zouine M, Frasse P, Bouzayen M, Broun P, Pétiard V, *et al.* 2010. High resolution synteny maps allowing direct comparisons between the coffee and tomato genomes. *Tree Genetics and Genomes* **6**: 565–577.

Leroy T, De Bellis F, Legnate H, Kananura E, Gonzales G, Pereira LF, Andrade AC, Charmetant P, Montagnon C, Cubry P, *et al.* 2011. Improving the quality of African robustas: QTLs for yield- and quality-related traits in *Coffea canephora*. *Tree Genetics & Genomes* **7**: 781–798.

Leung J, Valon C, Moreau B, Boeglin M, Lefoulon C, Joshi-Saha A, Chérel I. **2012**. The ABC of abscisic acid action in plant drought stress responses. *Biologie Aujourd'hui* **206**: 301–312.

Lin C, Mueller LA, Mc Carthy J, Crouzillat D, Pétiard V, Tanksley SD. **2005**. Coffee and tomato share common gene repertoires as revealed by deep sequencing of seed and cherry transcripts. *Theoretical and Applied Genetics* **112**: 114–130.

Ma Y, Szostkiewicz I, Korte A, Moes D, Yang Y, Christmann A, Grill E. **2009**. Regulators of PP2C phosphatase activity function as abscisic acid sensors. *Science* **324**: 1064–1068.

Marraccini P, Vinecky F, Alves GSC, Ramos HJO, Elbelt S, Vieira NG, Carneiro FA, Sujii PS, Alekcevetch JC, Silva VA, *et al.* 2012. Differentially expressed genes and proteins upon drought acclimation in tolerant and sensitive genotypes of Coffea canephora. *Journal of Experimental Botany* **63**: 4191–4212.

Mazzafera P. **2007**. Estresse hídrico e consequências na composição das sementes de café e qualidade da bebida. In: Salva T de JG, ed. Cafés de qualidade: aspectos tecnológicos, científicos e comerciais. Campinas, SP, Brazil: Instituto Agronômico de Campinas, 73–90.

Mondego JM, Vidal RO, Carazzolle MF, Tokuda EK, Parizzi LP, Costa GG, Pereira LF, Andrade AC, Colombo C a, Vieira LG, *et al.* 2011. An EST-based analysis identifies new genes and reveals distinctive gene expression features of Coffea arabica and Coffea canephora. *BMC Plant Biology* **11**: 30.

Ohkuma K, Lyon JL, Addicott FT, Smith OE. **1963**. Abscisin II, an abscission-accelerating substance from young cotton fruit. *Science* **142**: 1592–1593.

Park S-Y, Fung P, Nishimura N, Jensen DR, Fujii H, Zhao Y, Lumba S, Santiago J, Rodrigues A, Chow TFF, *et al.* 2009. Abscisic acid inhibits PP2Cs via the PYR/PYL family of ABA-binding START proteins. *Science* **324**: 1068–1071.

Poncet V, Dufour M, Hamon P, Hamon S, de Kochko A, Leroy T. **2007**. Development of genomic microsatellite markers in Coffea canephora and their transferability to other coffee species. *Genome* **50**: 1156–1161.

Poncet V, Rondeau M, Tranchant C, Cayrel A, Hamon S, De Kochko A, Hamon P. **2006**. SSR mining in coffee tree EST databases: potential use of EST-SSRs as markers for the *Coffea* genus. *Molecular Genetics and Genomics* **276**: 436–449.

Ribas AF, Dechamp E, Champion A, Bertrand B, Combes M-C, Verdeil J-L, Lapeyre F, Lashermes P, Etienne H. **2011**. Agrobacterium-mediated genetic transformation of Coffea arabica (L.) is greatly enhanced by using established embryogenic callus cultures. *BMC Plant Biology* **11**: 92.

Rodríguez-Gacio MDC, Matilla-Vázquez MA, Matilla AJ. **2009**. Seed dormancy and ABA signaling: the breakthrough goes on. *Plant Signaling & Behavior* **4**: 1035–1049.

Soon F, Ng L, Zhou XE, West GM, Xu Y, Chalmers MJ, Brunzelle JS, Zhang H, Yang H. **2012**. Molecular mimicry regulates ABA signaling by SnRK2 kinases and PP2C phosphatases. *Science* **335**: 85–89.

Thompson AJ, Mulholland BJ, Jackson AC, Mckee JMT, Hilton HW, Symonds RC, Sonneveld T, Burbidge A, Stevenson P, Taylor IB. **2007**. Regulation and manipulation of ABA biosynthesis in roots. *Plant, Cell and Environment* **30**: 67–78.

Umezawa T, Fujita M, Fujita Y, Yamaguchi-Shinozaki K, Shinozaki K. **2006**. Engineering drought tolerance in plants: discovering and tailoring genes to unlock the future. *Current Opinion in Biotechnology* **17**: 113–122.

Umezawa T, Nakashima K, Miyakawa T, Kuromori T, Tanokura M, Shinozaki K, Yamaguchi-Shinozaki K. **2010**. Molecular basis of the core regulatory network in ABA responses: sensing, signaling and transport. *Plant & Cell Physiology* **51**: 1821–1839.

Umezawa T, Sugiyama N, Mizoguchi M, Hayashi S, Myouga F, Yamaguchi-Shinozaki K, Ishihama Y, Hirayama T, Shinozaki K. **2009**. Type 2C protein phosphatases directly regulate Abscisic acid-activated protein kinases in Arabidopsis. *Proceedings of the National Academy of Sciences USA* **106**: 17588–17593.

Vieira LGE, Andrade AC, Colombo CA, Moraes AH de A, Metha Â, Oliveira AC de, Labate CA,

Marino CL, Monteiro-Vitorello CDB, Monte DDC, *et al.* 2006. Brazilian coffee genome project: an EST-based genomic resource. *Brazilian Journal of Plant Physiology* **18**: 95–108.

Wasilewska A, Vlad F, Sirichandra C, Redko Y, Jammes F, Valon C, Frey NF dit, Leung J. **2008**. An update on abscisic acid signaling in plants and more… *Molecular Plant* **1**: 198–217.

Zhang J, Davies WJ. **1989**. Abscisic acid produced in dehydrating roots may enable the plant to measure the water status of the soil. *Plant, Cell and Environment* **12**: 73–81.

1 Review of Related Literature

1.1 International market

Coffee is the most widely traded tropical agricultural commodity in the world, cultivated around 11 million hectares (ha) in over 80 countries from Africa, Asia, and the Americas. Small stakeholders account for approximately 70% of world coffee production and coffee trade has economic relevance as source of employment for millions of people worldwide.

In 2015/16, the annual world production was around 143.3 million bags of coffee beans (ICO 2016). The coffee trade statistic showed an increase of 0,7% in global coffee production in 2015/16 crop year compared to 2014/15. Estimated increase in global production of Robusta coffee represented 1.7% in 2015/2016 while no changes were estimated in global production of Arabica coffees in 2015/2016. Over the last four years, Robusta worldwide production increased from 39% to 42,15% as Arabica production decreased in 3,15% (Figure 1).

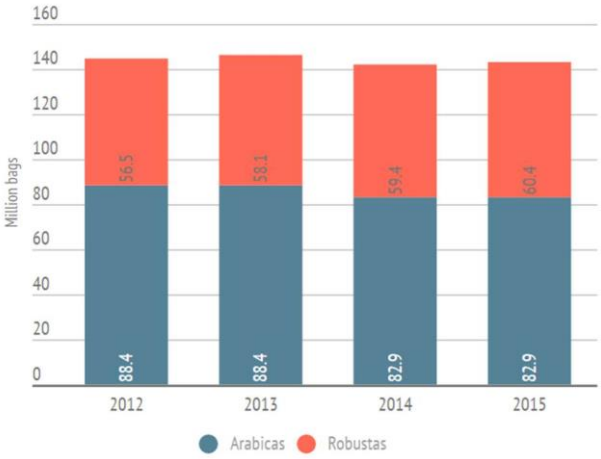


Figure 1 Coffee trade statistic for the last four years of crop production.
Source: ICO, 2016.

Brazil is the major coffee world producer for more than a century and currently responsible for a third of global production (30,2%), followed by Vietnam (19,2%), Colombia (9,42%), Indonesia (8,59%) and Ethiopia (4,67%) (ICO, 2016). Altogether, these exporting countries contributed around 72% of coffee world production in the 2015/2016 crop year (Table 1).

Table 1 Total coffee production by all exporting countries (in thousands 60 Kg bags) for the last six crop years.

Crop year	2010/11	2011/12	2012/13	2013/14	2014/15	2015/16
Brazil (A/R)	48.095	43.484	50.826	49.152	45.639	43.235
Vietnam (R/A)	20.000	26.500	23.402	27.610	26.500	27.500
Colombia (R/A)	8.523	7.652	9.927	12.124	13.333	13.500
Indonesia (R/A)	9.129	10.644	11.519	11.265	11.418	12.317
Ethiopia (A)	7.500	6.798	6.233	6.527	6.625	6.700
TOTAL	**134.246**	**140.617**	**144.960**	**146.506**	**142.278**	**143.306**

Source: ICO, 2016.

In 2016 crop year, Brazilian coffee yield is projected at 49,6 million bags and the total area planted is around 1.942,1 thousand ha (CONAB, 2015). Minas Gerais (MG) is the major coffee producer state with 28,5 million bags (57,46%) in which Arabica species represented 67,35% of total coffee area planted in Brazil. On the other hand, Espirito Santo (ES) is the second producer state with 9,5 million bags (19,15%) mainly planted with Robusta (CONAB, 2015).

Arabica trees are forecast to produce 38 million bags in 2015/2016 crop year, up 3.8 million bags compared to the previous season. On the other hand, Robusta production in 2015/16 is expected to decrease to 14.4 million bags, down 2.6 million bags from the previous crop year, especially due to lower agricultural yields in Espirito Santo as a result of a prolonged dry spell and above average temperatures during the summer months. In addition, Espirito Santo has also faced shortage of water resources, limiting the use of irrigation in coffee plantations which are fairly common in that state (GAIN, 2016). Coffee is also growing in other Brazilian states like São Paulo (10%), Bahia (7,6%), Paraná (2,18%), Rondônia (4,44%) and Goias (6,1%) (CONAB, 2015).

The benefits of coffee consumption are being perceived by consumers and the demand is currently rising. Several epidemiological studies suggest that moderate coffee consumption (3-4 cups/day) may prevent several chronic diseases (Higdon and Frei, 2006) such as diabetes (including type 2 diabetes mellitus) (van Dam and Feskens, 2002; Carlsson, 2004), cardiovascular (coronary heart disease, congestive heart failure, arrhythmias) (O'Keefe *et al.*, 2013), chronic liver illnesses (cirrhosis and hepatocellular carcinoma) (Gallus *et al.*, 2002) and neurodegenerative (Parkinson's, Alzheimer) (Lindsay *et al.*, 2002; van Gelder *et al.*, 2007; Campdelacreu, 2014) ones.

The first bitter mouthful in the morning which gives energy to the planet daily is coffee, one of the most consumed beverages in the world with more than 2.25 billion cups consumed every day. The global coffee consumption was estimated to 149.3 million bags (60kg of green beans) in 2014 (ICO, 2016). Since 2011, coffee consumption averaged annual growth rate of 2.3% (Figure 2).

Figure 2 Global coffee consumption. During last years the demand increase in many countries including traditional markets, exporting counties and emerging markets.
Source: ICO, 2016.

During the last few years, the demand increased in many countries, particularly in traditional markets (Canada, European Union [EU], Japan, Norway, Switzerland, USA and others), but was also sustained by emerging markets (Algeria, Australia, Russia, South Korea, Turkey, Ukraine, others) and exporting countries (e.g. Brazil) (ICO 2016). A total of 112.372 thousand bags was imported in 2016, USA being the first in the rank of importing countries with 27.016 thousand of bags (24%). On the other hand, EU imported 72.246 thousand bags (64,2%), among them Germany (18,8%), Italy (7,86%) and France (5,97%) stands out as coffee importers. Japan is closer to Italy with 7,46% of world importations (ICO, 2016).

Currently, the total domestic consumption by all exporting countries is 47.633 thousand bags (Table 2). Besides Brazil being the main producer, it also leads consumption among exporting countries (42.9%) followed by Indonesia (9,36%) and Ethiopia (7,73%) (Table 2). On the other hand, the European Union stands out (39,82%) the ranking of consumption among importing countries, USA (23,37%) and Japan (7,36%) are in the second and third position, respectively (Table 2).

Table 2 World coffee consumption (in thousand 60Kg bags) for the last four calendar years.

Calendar year	2012	2013	2014	2015
Exporting countries	**44.711**	**455.222**	**46.649**	**47.633**
Brazil	20.178	20.146	20.271	20.458
Indonesia	3.842	4.100	4.292	4.458
Ethiopia	3.387	3.463	3.656	3.681
Importing countries	**98.719**	**102.289**	**103.740**	**104.572**
European Union	41.018	41.875	42.215	41.638
USA	22.232	23.417	23.767	24.441
Japan	7.131	7.435	7.494	7.695
TOTAL	**143.430**	**147.811**	**150.389**	**152.204**

Source: ICO, 2016.

To attend the increasing world consumption of coffee, it is necessary to overcome some challenges in production. Nowadays, drought and high temperatures are the major climatic limitations for world coffee production (DaMatta and Ramalho, 2006). These abiotic stresses are expected to become increasingly important in several coffee growing regions due to the recognized changes in global climate and also because coffee cultivation has spread towards marginal lands, where water shortage and unfavorable temperatures constitute major constraints to coffee yield.

1.2 Global Climate Change: impacts in coffee production

1.2.1 Impacts in coffee areas

Global climate change is becoming more unpredictable and abiotic stresses are the major cause of decreasing the average yield of principal crop species (Hazarika *et al.*, 2013). Climate changes is occurring at rates never experienced before by modern agriculture, with temperatures planned to increase of 2-3°C over the next 40 years (Hatfield, 2013). This will affect all not only growth and development of plants, but also the quality of their products. When evaluating the effects of climate changes on plants, it is important to include the direct effects of perennial plants because adaptation strategies for these production systems are more complex than in annual crops.

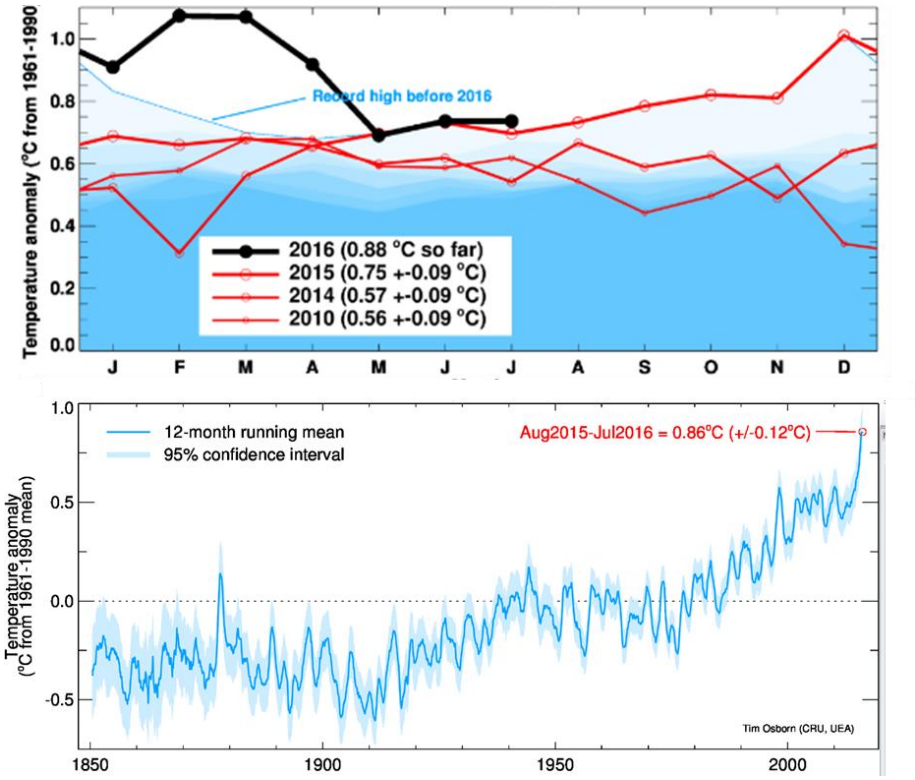


Figure 3 The top panel shows global-mean temperature anomalies for the current year so far (black). The red lines show the monthly temperature anomalies for the 3 warmest years. The blue line near the top shows the record high for each individual month prior to the current year. The bottom graph shows series and 12-month running means values yearly global temperature graphs anomaly time series 1850-2010.
Source: available at https://crudata.uea.ac.uk/~timo/diag/tempdiag.htm (Morice *et al.*, 2012).

As a consequence of global warming, coffee-growing geographical regions could also suffer important geographical delocalization (Assad *et al.*, 2004). In marginal regions without irrigation or during dry seasons, this could led in decreasing coffee yields as much as 80% (DaMatta & Ramalho, 2006). Analyzing the effects of recent climate change by extrapolating the historical tendencies in temperature and precipitation to 2020 in coffee producing areas in Veracruz, Mexico, the analysis predict that coffee production is likely to decline about 34%. The suitability for coffee crops in Costa Rica, Nicaragua and El Salvador will be reduced by more than 40% (Glenn *et al.*, 2014) while the loss of climatic niches in Colombia will force the migration of coffee crops towards higher altitudes by mid-century (Ramirez-Villegas *et al.*, 2012).

In Brazil, it is expected that coffee areas will migrate towards more favorable zones in the South of country under future climate change (Assad *et al.*, 2004). Some studies have mapped the changes in area suitable for coffee production in the four main coffee producing states as a consequence for global warming (Assad *et al.*, 2004; Pinto *et al.* 2007). According to the last report of the Intergovernmental Panel on Climate Change (IPCC, 2014), an increase of 3º C in temperature would lead to major changes in the distribution of coffee producing zones. In the main coffee producing states of Minas Gerais and

São Paulo, the potential area for production would decline from 70-75% of the states to 20-25%, while coffee area would be reduced by 10% in Paraná and production would be eliminated in Goias state (Figure 4). The new areas suitable for coffee production that could emerge in Santa Catarina and Rio Grande do Sul will only partially compensate the loss of area in other states (Pinto & Assad, 2008).

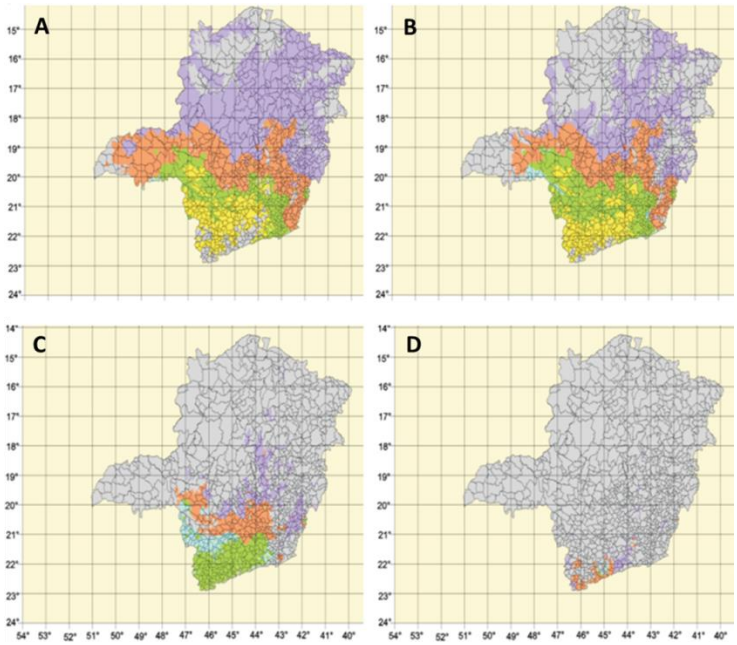


Figure 4 Current coffee zoning for Minas Gerais State (A), with the increase of 1°C in temperature and 15% in rainfall (B), considering 3°C rise in temperature and 15% in rainfall (C); with the increase of 5,8°C in temperature and 15% in rainfall. The colored regions indicates: irrigation required (purple); suitable for cultivation (green); irrigation recommended (orange); frost risk (yellow); thermal excess (light blue); unsuitable for cultivation (gray).
Source: Adapted from Assad *et al.*, 2004.

These forthcoming scenarios require new approaches that develop innovative strategies to manage the crop production system and reduce the impact of climate change. Strategies such as the developing agroforestry production systems, increase irrigation, and modify agricultural practices maintaining cover crops are projected to become more frequent.

1.2.2 Impact in term of abiotic stress

Under drought and high temperatures, some coffee pests and diseases should also become more severe. The occurrence of leaf miners (*Leucoptera coffeella*) disease has been increasing over recent years in coffee yields as a consequence of dry conditions (Assis *et al.*, 2012). On the other hand, leaf rust disease is rising with warmer temperatures. Likewise, the number of cycle life generations of *Hyphotenemus hampei* has been increasing under the same climatic conditions, as a result, a thermal tolerance of the coffee berry borer has been demonstrated (Jaramillo *et al.*, 2009). In the case of

Leucoptera coffeella and *Meloidogyne incognita*, the same circumstances have been predicted in Brazil under these climate change conditions. Therefore, the coffee production demands nowadays plants better adapted to both abiotic and biotic stresses.

1.2.3 Impact on coffee plants

Long periods of drought can beget diverse effects on coffee plants. Moderate drought can promote leaf falling, delay and un-synchronize flowering, reduce vegetative growth of plagiotropic branches and consequently production potential in following crop year, upon severe drought yet major effects are expected up to plant death, abortion of flowering and fruits.

Besides the loss of coffee production and changes in distribution of coffee producing zones, the biochemical composition of beans could also be modified by drought. Variations in rainfall and temperatures affect sugar, proteins and caffeine contents (Mazzafera, 2007) and consequently the beverage quality (Camargo *et al.*, 1992; Vinecky *et al.*, 2016). Moreover, the predicted climate change and the increasing world population will lead to a growing demand for water and reveal the urgent need for drought tolerant crops (Alter *et al.*, 2015).

Nowadays, coffee production demands plants better adapted to both abiotic and biotic stresses. In such way, it is worth noting that the drought-tolerant (D^T) clone 14 of *C. canephora* (Marraccini *et al.*, 2012) was also recently reported to present durable multiple resistant plant to root-knot nematodes of *Meloidogyne* spp. (Lima *et al.*, 2015).

1.3 *Coffea* genus

The *Coffea* genus belongs to Rubiaceae family, the fourth largest flowering plant family in the world, consisting of more than 11.000 thousand species in 660 genera (Robbrecht & Manen, 2006) which represent 10 to 20% of the total plant species diversity. The most economically valuable genus is *Coffea* that contains 124 species which comprises perennial species all native from Madagascar, Africa, the Mascarene Island, the Comoros Island, Asia and Australia (Davis *et al.*, 2006, 2012).

Among all species, *C. arabica* and *C. canephora* are the two economically important species corresponding to 65% and 35% of the international market, respectively (ICO, 2016). The two species are perennial woody trees and display considerable variation in morphology, size, and ecological adaptation (Combes *et al.*, 2015). Nevertheless, *C. arabica* is an allotetraploid (2n = 4x = 44) that was originated 1 million years from the natural hybridization of two ancestral diploid genomes, *C. canephora* and *C. eugenioides* (Figure 5). As provider of a higher quality beverage *C. arabica* is the most cultivated specie (Poncet *et al.*, 2007).

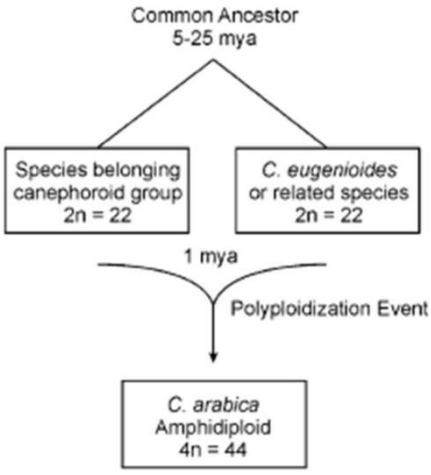


Figure 5 Evolutionary history of allotetraploid *C. arabica*. The progenitor genomes are represented by diploid *C. eugenioides* and *C. canephora*. *C. arabica* arose 1 to 2 million years ago (mya) from the fusion of *C. canephora* (or related species) and *C. eugenioides*.
Source: Vidal *et al.*, 2010.

1.3.1 *Coffea arabica*

Originally from Southwest Ethiopia and Plateau of Sudan, *C. arabica* was cultivated about 1,500 years ago, firstly in Ethiopia. The genetic background of the current *C. arabica* cultivars comes from Typica and Bourbon (Anthony *et al.*, 2002). As a predominant autogamous (natural self-pollinating) specie, *C. arabica* present low genetic diversity (Hatanaka *et al.*, 1999) and has a total genome size estimated by flow cytometry at around 2.62 x 10^3 Mb (Clarindo & Carvalho, 2009). The breeding programs nowadays have been search new cultivars with improved traits such as beverage cup quality, flowering time synchronicity, resistance to pests, and drought stress tolerance.

As *C. arabica* is an amphidiploid species (originating from a natural hybridization event between *C. canephora* and *C. eugenioides*), its transcriptome is a mixture of homologous genes expressed from these two subgenomes in which *C. eugenioides* is assumed to expressed genes mainly for proteins involved in basal biological process as photosynthesis, while *C. canephora* sub-genome is assumed to regulate Arabica gene expression by expressing genes for regulatory proteins and adaptation process (Vidal *et al.*, 2010).

1.3.2 *Coffea canephora*

C. canephora is a cross-pollinated diploid species (2n = 2x = 22) that has high genetic variability in its haploid genome of 710 Mb (Denoeud *et al.*, 2014). Thereby, exist genetic variability within the *Coffea* genus that could be used to increase drought tolerance and among commercial species *C. canephora* stands out. Despite the ability of *C. canephora* to adapt regarding various climatic conditions

(Bertrand *et al.*, 2003), it produces beans giving lower quality beverage that are more used in instant coffee drinks (Hatanaka *et al.*, 1999).

C. canephora genetic diversity can be divided in two major clades according to their geographical origins: the Guinean group (G) and the Congolese group. The Congolese group can be subdivided into SG2/B, C, SG1 and UW (Montagnon and Leroy, 1993) (Figure 6). Guinean genotypes are considered the most tolerant to drought and genotypes from the SG1 Congolese group are more tolerant to drought than those from the SG2 Congolese group (Montagnon and Leroy, 1993). The considerable genetic diversity observed in *C. canephora* is still largely unexploited. During the last decade, several breeding programs to development of new *C. canephora* clones have attempted to explore the genetic diversity of *C. canephora*. In Brazil, a genetic improvement program for the development of new cultivars, using SG1 genotypes as source of genetic variability, characterized a clonal variety of *C. canephora* Conilon highly productive under drought conditions (Ferrão *et al.*, 2000).

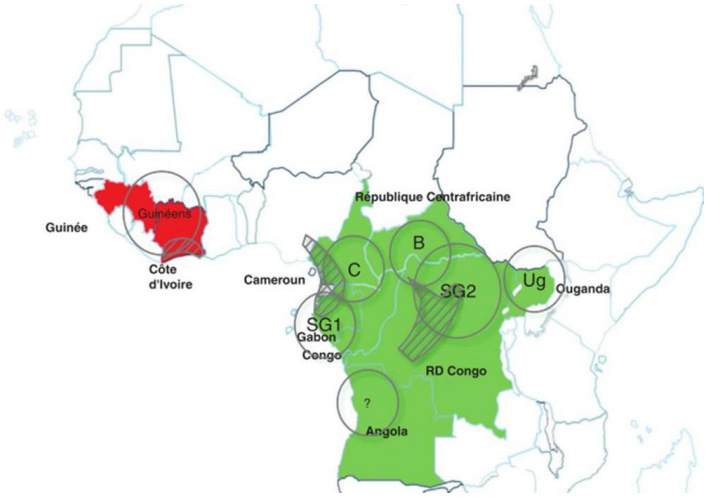


Figure 6 Geographic origin of the two main genetic group of *C. canephora*. In red: geographic origin of the Guinean group. In green: geographic origin of the Congolese subgroups (SG). The circles highlight the identification of each subgroup.
Source: Montagnon *et al.*, 2012.

1.3.3 Other *Coffea* species

Even though *Coffea* genus diverged recently (5 to 25 million years ago) from others plants, most of their species are genetically highly related thus permitting natural or manual hybridizations that could be used in coffee breeding programs. For instance, it has been introduced in *C. arabica* by breeding programs resistance genes for leaf rust (*Hemileia vastatrix*), for the *Meloidogyne* nematodes, and to *Colletotrichum kahawae* fungus agent of Coffee Berry Disease (CDB) (Bertrand *et al.*, 2003).

In this sense, the diploid species *C. racemosa* presents high resistance to drought and elevated temperatures. In its native habitat, *C. racemosa* is able to adapt to regions where the annual rainfall does

not exceed 1000 mm and where dry seasons vary from four to six months (Krug, 1965; Dublin, 1968). This specie presents deeper growth of primary root and lower growth of secondary roots allowing this specie to explore deeper soil layers in water deficit conditions (Fazuoli, 1975). *C. racemosa* had the longest root system in comparative analyses with other coffee species (*C. canephora, C. arabica, C. liberica and C. congensis)* and the root system is mainly contrasting with *C. congensis* root system which survived in a natural environment completely different of *C. racemosa* (Dublin, 1968).

Medina Filho *et al.* (1977b) had evaluated the genetic material of *C. racemosa* from Campinas (Brazil), and they verify that triploids (*C. arabica* x *C. racemosa*) as well as individuals belonging to the second generation backcrosses to *C. arabica* were highly resistant to drought, while Catuai and Acaia cultivars of *C. arabica* (positive controls of the experiment), were highly sensitive. While these cultivars lose a lot of leaves the plants which derivate of *C. racemosa* keep their leaves notably turgid.

1.4 Drought responses in plants

Drought is one of the major constraints of plant productivity worldwide. Under field conditions, plant performance in terms of growth, development, biomass accumulation and yield depends on acclimation ability to the environmental changes and stresses, exercising specific tolerance mechanisms that involve a complex network of biochemical and molecular processes (Wang *et al.*, 2003). When exposed to reduce water availability plants exhibit various physiological responses. For instance, a pivotal reaction is stomatal closure to avoid water loss by transpiration. The resulting reduced availability of carbon dioxide together with a down regulation of photosynthesis-related genes lead to decrease in carbon assimilation restricting plant growth and productivity (Alter *et al.*, 2015). Under drought stress conditions, an increase in photorespiration leads to an accumulation of reactive oxygen species (ROS), which are toxic for cellular components and will eventually lead to cell death (Mittler, 2002). Plants have evolved a number of molecular and physiological adaptation mechanisms to cope with reduced water availability which can be categorized into drought avoidance and drought tolerance (Verslues *et al.*, 2006) (Figure 7).

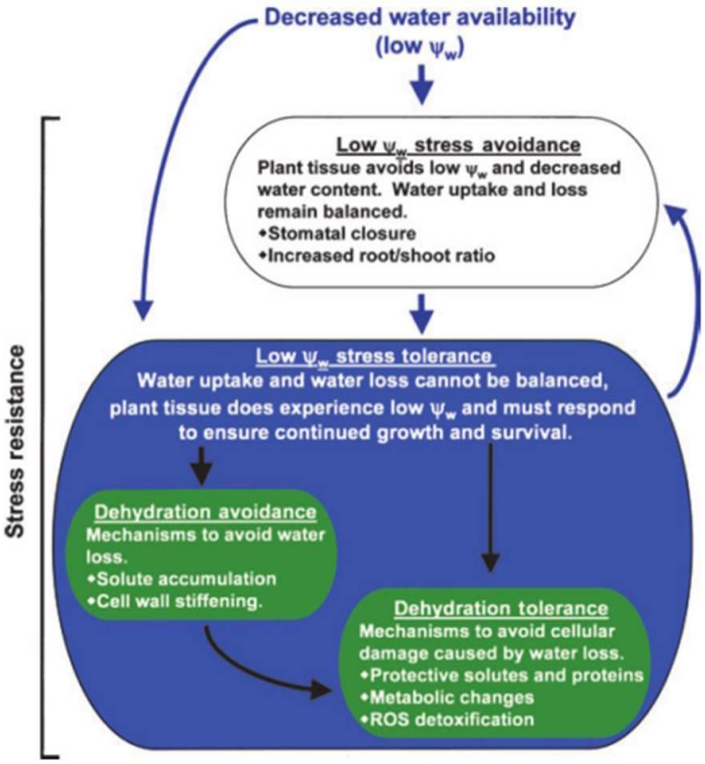


Figure 7 Conceptual diagram of the stress tolerance/stress avoidance model of low-Ψ_w responses.
Source: Verslues *et al.*, 2006.

In most cases, the plant first response is avoid low Ψ_w. Tissue Ψ_w and water content are maintained close to the unstressed level by increasing water uptake or limiting water loss by such that the rates of water loss and water uptake remain balanced. Such a balance is achieved in the short term mainly by stomatal closure. In long term, changes in root and shoot growth, leading to an increased root/shoot ratio, tissue water storage capacity and cuticle thickness and water permeability are also of potential importance. Of these, changes in root growth to maximize water uptake are of the greatest importance for crop plants (Verslues *et al.*, 2006).

Furthermore, these mechanisms for avoiding water loss do not themselves offer any protection from the effects of low Ψ_w if the stress becomes more severe and the plant is no longer able to maintain a balance between water uptake and loss. When stomata are closed because of stress, transpiration is minimized, the Ψ_w of the plant will equilibrate with that the water source (most of cases Ψ_w of the soil). When soil water content and Ψ_w are low, the Ψ_w of the plant tissue must also decrease, either through water loss or by adjustment made by the plant (dehydration avoidance) to achieve a low Ψ_w while avoiding a water loss. The main mechanism of dehydration avoidance are accumulation of solutes and cell wall hardening (Verslues *et al.*, 2006).

The Ψ_w of a walled cell, such as a plant cell, is governed by the equation: $\Psi_w = \Psi_s + \Psi_p$, where Ψ_s is the osmotic potential and Ψ_p is the pressure potential (turgor pressure). At a given Ψ_w, a higher Ψ_p

can be achieved by accumulating solutes inside the cell, thus lowering Ψ_s. The accumulation of additional solutes in response to low Ψ_w is termed osmotic adjustment (Zhang *et al.*, 1999). Osmotic adjustment refers to the active accumulation of additional solutes in response to low Ψ_w (after the effect of reduced water content on the concentration of existing solutes has been factored out). Thus, many plants accumulate one or more types of compatible solutes, such as proline or glycine betaine, in response to low Ψ_w (Verslues *et al.*, 2006). Compatible solutes can also protect protein and membrane structure under dehydration (Hincha & Hagemann, 2004).

In this way, a key regulatory which control plant responses to many types of abiotic stress (including low Ψ_w) is the phytohormone abscisic acid (ABA). It accumulates in response to abiotic stress and regulates the processes involved several the aspects of the low-Ψ_w response. For instance, ABA-regulated stomatal conductance, root growth and seed dormancy (Schroeder, *et al.* 2001; Sharp & LeNoble, 2002; Kermode, 2005) which are important in avoidance of low Ψ_w. Moreover, ABA induces accumulation of compatible solutes which can be crucial for dehydration avoidance (Ober & Sharp, 1994) and ABA also regulates dehydrins and LEA proteins synthesis, important for dehydration tolerance (Sivamani *et al.*, 2000). Thus, at the level of the organism, it seems that a main function of ABA is to coordinate the various aspects of low-Ψ_w response.

1.4.1 Coffee genetic diversity and drought

Among the strategies displayed by coffee plants to cope with drought, leaf folding and inclination that reduce the leaf surface (Figure 8), water loss by transpiration and exposure to high irradiance were commonly observed for Guinean and SG1 genotypes (Montagnon & Leroy, 1993). Leaf abscission is then reduced, favoring a rapid recovery of vegetation with the return of the rains. Such a trait can be considered as a selective advantage when compared with the leaf abscission that characterizes SG2 genotypes (Marraccini *et al.*, 2012).

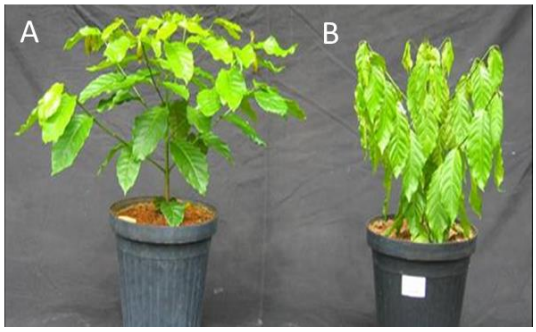


Figure 8 *C. canephora* clones (A: Drought tolerant, D^T; B: Drought susceptible, D^S) grown in greenhouse and submitted to drought conditions.

Several drought-tolerant clones (D^T) of *C. canephora* var. Conilon have been characterized as vigorous plants with high productivity throughout years under drought stress (Ferrão, 2000; Fonseca, 2004). Fingerprint analyses also revealed that these Conilon clones belong to the SG1 group of *C. canephora* (Lambot *et al.*, 2008; Montagnon *et al.*, 2012).

Regarding *Coffea arabica*, the study of populations from Ethiopia growing under contrasting climatic conditions also revealed that this species exhibited phenotypic plasticity in response to varying soil moisture conditions (Beining 2007). It is well known that a genetic variability for drought tolerance also exists in *C. arabica*. For instance, the cultivar IAPAR59 (I59), which the result of a cross between the Timor hybrid HT832/2 and the Villa Sarchi cultivar is considered more tolerant to drought than the Rubi cultivar that did not undergo recent introgression with *C. canephora* genomic DNA (Marraccini *et al.*, 2001; Mofatto *et al.*, 2016)

Figure 9 Contrasting phenotypes of the drought-tolerant I59 (A) and drought-susceptible Rubi (B) cultivars of *C. arabica* in response to a drought period of around 200 days without rainfalls (Embrapa Cerrados).

Major differences between these two cultivars concerned their phenotypic behavior (Figure 9) as predawn leaf water potential, Ψ_{pd} (Figure 10) and transcriptome expression profiles. Marraccini *et al.* (2011) evaluated the effect of drought in leaves of young plantas of *C. arabica* cv. I59 and Rubi cultivars grown in field with irrigation (I) or without (NI) irrigation during two consecutive years (2008 and 2009).. As result, the Ψ_{pd} values measured during the dry season of 2008 and 2009 were almost less negative for the D^T I59 than for D^S Rubi, indicating a better access to soil water for the former compared to the latter (Marraccini *et al.*, 2011).

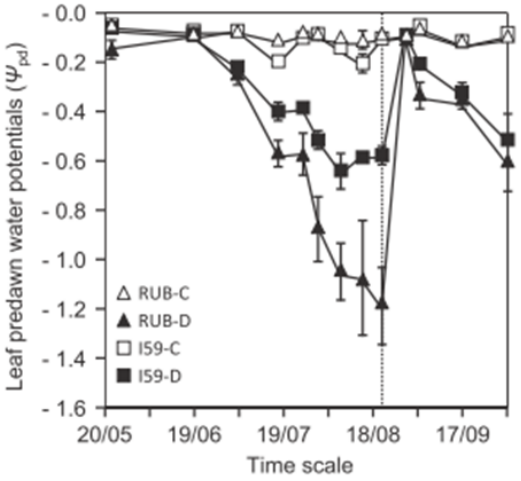


Figure 10 Predawn leaf water potential (Ψ_{pd}) measured in plants of *C. arabica*, Rubi (RUB, triangle) and IAPAR (I59, square) cultivars were grown under control (C, open symbols) and drought (D, black symbols) conditions. Ψ_{pd} values (expressed in mega-Pascal, MPa) were measured once a week during the 2009 dry season (23-month-old plants).
Source: Mofatto *et al.*, 2016.

1.4.2 Physiological responses

The Kouillou (SG1) group of *C. canephora* appears to be more tolerant to water deficit than Robusta (SG2) (Montagnon & Leroy, 1993). SG1 group maintain stomatal opening and consequently active photosynthesis, while stomata of SG2 plants were completely closed under drought conditions.

Besides that, more efficient root water absorption for the SG1 plants could explain its drought tolerance albeit its maintenance of stomatal opening (Boyer, 1969). Physiological analyses also suggested that drought tolerance could be a direct consequence of better root development (Pinheiro *et al.*, 2005) (Figure 11).

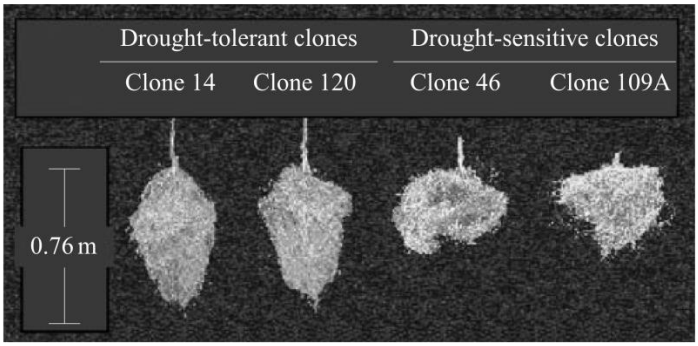


Figure 11 Typical root systems of four clones of Robusta coffee grown under full irrigation.
Source: Pinheiro *et al.*, 2005.

One of the physiological parameters that distinguish the drought-susceptible (D^S) clone 22 of *C. canephora* var. Conilon from the D^T clones 14, 73 and 120 is the rate of decrease in the predawn leaf

water potential (Ψ_{pd}) (RDPW) (Pinheiro *et al.*, 2004). To reach the imposed Ψ_{pd} of -3.0 MPa under the stressed (NI) conditions in the greenhouse, the RDPWP decrease faster for the D^S clone 22 than for the D^T clones (Figure 12). In this condition, the clone D^S 22 reached the Ψ_{pd} of -3.0 MPa within six days, while clones 14, 73 and 120 reached the same within 12, 15 and 12 days, respectively (Marraccini *et al.*, 2011).

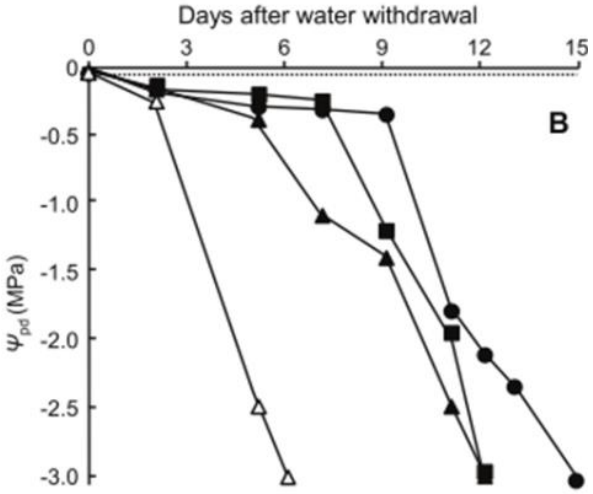


Figure 12 The evolution of predawn leaf water potential (Ψ_{pd}) in the leaves of *C. canephora*. The clones 14, 22, 73 and 120 of *C. canephora* var. Conilon were grown in a greenhouse under water stress. For each clone, Ψ_{pd} evolutions are presented.
Source: Marraccini *et al.*, 2011.

According to DaMatta *et al.* (2003), the better crop yield of a drought-tolerant clone compared with a drought-sensitive clone is mainly associated with the maintenance of leaf area and tissue water potential that are consequences of reduced stomatal conductance (g_s). The D^T and D^S clones of *C. canephora* are important models of study once a lot of physiological and molecular parameters were already evaluated in these plants concerning drought stress under controlled conditions. It is worth noting that the drought-tolerant (D^T) clone 14 of *C. canephora* (Marraccini *et al.*, 2012) was also recently reported to present durable multiple resistant plant to root-knot nematodes of *Meloidogyne* spp. (Lima *et al.*, 2015).

1.4.3 Biochemical responses

The activity of antioxidant enzymes might also be involved in the drought tolerance mechanism (Vieira *et al.*, 2006). A key role of ascorbate peroxidase (APX) was postulated to allow clone 14 to cope with potential increases of H_2O_2 under drought conditions, as an increased (38%) activity of this enzyme was found for this clone upon drought stress (Lima *et al.*, 2012; Pinheiro *et al.*, 2004). Praxedes *et al.*

(2005) observed a maintenance of SPS activity with the decrease of pre-dawn leaf water potential (Ψ_{pd}) for the drought-tolerant clone 120 but not for the drought-sensitive clones.

1.4.4 Molecular responses

Several differentially expressed genes and proteins were investigated in leaves of drought-tolerant and susceptible *C. canephora* clones upon drought acclimation. Genes coding for protein functioning as secondary messengers (*CcNSH1*, *CcEDR1* and *CcEDR2*), related to abscisic acid (ABA) perception and signal transduction (*CcPYL3*, *CcPYL7* and *CcPP2C*), transcription factors (*CcABI5*, *CcAREB1*, *CcRD26*, *CcDREB1*), photosynthesis (*CcPSBP*, *CcPSBQ*, *CcRBCS1*), and drought protection (*CcHSP1*, *CcDH3*, *CcAPX1*), were previously characterized (Marraccini *et al.*, 2012; Vieira *et al.*, 2013).

Recently, among the 42 genes showing up-regulated expression in plagiotropic buds of plants submitted to drought were *CaSTK1* (coding a protein kinase), *CaSAMT1* (coding a protein involved in abscisic acid biosynthesis), *CaSLP1* (coding a protein involved in plant development) and several "no-hit" (orphan) genes of unknown function. Under water scarcity, the expression of *nsLTPs* (coding non-specific lipid-transfer proteins) was greatly up-regulated specifically in plagiotropic buds of I59 which could explain the thicker cuticle observed on the abaxial leaf surface in the D^T I59 compared with the D^S Rubi (Mofatto *et al.*, 2016).

All this information could be used to generate molecular markers to be used in *Coffea* breeding programs for both *C. arabica* and *C. canephora* plant. In this context, 436 plants of *C. canephora* (LxPy) were selected among a population of 3500 individuals from 48 progenitors based on traits of interest such as precociousness of fruit, plant vigor, productivity in field (Carneiro *et al.*, 2015). These plants grown in field conditions since 2009/2010 were submitted to drought conditions and evaluated for their productivity and Ψ_{pd} under drought (winter) season (Figure 13). This allowed the identification of productive and drought-tolerant plants (e.g. L13P63, L8P68 and L5P47) that contrasted with drought-susceptible and lower productive plants (L12P57, L12P100 and L15P14).

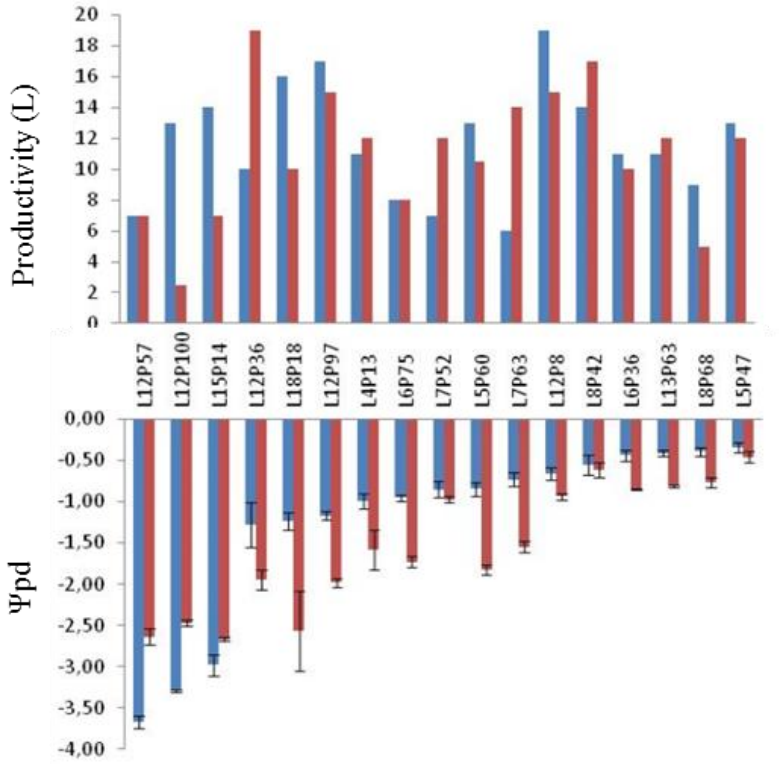


Figure 13 The productivity (measured in liters of cherries per plant) and Ψ_{pd} of LxPy plants of *C. canephora* Conilon grown in field conditions (Embrapa Cerrados) under drought stress. These values were measured during two years (2009: blue isobars and 2010: red isobars).
Source: Carneiro *et al.*, 2015.

1.5 ABA structure and biological roles

The abscisic acid (ABA), discovered in the 1960's (Ohkuma *et al.*, 1963; Cornforth *et al.*, 1965) is a vital hormone synthesized mainly in leaves and roots of the plants (Zhang & Davies, 1989; Thompson *et al.*, 2007), acting as central regulator that protects plants against abiotic stresses such as drought (Wasilewska *et al.*, 2008; Soon *et al.*, 2012). This sesquiterpenoid molecule ($C_{15}H_{20}O_4$) naturally occur in its S-(+)-ABA form, despite the R-(-)-ABA form is active in some assays (Cutler *et al.*, 2010) (Figure 14).

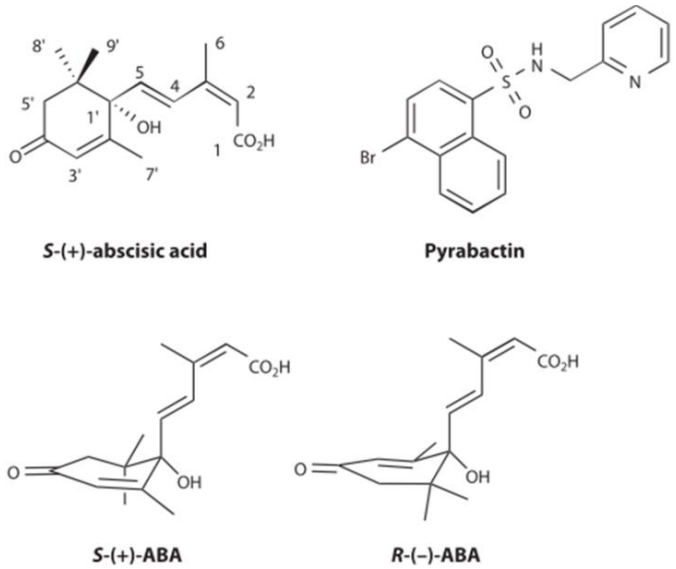


Figure 14 Chemical structures. At bottom is an illustration of the ability of an abscisic acid (ABA) stereoisomer to be rotated along its lengthwise plane to maintain positioning of polar functional groups.
Source: Cutler *et al.*, 2010.

ABA has been shown to control many aspects of plant growth and development as embryo maturation, seed dormancy, germination, cell division and elongation and floral induction (Finkelstein, 2013). ABA is well known as 'stress hormone' and it plays a key role not only during drought (Santiago *et al.*, 2009; Gonzalez-Guzman *et al.*, 2014) but under other abiotic stresses such as salinity (Pons *et al.*, 2013), cold (Bhyan *et al.*, 2012; Shinkawa *et al.*, 2013) and UV radiation (Tossi *et al.*, 2012; Chen *et al.*, 2013). Moreover, ABA has an important function as well in biotic stresses acting in plant immunity (Adie *et al.*, 2007; Fan *et al.*, 2009; Robert-Seilaniantz *et al.*, 2011; Ramegowda & Senthil-Kumar, 2015).

1.5.1 ABA biosynthesis, catabolism, conjugation and transport

The increase of ABA levels in the leaves and roots after drought stress was very limited in the ABA-deficient *Arabidopsis* mutant *aao3-1,* which has a defect in a final step of ABA biosynthesis, indicating that the increase in ABA levels after stress treatment is due to the activation of de novo ABA biosynthesis (Ikegami *et al.*, 2009). ABA can also be rapidly release from cellular stores of conjugated glycosyl ester form by glucanases activated or stabilized by dehydrating stress (Lee *et al.*, 2006; Xu *et al.*, 2012)

Similarly the most plant hormones, ABA levels reflect a balance of ABA biosynthesis and inactivation by turnover or conjugation, further modified by compartimentation and transport (Figure 15). In plants, ABA is synthesize from carotenoids and it is known to be transported over long distances (Jiang & Hartung, 2008). As a weak acid, ABA is mostly uncharged when present in the relatively acid

apoplastic compartiment of plants and analyses uptake does not occur solely by a diffusive process since active ABA transporters were also reported to participate to its uptake (Jiang & Joyce, 2003). Among multiple plasma membrane-localization transporters that have been recently identified, two ATP-binding cassete (ABC) transporters were identified as an importer (AtABCG40) and exporter (AtABCG25) of ABA, and genetic analyses demonstrated their importance for ABA responses including stomatal regulation, gene regulation, germination inhibition and stress tolerance (Kang *et al.*, 2010; Kuromori *et al.*, 2010).

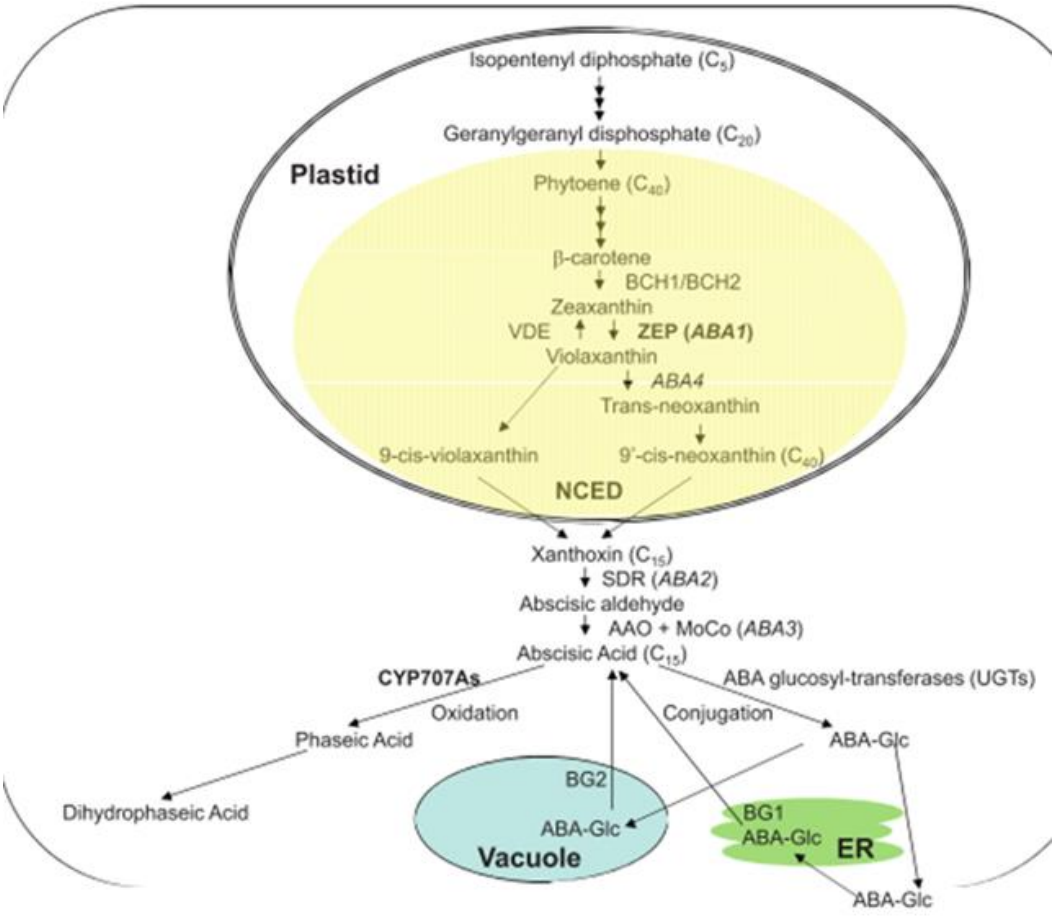


Figure 15 ABA metabolic pathways. ABA biosynthesis, degradation and conjugation pathways are shown in relation to the cellular compartments where these events occur. Carotenoid intermediates are highlighted in yellow. Enzymes regulating key regulatory steps are shown in bold. Individual loci identified based on ABA deficiency are shown in italics.
Source: Finkelstein, 2013

The site of stress perception and that of ABA biosynthesis during the drought stress have been extensively discussed (Sauter *et al.*, 2001; Ikegami *et al.*, 2009; Hartung, 2002; Jeschke *et al.*, 1997). There are evidences that shoot transpiration rate is largely dependent of the delivery of ABA from the roots and the sensitivity to ABA in response to water deficit. In this context, roots are able to 'measure' decreasing soil water availability during a period of drought which results in an increased release of

ABA from the roots tissues to the xylem vessels. After xylem transport to the shoot, guard cells respond rapidly and sensitively to increased ABA concentrations resulting in reduced transpirational water loss (Sauter *et al.*, 2001). Some of the ABA synthesised in the dry roots may be transported to the shoot through the xylem with the transpiration stream and accumulate in high levels in the leaves (Hartung, 2002). Under conditions of soil drying and salt stress large amounts of ABA are deposited in root tissues and loaded into the xylem. Sometimes ABA synthesis by roots is increased substantially but root ABA concentrations may not increase because most of this newly synthesized ABA is loaded to the xylem and transported to the leaves (Jeschke *et al.*, 1997). It could also occur once ABA may move freely from plant to soil and to soil from plant (Sauter *et al.*, 2001).

On the other hand, it have been demonstrated also that ABA is synthesized mainly in the leaves in response to drought stress and that some of the ABA accumulated in the leaves is transported to the roots (Ikegami *et al.*, 2009). In this work, tracer experiments using isotopelabeled ABA indicate that the movement of ABA from leaves to roots is activated by water deficit in roots (Ikegami *et al.*, 2009). When roots were kept in well-watered conditions and drought stress was localized to the leaves only, the ABA level in the leaves increased as in the case of intact plants and detached leaves. Further, under these conditions, the ABA level in the roots did not differ from that in the well-watered control. On the other hand, when drought stress was localized to the roots only, the ABA level in the leaves was slightly higher than that in the well-watered control. Consistent with the ABA levels, leaf stomata closure was almost complete after localized stress treatment to leaves, and was partially induced when drought stress was localized to roots only (Ikegami *et al.*, 2009).

The role of ABA in controlling plant responses likely involves actions at several levels, including effects on transcription, RNA processing, post-translational protein modifications, and the metabolism of secondary messengers (Figure 16). Almost 200 loci regulating ABA response and thousands of genes are regulated by ABA under different contexts (Finkelstein, 2013).

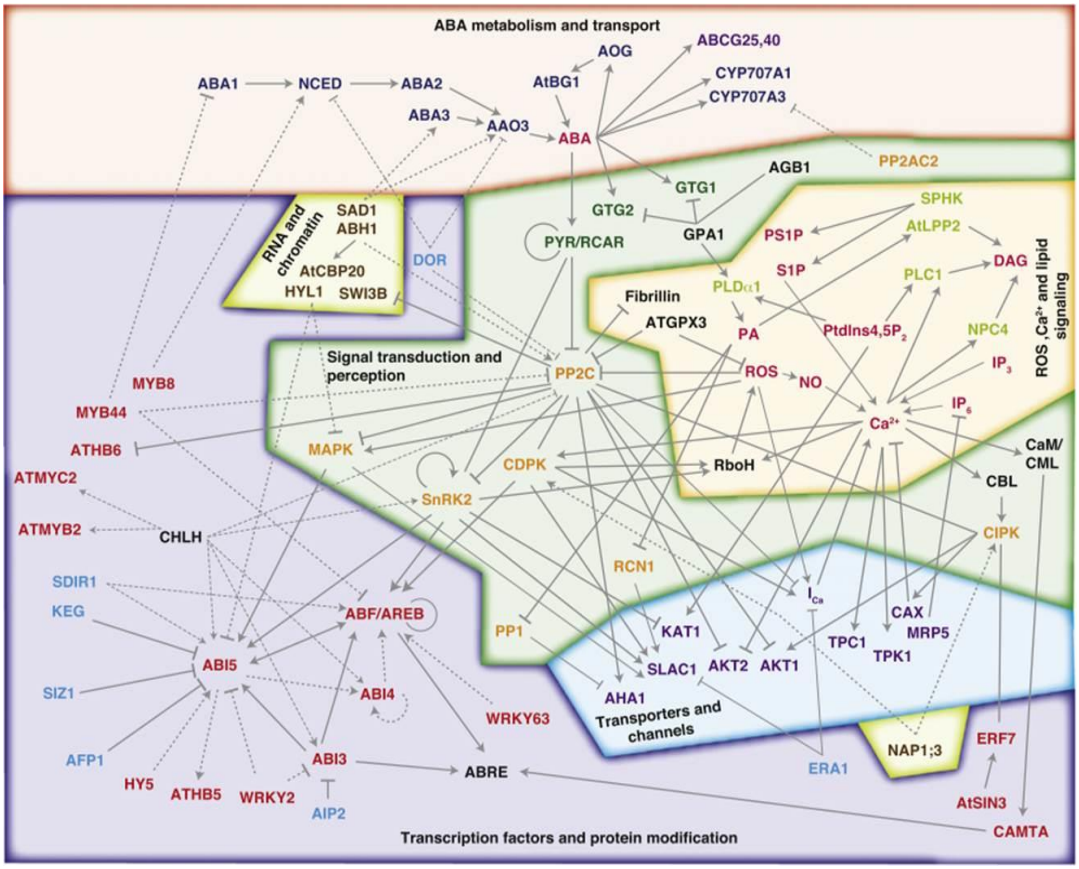


Figure 16 The ABA signaling network. The network is divided into six main functional categories: ABA metabolism and transport (red); perception and signal transduction (dark green); ROS, Ca^{2+} and lipid signaling (orange); transporters and channels (blue); transcription factors and protein modification (purple); and RNA processing and chromatin remodeling (light green).
Source: Hauser *et al.*, 2011.

1.6 The PYL/PP2C/SnRK2: the first steps of ABA sensing and signaling

Over the past few decades, a lot of work was done elucidating the molecular mechanisms underlying ABA sensing and signaling (Umezawa *et al.*, 2010). Several putative ABA receptors, including FCA (Razem *et al.*, 2006), CHLH (Shen *et al.*, 2006), GCR2 (Liu *et al.*, 2007), GTG1 and GTG2 (Pandey *et al.*, 2009) were reported to bind ABA with varying affinities. The discovery of PYLs candidate ABA receptors was different from that of the earlier putative ABA receptors, once independent findings from several groups converged upon this novel class of ABA binding proteins, which fit elegantly into a model that connected the core components of the ABA signal transduction pathway (Ng *et al.*, 2014).

The tripartite ABA signaling pathway is initiated by ABA perception through the PYRABACTIN RESISTANCE1 (PYR1)/PYR1-LIKE (PYL)/REGULATORY COMPONENTS OF ABA RECEPTORS (RCAR) family of proteins (Ma *et al.*, 2009; Park *et al.*, 2009). These novel intracellular ABA receptors (PYL/RCARs) are involved in ABA sensing and signaling via their direct

interaction with clade A protein phosphatase type 2C (PP2Cs), such as ABA INSENSITIVE1 (ABI1) and ABI2, HYPERSENSITIVE TO ABA1 (HAB1) and HAB2, and PROTEIN PHOSPHATASE 2CA/ABA-HYPERSENSITIVE GERMINATION3 (PP2CA/AHG3), thereby releasing their inhibition on three ABA-activated SNF1-related protein kinases (SnRK2s), SnRK2.2/D, 2.3/I and 2.6/E/OST1 (Umezawa *et al.*, 2009; Vlad *et al.*, 2009).

The current ABA signal transduction model can be described as follow: in the absence of ABA, SnRK2 kinases are inactivated by PP2Cs which physically interact with SnRK2 and dephosphorylate a serine residue in the kinase activation loop, a phosphorylation essential for kinase activity (Belin *et al.*, 2006). On the other hand, when ABA binds to the ABA receptors family PYR/PYL/RCAR, this allows the bounds of the receptors in the catalytic site of PP2Cs to inhibit their enzymatic activity. In that case, ABA-induced inhibition of PP2Cs that leads to SnRK2 activation (Boudsocq *et al.*, 2007; Soon *et al.*, 2012; Leung, 2012).

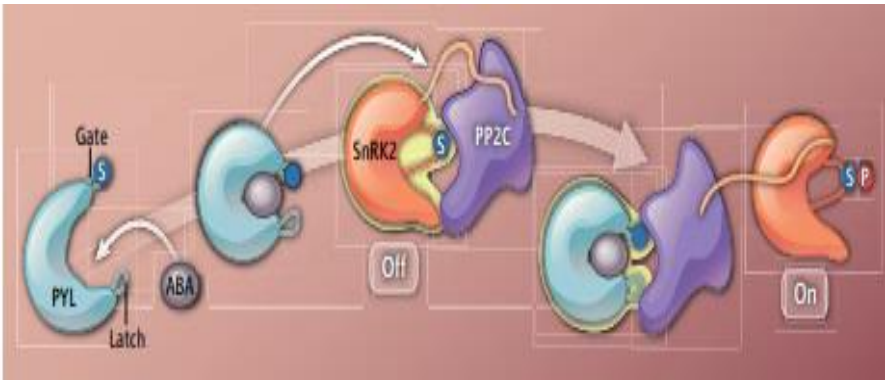


Figure 17 Molecular mimicry between the kinase SnRK2 and the hormone receptor PYL bound to ligand ABA permits alternate binding to the PP2C phosphatase. This change in partners activates (on) or deactivates (off) SnRK2, allowing it to phosphorylate downstream signals.
Source: Leung, 2012.

A crucial event in the receptor's activation was found to be an open-to-closed conformational change in the gate loop of the receptor protein. More recent progress has provided strategies for controlling the gate's closure using chemical agonists (Melcher *et al.*, 2010; Todoroki & Hirai, 2002) or protein engineering approaches. On the other hand, ABA antagonist could be used inhibiting ABA signaling in vivo and further investigations using this approach may reveal the function of ABA in diverse plant species. ABA antagonists may provide new insights into the function of ABA in desiccation tolerance during the evolution of plants on land (Takeuchi *et al.*, 2014).

1.7 Evolution of ABA sensing and signaling

ABA is ubiquitious in plants and it is also produced by some phytopathogenic fungi, bacteria and metazoans ranging from sea esponges to humans (Wasilewska *et al.*, 2008). Based on the available

fossil record, the first land plants (embryophytes) colonized the terrestrial habitat about 500 million to 470 million years ago (Sanderson *et al.*, 2004; Lang *et al.*, 2010). Regarding cellular dehydration in plants, the core ABA signaling components found in *Arabidopsis* are conserved only in land plants (Figure 18), unlike the auxin and ethylene signaling components (Klingler *et al.*, 2010; Umezawa *et al.*, 2010; Hauser *et al.*, 2011), supporting the idea that ABA signaling components may have played a crucial role in land colonization by plants. Furthermore, phylogenetic and transcriptome data suggest that plants have developed a highly sophisticated stress tolerance system through the expansion of duplicate gene families implicated in ABA signaling (Hanada *et al.*, 2011).

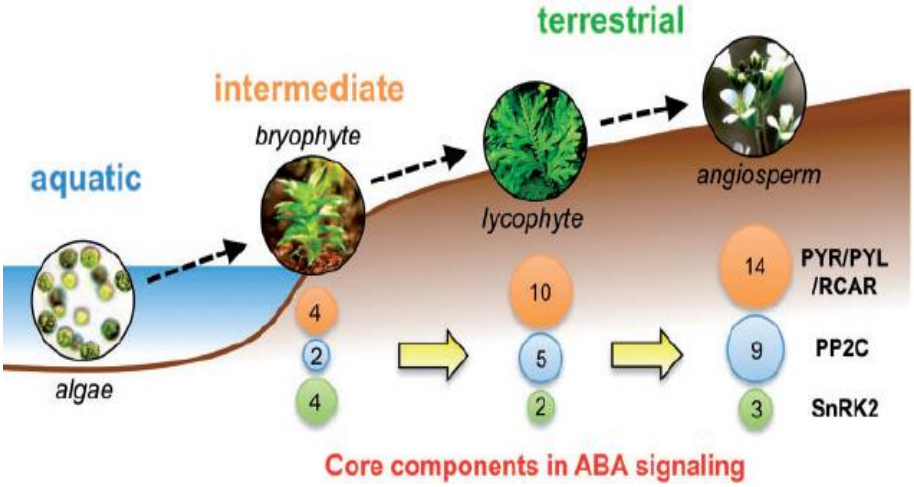


Figure 18 Evolution of core components of ABA signaling. The PYR/PYL/RCAR, group A PP2C and subclass III SnRK2 are conserved from bryophytes. The development of an ABA signaling system seems to be highly correlated with the evolution from aquatic to terrestrial plants. As representatives, component numbers of bryophyte, lycophyte and angiosperm were obtained from *Physcomitrella patens*, *Selaginella moellendorffii* and *Arabidopsis thaliana*, respectively.
Source: Umezawa *et al.*, 2010.

ABA was characterized like an important endogenous small molecule that mediates stress-responsive gene expression, stomatal closure, and vegetative growth modulation (Rodriguez-Gacio *et al.*, 2009) in water deficit conditions. Overall, the core ABA signaling components play an essential role in both fast and slow response to cellular dehydration (Figure 19). To maintain water, ABA promotes stomatal closure through the control of membrane transport systems (Osakabe *et al.*, 2014), shoot growth is inhibited whereas the root growth rate is maintained to gain access to water (Des Marais *et al.*, 2012). Thus, fast ABA signaling involves stomatal closure responses in guard cells, whereas the comparatively slow signaling pathways involve transcriptional regulation in both seeds and vegetative tissues (Miyakawa *et al.*, 2013).

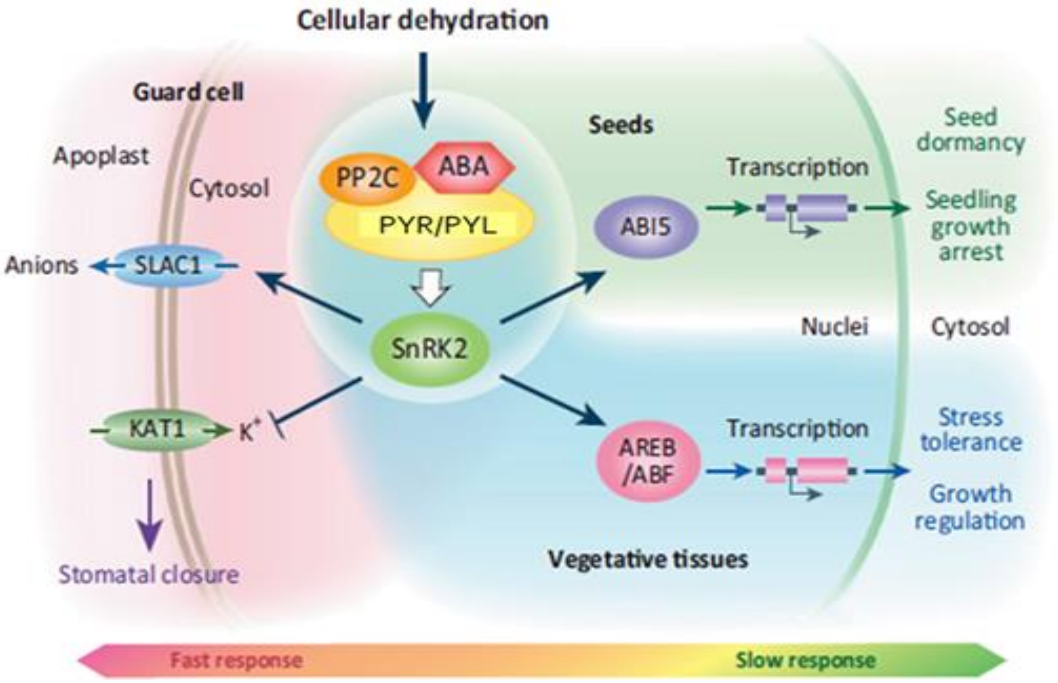


Figure 19 Current model for the major abscisic acid (ABA) signaling pathways in response to cellular dehydration. Core ABA signaling components [ABA, ABA receptors, protein phosphatases 2C (PP2Cs), and subclass III sucrose non-fermenting-1 (SNF1)-related protein kinase 2 (SnRK2s)] control both fast and slow ABA signaling pathways in response to cellular dehydration. Fast signaling involves stomatal closure responses in guard cells, whereas the comparatively slow signaling pathways involve transcriptional regulation in both seeds and vegetative tissues.
Source: Adapted from Miyakawa *et al.*, 2013.

In guard cells, SnRK2 protein kinases activate the anion channel SLOW ANION CHANNEL-ASSOCIATED 1 (SLAC1) and inhibit the cation channel POTASSIUM CHANNEL IN ARABIDOPSISTHALIANA 1 (KAT1) through phosphorylation to release anions, causing stomatal closure (Cutler *et al.*, 2010). In seeds, the post-germination phase induce cellular dehydration (Fujita *et al.*, 2012) which cause an increase in plant ABA content through increase ABA synthesis in vascular tissues, adjustment of ABA metabolism (Nambara *et al.*, 2005), and transport to sites of ABA action (Kanno *et al.*, 2012). In roots, ABA signaling plays an important role to regulate root growth and root system architecture and this system is required for both hydrotropism and osmoregulation of water-stressed roots (Sharp *et al.*, 2004; Gonzalez-Guzman *et al.*, 2014). So, to regulate ABRE-dependent gene expression in seeds and vegetative tissues, respectively subclass III SnRK2s released from inhibition by PP2Cs activate ABA-INSENSITIVE 5 (ABI5) and ABA-responsive element (ABRE) binding protein (AREB)/ABRE-binding factor (ABF) transcription factors (TFs) (Miyakawa *et al.*, 2013).

1.7.1 The tripartite system: PYL-PP2C-SnRK complex

Abscisic acid (ABA) has a central role regulating adaptive responses in plants (Gonzalez-Guzman *et al.*, 2014). Under drought, this phytohormone, synthesized in roots and leaves during periods of water scarcity (Thompson *et al.*, 2007), is perceived by ABA receptors that are the first component of the ABA tripartite systems (Klingler *et al.*, 2010). Further, the PYL-ABA complex bind to the clade A phosphatase type 2C (PP2C) inactivating them (Hao *et al.*, 2011; Ma *et al.*, 2009; Park *et al.*, 2009). Then, the subclass III SNF1-related kinase (SnRK2) proteins are activated by dephosphorylation allowing expression of downstream stress responsive genes (Cutler *et al.*, 2010). In this system, SnRK2 and PP2C proteins function therefore as positive and negative regulators of ABA pathway, respectively.

1.7.2 PYR-PYL/RCARs: ABA receptors

Concerning ABA receptors, PYR/PYL/RCAR proteins are members of the large superfamily of soluble ligand-binding proteins defined as the START-domain superfamily (Iyer *et al.*, 2001), more recently named Bet v I-fold superfamily (Radauer *et al.*, 2008). After the genetic and biochemical identification of PYL/RCARs, several groups have determined the protein structure of the complex between PYL/RCARs and PP2Cs via X-ray crystallography. To date, the crystal structures of PYR1 (Nishimura *et al.*, 2010; Santiago, *et al.*, 2009), PYL1 (Miyazono *et al.*, 2009), PYL2 (Melcher *et al.*, 2009; Yin *et al.*, 2009), PYL3 (Zhang *et al.*, 2013; Zhang *et al.*, 2012), PYL5 (Zhang *et al.*, 2013), PYL9 (Zhang *et al.*, 2013; Nakagawa *et al.*, 2014), PYL10 (Hao *et al.*, 2011; Sun *et al.*, 2012), and PYL13 (Li *et al.*, 2013) have been reported.

Cellular ABA receptor PYL/RCAR orthologs appear to be highly evolutionarily conserved in plants. For example, the *A. thaliana* genome encodes 14 PYR/RCAR proteins, named PYR1 and PYR1-like (PYL) 1-13 or RCAR1-RCAR14 (Ma *et al.*, 2009; Park *et al.*, 2009). The receptor family can be classified into different sub-types based on the sequence similarity, ABA sensitivity, oligomeric state, basal activation level and function. For instance, PYR1/RCAR11, PYL1/RCAR12, PYL2/RCAR14 and PYL3/RCAR13 proteins of Arabidopsis, which form homodimers in the absence of ABA, were released as monomers following ABA binding and subsequently interacted with group-A PP2Cs. In contrast, PYL4/RCAR10, PYL5/RCAR8, PYL6/RCAR9, PYL8/RCAR3, PYL9/RCAR1 and PYL10/RCAR4 behave as monomers in both the presence and absence of ABA, and these monomers can inhibit group-A PP2Cs regardless of ABA binding (Yoshida *et al.*, 2015). There are at least 10 functional orthologs in *Oryza sativa* (Kim *et al.*, 2012), 14 in *Solanum lycopersicum* (Sun *et al.*, 2011; Gonzalez-Guzman *et al.*, 2014), 7 in *Vitis vinifera* (Boneh *et al.*, 2012) and 6 in *Citrus sinensis* (Romero *et al.*, 2012).

A series of mutations in *PYR1/RCAR11* increase its basal activity. Once the combination of these mutations was incorporated into *PYL2* this was sufficient for the activation of ABA signaling in seeds (Mosquna *et al.*, 2011) suggesting that a single receptor modified is sufficient to activate this

signaling. In this sense, a useful tool that activate individual family members selectively and explore phenotypic consequences (Ben-Ari, 2012).

ABA receptors *PYL4* and *PYL5* are known to be involved in the regulation of *ABI1* and *ABI2* genes, ABA normally lowers wild type PP2C activity via PYR/PYL proteins, but ABI PP2Cs escape this and disrupt signaling due to their residual activity (Park *et al.*, 2009). Furthermore, *PYL4* and *PYL5* have been pointed as components of the crosstalk between the JA and ABA signaling pathways (Figure 20) (Lackman *et al.*, 2011). In *N. tabacum* and *A. thaliana*, the *PYL4* gene is regulated by JA. The loss-of-function mutants in *PYL4* and *PYL5*, which were hypersensitive to JA treatment, showed reduced growth in comparison to wild type plants (of *A. thaliana*. Both mutants *pyl4* and *pyl5* displayed reduced anthocyanin accumulation in response to JA compared to wild type (Lackman *et al.*, 2011). Interestingly, *PYL4* and *PYL5* stand out among the genes that were up-regulated at 3 hours after under drought and inoculation by *Pieris rapae* (Davila Olivas *et al.*, 2016) showing that these genes could act in different hormonal pathways intermediating both abiotic and biotic stresses. The *OsPYL/RCAR5* gene stands out as positive regulator of the ABA signal transduction pathway in seed germination and early seedling growth (Kim *et al.*, 2012).

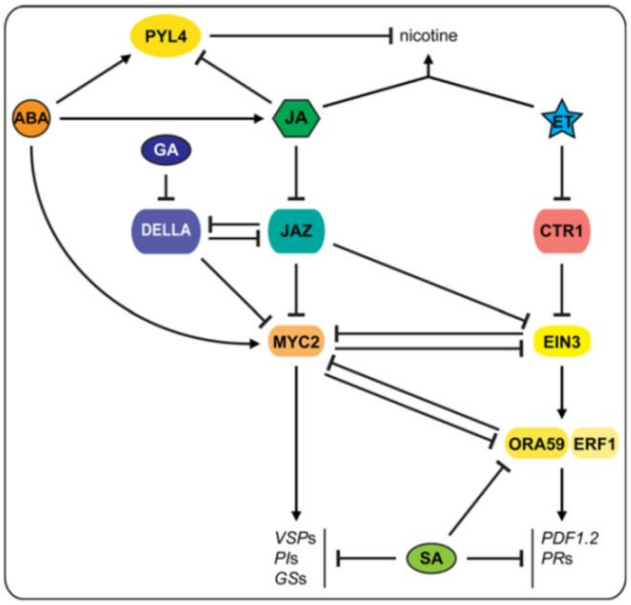


Figure 20 Schematic representation of interactions between hormonal cascades regulating induced defense against biotic agents. Insect herbivores induce JA-dependent MYC2 regulation of defense-related genes, which is enhanced by ABA signaling. Necrotrophic pathogens induce JA/ET-dependent signaling to regulate ERF1 and ORA59 and downstream defense-related genes. The two branches of defense responses mutually antagonize one another. GA and SA signaling generally inhibit JA-dependent defense responses.
Source: Nguyen *et al.*, 2016.

In *Arabidopsis* the overexpression of *PYL9/RCAR1*, *PYL5/RCAR8* and PYL8/RCAR3 genes produced enhanced ABA responses or elevated drought tolerance (Ma *et al.*, 2009; Santiago *et al.*, 2009;

Saavedra *et al.*, 2010). Several recent studies have suggested that the role of PYL8/RCAR3 is overlapping with but distinct from that of other PYR/PYL/RCAR. The PYL8/RCAR3 interacts with transcription factors such as MYB77 which lead to the transcriptional activity of MYB77 which modulates auxin signaling during lateral root development (Shin *et al.*, 2007).

The RCAR7/PYL13 family member regulated the phosphatase activity of the PP2C ABI1, ABI2, and PP2CA proteins *in vitro* at nanomolar ABA levels. However, it appeared to differ from the majority of other RCARs once it failed to bind to the hypersensitive to ABA 1 (HAB1) PP2C in a heterologous system (Bhaskara *et al.*, 2012). Of the 14 RCARs, it has been shown that RCAR7 was the only one that had a variant ABA-binding pocket, with three non-consensus amino acids (Fuchs *et al.*, 2013).

Despite ABA receptor function of RCAR7 has been questioned it was recently demonstrated and the structural constraints that contribute to specific pairing of RCAR7 with PP2Cs was identified (Fuchs *et al.*, 2013).

1.7.3 PP2Cs phosphatases

Otherwise, protein phosphatases are already well known to function as negative regulators of ABA signaling pathway. The physiological functions of PP2Cs were clearly determined genetically in the beginning of XXI century (Umezawa *et al.*, 2010a). Model plants such as *A. thaliana* and rice contained for example 80 and 78 PP2C genes, respectively (Xue *et al.*, 2008). Phylogenetic analyses from *Arabidopsis* and soybean were supported by gene structure and protein motifs and led to subdivide the *PP2C* genes (Figure 21).

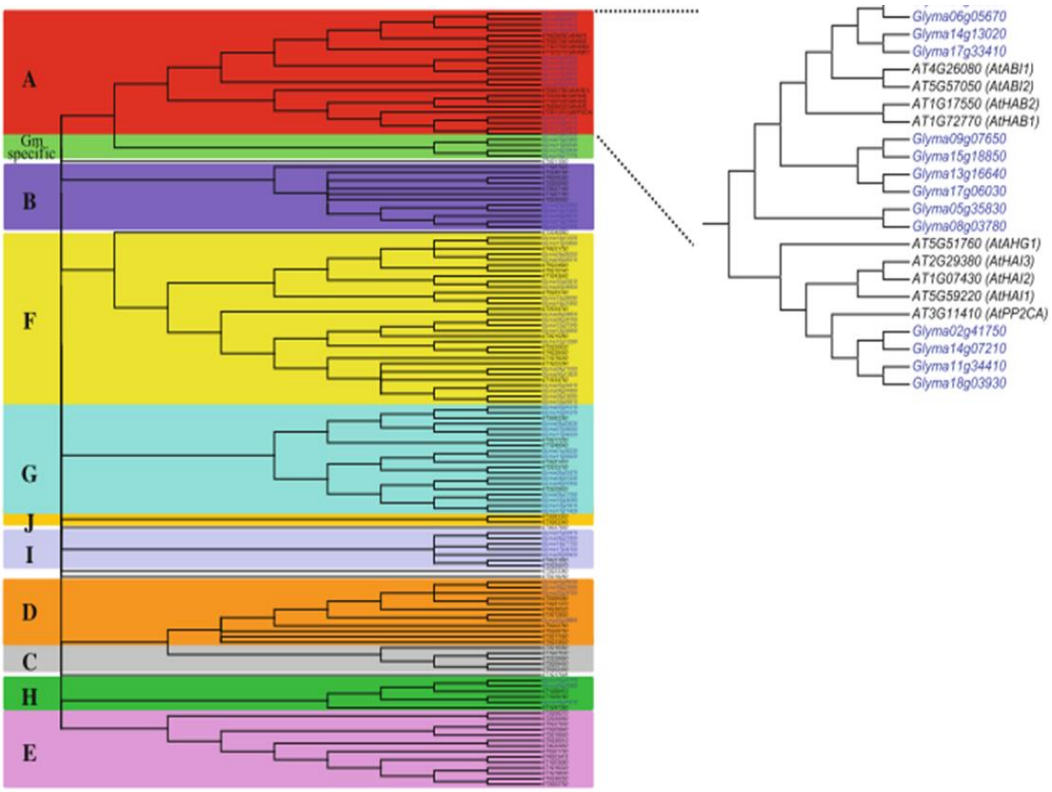


Figure 21 An unrooted phylogenetic tree based on sequence alignment of the catalytic domains encoded by soybean and *Arabidopsis* PP2C. Each cluster was categorized according to the phylogenetic analysis of Arabidopsis PP2C genes (Schweighofer *et al.*, 2004). The cluster of Arabidopsis (black font) and soybean (blue font) group A PP2C is enlarged.
Source: Adapted from Ben-Ari *et al.*, 2012.

PP2C proteins are classified according to the substrate into Ser/Thr, Tyr or dual-specificity classes. Depending on their biochemical and structural features, plant Ser/Thr phosphatases are further divided into PP1, PP2A and PP2C groups (Luan, 2003). The PP2C proteins contain both catalytic and regulatory domains (Figure 22) within the same polypeptide chain (Shi, 2009).

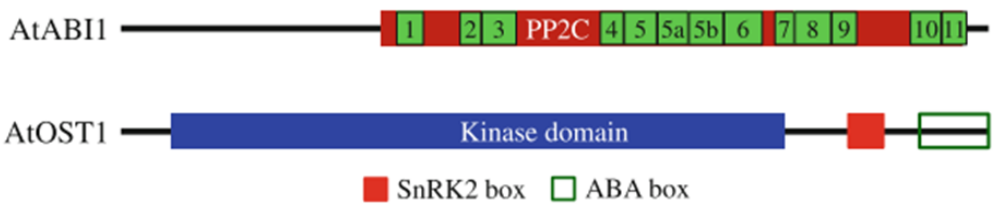


Figure 22 A schematic representation of the group A PP2C, AtABI1 and the SnRK2, AtOST1. AtABI1consit of a PP2C (catalytic) domain (brown) in addition to the 11 motifs (green) (Bork *et al.*, 2006) at its C-terminal. AtOST1 consist of a kinase domain (blue) at its N-terminal followed by a SnRK2 box (red) and an ABA box (green). The ABA box appears with an empty green box to emphasize that this domain is not used for SnRK2 identification.
Source: Adapted from Ben-Ari *et al.*, 2012.

Gene duplication analyses reveals that whole genome and chromosomal segment duplications mainly contributed to the expansion of both *OsPP2C*and *AtPP2C* genes, however, tandem or local duplication occurred less frequently in *Arabidopsis* than rice (Xue *et al.*, 2008).

PP2C phosphatases belong to the Mn^{2+}/Mg^{2+} metal-dependent protein phosphatases PPM family and negative regulatory roles of PP2C subgroup A in ABA signaling have been demonstrate after 2009 and suggesting that PP2C functions are well conserved in different plant species (Saez *et al.*, 2003; Komatsu *et al.*, 2009). Two homologous members of clade B PP2Cs were also reported to be involved in ABA signaling (Ben-Ari, 2012). Regarding, group-A PP2Cs are functionally redundant at the molecular level, but they have distinctive roles in different tissues and organs, as indicated by tissue-specific expression patterns (Umezawa *et al.*, 2010). The PP2C functions emphasized the existence of sophisticated signaling pathways in plants, in which protein dephosphorylation played a crucial role towards determining specificities (Schweighofer *et al.*, 2004).

At least six *A. thaliana* PP2Cs belonging to the group A act as negative regulators of the ABA pathway (Gosti *et al.*, 1999; Merlot *et al.*, 2001; Leonhardt *et al.*, 2003; Saez *et al.*, 2003; Yoshida *et al.*, 2006). On the other hand, ten *VvPP2Cs* and two *CsPP2C* were identified in *V. vinifera* and *C. sinensis* from group A, respectively, while a family of 23 group A-PP2C genes was found in *S. lycopersicum* (Wang *et al.*, 2013a). In *V. vinifera* and *C. sinensis* all these genes were shown to be up-regulate in response to drought (Gambetta *et al.*, 2010; Boneh *et al.*, 2012a). Interestingly, the expression pattern of the OsPP2C subfamily A genes plants treated with ABA, salt, osmotic (mannitol) and cold stress is in good agreement with the microarray data for Arabidopsis subfamily A members, suggesting that the members of this subfamily play foremost roles in ABA-mediated processes related to stress responses both in monocots and eudicots (Xue *et al.*, 2008).

1.8 SnRK2 kinases

The reversible phosphorylation of proteins is a fundamental mechanism by which living organisms modulate signal transduction events (Cutler *et al.*, 2010). Once active, SnRK2 kinases can phosphorylate downstream effectors (Figure 23) such as the basic leucine zipper transcription factors ABFs/AREBs, thus switching-on the transcription of ABA-responsive genes (Furihata *et al.*, 2006).

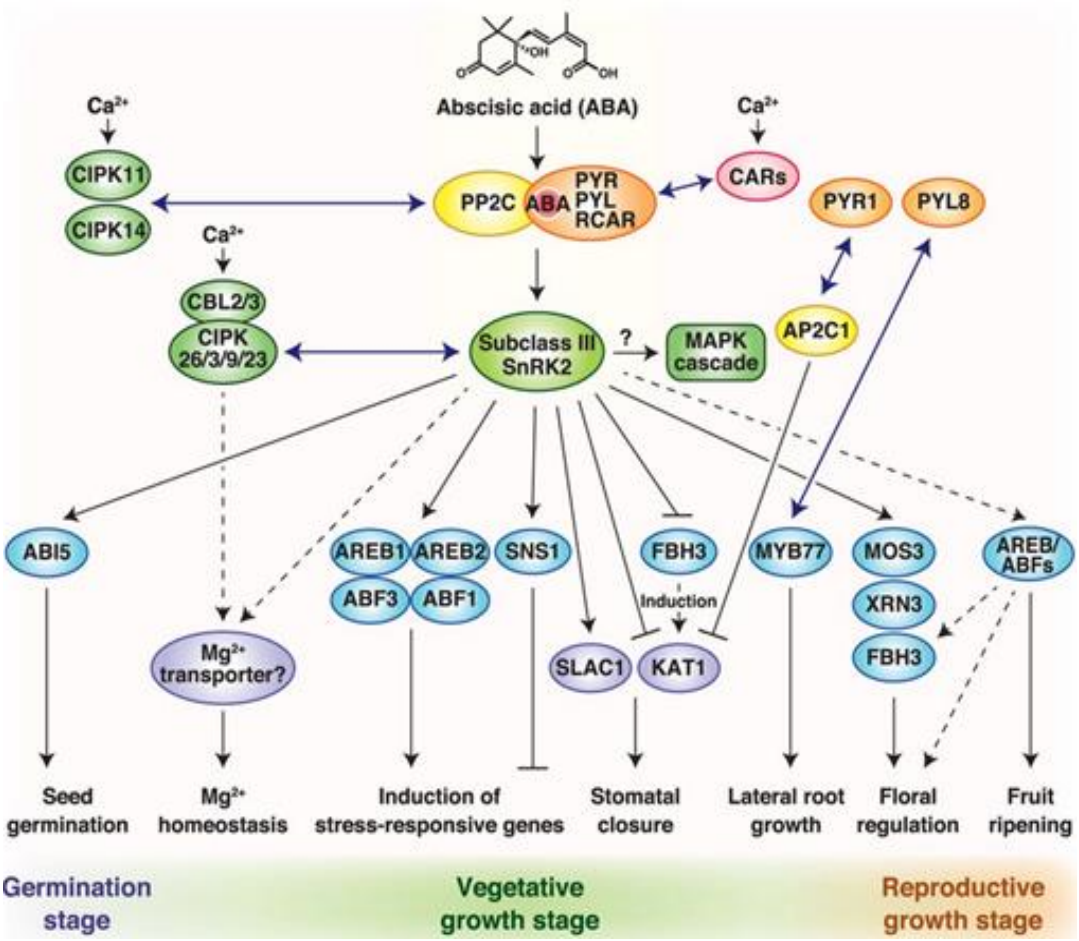


Figure 23 Schematic model of the ABA signaling pathway, which is mediated by novel signaling components discovered in recent omics studies as well as by the core components PYR/PYL/RCAR, group-A PP2Cs and subclass III SnRK2. In addition to the core components, several protein kinases/protein phosphatases (green and yellow ellipses, respectively) are key players in the regulation of ABA-mediated physiological responses during the life cycle of plants. Several PYR/PYL/RCAR proteins (represented by orange ellipses) are also able to regulate ABA responses independent of group A PP2Cs. C2-domain ABA–related (CAR) proteins are shown as pink ellipses. Downstream targets involved in transcriptional regulation and ion transport are shown as blue and purple ellipses, respectively. Physical interactions identified by interactome analyses are depicted as bidirectional blue arrows. The dashed lines indicate possible but unconfirmed routes. Due to space constraints, not all interacting protein and /or substrates of the core components are shown.
Source: Adapted from Yoshida *et al.*, 2015.

The first positive regulators termed SnRK2 (Subfamily 2 of sucrose non-fermenting 1 related protein kinases SNF1) gene was isolated and characterized 20 years ago in wheat and called PKABA1 (Anderberg & Walker-Simmons, 1992). At least 10 SnRK2-encoding genes were found in *A. thaliana* genome, with *SnRK2.2, SnRK2.3* and *SnRK2.6* being associated with ABA signaling (Fujii and Zhu, 2009). The entire *SnRK2* gene family was also identified in many crops such *O. sativa* (Kobayashi *et al.*, 2004), *S. lycopersicum* (Sun *et al.*, 2011; Sato *et al.*, 2012; Wang *et al.*, 2013), *V. vinifera* (Boneh *et al.*, 2012) and *C. sinensis* (Romero *et al.*, 2012).

Among SnRK superfamily proteins, SnRK2s plays a major part in ABA signalling and it were divided into three subclasses (Figure 24), which differed by their activation in response to ABA (Kobayashi *et al.*, 2004; Boudsocq *et al.*, 2004). Subclass I corresponded to genes not activated in the presence of ABA. On the other hand, SnRK2s proteins of subclass II were activated to a lesser extent by ABA. In turn, those of subclass III are strongly activated by ABA.

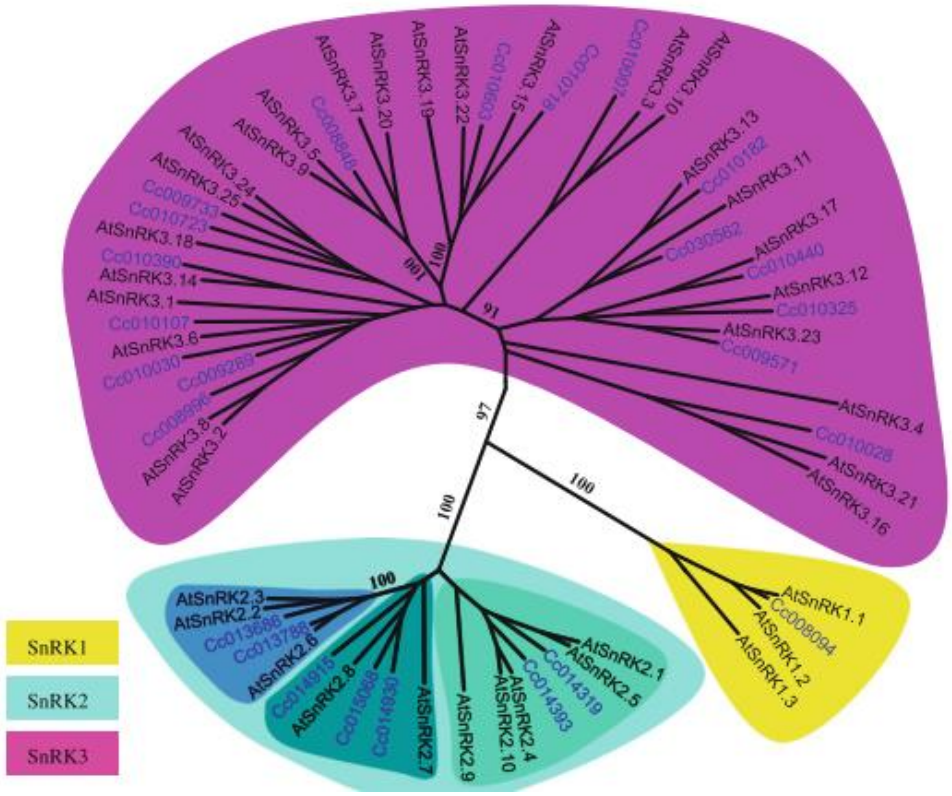


Figure 24 All SnRKs from *Arabidopsis* (black font) and Clementine (blue font) are presented with yellow (SnRK1), blue (SnRK2) and purple (SnRK3) backgrounds. The SnRK2s were clustered into three subgroups, each of which appears with a different background color.
Source: Ben-Ari, 2012.

The C-terminal extremity of SnRK2 subclass III contain an Asp-enriched domain required for both the hormone specific activation of the kinase (Belin *et al.*, 2006) and interaction with PP2C (Hubbard *et al.*, 2010). Domain I represent the SnRK2 box, which is conserved in all members of the SnRK2 gene family. The kinase domain presents an ATP-binding and the activation loop. Domain II is ABA box is conserved only in subclass III of the SnRK2 gene family.

References

ADIE, B. A. T. *et al.* ABA is an essential signal for plant resistance to pathogens affecting JA biosynthesis and the activation of defenses in *Arabidopsis*. **The Plant Cell**, Cambridge, v. 19, p. 1665-1681, 2007

ALTER, S. *et al.* DroughtDB: an expert-curated compilation of plant drought stress genes and their homologs in nine species. **Database**, Oxford, v. 15, p. bav046, 2015.

ANDERBERG, R. J.; WALKER-SIMMONS, M. K. Isolation of a wheat cDNA clone for an Abscisic acid-inducible transcript with homology to protein kinases. **Proceedings of the National Academy of Sciences**, Washington, v. 89, p. 10183-10187, 1992.

ANTHONY, F. *et al.* The origin of cultivated *Coffea arabica* L. varieties revealed by AFLP and SSR markers. **Theoretical and Applied Genetics**, Berlin, v. 104, p. 894-900, 2002.

ASSAD, E. D. *et al.* Impacto das mudanças climáticas no zoneamento agroclimático do café no Brasil. **Pesquisa Agropecuária Brasileira**, Brasília, DF, v. 39, p. 1057-1064, 2004.

ASSIS, G. A. *et al.* Leaf miner incidence in coffee plants under different drip irrigation regimes and planting densities. **Pesquisa Agropecuária Brasileira**, Brasília, DF, v. 47, n. 2, p. 157-162, 2012.

BEINING, A. Ecophysiological diversity of wild Coffea arabica populations in Ethiopia**:** drought adaptation mechanisms. PhD Thesis Bonn University, 143 p., 2007.

BELIN, C. *et al.* Identification of features regulating OST1 kinase activity and OST1 function in guard cells. **Plant Physiology**, Bethesda, v. 141, p. 1316-1327, 2006.

BEN-ARI, G. The ABA signal transduction mechanism in commercial crops: learning from *Arabidopsis.* **Plant Cell Reports**, Berlin, v. 31, n. 8, p. 1357-1369, 2012.

BERTRAND, B. *et al.* Impact of the *Coffea canephora* gene introgression on beverage quality of *C. arabica*. **Theoretical and Applied Genetics**, Berlin, v. 107, p. 387-394, 2003.

BHASKARA, G. B.; NGUYEN, T. T.; VERSLUES, P. E. Unique drought resistance functions of the highly ABA-induced clade A protein phosphatase 2Cs. **Plant Physiology**, Bethesda, v. 160, n. 1, p. 379-395, 2012.

BHYAN, S. B. *et al.* Cold acclimation in the moss *Physcomitrella patens* involves abscisic acid-dependent signaling. **Journal of Plant Physiology**, Bethesda, v. 169, p. 137-145, 2012.

BONEH, U. *et al.* Characterization of potential ABA receptors in *Vitis vinifera*. **Plant Cell Reports**, Berlin, v. 31, n. 2, p. 311-321, 2012.

BORK, P. *et al.* The protein phosphatase 2C (PP2C) superfamily: detection of bacterial homologues. **Protein Science**, New York, v. 5, p. 1421-1425, 1996.

BOUDSOCQ, M.; BARBIER-BRYGOO, H.; LAURIERE, C. Identification of nine sucrose nonfermenting 1-related protein kinases 2 activated by hyperosmotic and saline stresses in Arabidopsis thaliana. **Journal of Biological Chemistry**, Bethesda, v. 279, p. 41758-41766, 2004.

BOUDSOCQ, M. *et al*. Different phosphorylation mechanisms are involved in the activation of sucrose non-fermenting 1 related protein kinases 2 by osmotic stresses and abscisic acid. **Plant Molecular Biology**, Dordrecht, v. 63, n. 4, p. 491-503, 2007.

BOYER, J. Étude expérimentale des effets du régime d'humidité du sol sur la croissance végétative, la floraison et la fructification des caféiers Robusta. **Café Cacao Thé**, Paris, v. 13, p. 187-200, 1969.

CAMARGO, A. P.; SANTINATO, R.; CORTEZ, J. G. Aptidão climática para qualidade da bebida nas principais regiões cafeeiras de Arábica no Brasil. In: BRAZILIAN CONGRESS OF COFFEE RESEARCH, 18., 1992, Araxá. **Proceedings** Araxá, 1992. p. 70-74.

CAMPDELACREU, J. Parkinson disease and Alzheimer disease: environmental risk factors. **Neurologia**, Barcelona, v. 29, n. 9, p. 541-549, 2014.

CARLSSON, S. Coffee consumption and risk of type 2 diabetes in finnish twins. **International Journal of Epidemiology**, London, v. 33, n. 3, p. 616-617, 2004.

CARNEIRO, F. de A. *et al*. Estrutura e diversidade genética de uma população de *Coffea canephora* conilon utilizando marcadores SNPs identificados por nextRAD. In: SIMPÓSIO DE PESQUISA DOS CAFÉS DO BRASIL, 9., 2015, Curitiba. **Resumos Expandidos** Curitiba, 2015. Available from: <http://www.sbicafe.ufv.br/handle/123456789/4136>. Access in: 10 Mar. 2016.

CHEN, C. *et al.* Light-regulated stomatal aperture in *Arabidopsis* **Mol. Plant**, Berlin, v. 5, p. 566-572, 2012.

CLARINDO, W. R.; CARVALHO, C. R. Comparison of the *Coffea canephora* and *C. arabica* karyotype based on chromosomal DNA content. **Plant Cell Reports**, Berlin, v. 28, p. 73-81, 2009.

COMBES, M. C. *et al.* Regulatory divergence between parental alleles determines gene expression patterns in hybrids. **Genome Biology and Evolution**, Oxford, v. 7, n. 4, p. 1110-1121, 2015.

COMPANHIA NACIONAL DE ABASTECIMENTO. **3º levantamento:** safra. Brasília, DF, 2015. 104 p.

CORNFORTH, J. W. *et al.* Chemistry and Physiology of 'Dormins' in sycamore: identity of sycamore 'Dormin' with Abscisin II. **Nature**, London, v. 205, p. 1269-1270, 1965.

CUTLER, S. R. *et al.* Abscisic acid: emergence of a core signaling network. **Annual Review of Plant Biology**, Palo Alto, v. 61, p. 651-679, 2010.

DAMATTA, F. M. *et al.* Drought tolerance of two field-grown clones of *Coffea canephora*. **Plant Science**, Shannon, v. 164, p. 111-117, 2003.

DAMATTA, F. M.; RAMALHO, J. D. C. Impacts of drought and temperature stress on coffee physiology and production: a review. **Brazilian Journal of Plant Physiology**, Piracicaba, v. 18, p. 55-81, 2006.

DAVILA OLIVAS, N. H. *et al.* Effect of prior drought and pathogen stress on *Arabidopsis* transcriptome changes to caterpillar herbivory. **New Phytologist**, Cambridge, v. 210, p. 1344-1356 2016.

DAVIS, A. P. *et al.* An annotated taxonomic conspectus of the genus *Coffea* (Rubiaceae). **Botanical**

Journal of the Linnean Society, London, v. 152, p. 465-512, 2006.

DAVIS, A. P. *et al.* The impact of climate change on indigenous Arabica coffee (*Coffea arabica*): predicting future trends and identifying priorities. **PLoS ONE**, San Francisco, v. 7, p. e47981, 2012.

DENOEUD, F. *et al.* The coffee genome provides insight into the convergent evolution of caffeine biosynthesis. **Science**, New York, v. 345, n. 6201, p. 1181-1184, Sept. 2014.

DES MARAIS, D. L. *et al.* Physiological genomics of response to soil drying in diverse *Arabidopsis* accessions. **The Plant Cell**, Rockville, v.24, p. 893-914, 2012.

DUBLIN, P. Le rapport longuer pivot/longuer hypocotyle des plantules de *Coffee racemosa Lour*. et de quelques autres espècies du genre. **Café Cacao Thé**, Paris, v. 12, p. 127-134, 1968.

FAN, J. *et al.* Abscisic acid has a key role in modulating diverse plant-pathogen interactions. **Plant Physiology**, Bethesda, v. 150, p. 1750-1761, 2009.

FAZUOLI, L. C. Resistência de *Coffea racemosa* ao *Meloidogyne exigua*. **Bragantia**, Campinas, v. 36, n. 29, p. 297-307, 1975.

FERRÃO, R. *et al.* EMCAPA 8141 - Robustão Capixaba, variedade clonal de café Conilon tolerante à seca, desenvolvida para o estado do Espírito Santo. **Ceres**, Viçosa, MG, v. 47, p. 555-560, 2000.

FINKELSTEIN, R. Abscisic acid synthesis and response. **The Arabidopsis Book**, Bethesda, v. 11, p. e0166, 2013.

FONSECA, A. F. A. *et al.* Conilon Vitória-Incaper 8142: improved *Coffea canephora* var. Kouillou clone cultivar for the State Espírito Santo. **Crop Breeding and Applied Biotechnology**, Londrina, v. 4, p.1-3, 2004.

FUCHS, S. *et al.* Type 2C protein phosphatases in plants. **FEBS Journal**, Amsterdam, v. 280, n. 2, p. 681-693, 2013.

FUJII, H.; ZHU, J. K. Arabidopsis mutant deficient in 3 abscisic acid-activated protein kinases reveals critical roles in growth, reproduction, and stress. **Proceedings of the National Academy of Sciences USA**, Washington, v. 106, p. 8380-8385, 2009.

FUJITA, Y.; YOSHIDA, T.; YAMAGUCHI-SHINOZAKI, K. Pivotal role of the AREB/ABF-SnRK2 pathway in ABRE-mediated transcription in response to osmotic stress in plants. **Physiology Plantarum**, Copenhagen, v. 147, p. 15-27, 2013.

FURIHATA, T. *et al.* Abscisic acid-dependent multisite phosphorylation regulates the activity of a transcription activator AREB1. **Proceedings of the National Academy of Sciences USA**, Washington, v. 103, p. 1988-1993, 2006.

GALLUS, S. Does coffee protect against hepatocellular carcinoma? **British Journal of Cancer**, London, v. 87, n. 9, p. 956-959, 2002.

GAMBETTA, G. A. *et al.* Sugar and abscisic acid signaling orthologs are activated at the onset of ripening in grape. **Planta**, Berlin, v. 232, p. 219-234, 2010.

GELDER, B. van *et al.* Coffee consumption is inversely associated with cognitive decline in elderly European men: the FINE Study. **European Journal of Clinical Nutrition**, Cleveland, v. 61, p. 226-232, 2007.

GLENN, D. M. *et al.* Response of perennial horticultural crops to climate change. **Horticultural Reviews**, New York, v. 47, p. 130, 2014.

GONZÁLEZ-GUZMÁN, M. *et al.* Tomato PYR/PYL/RCAR abscisic acid receptors show high expression in root, differential sensitivity to the abscisic acid agonist quinabactin, and the capability to enhance plant drought resistance. **Journal of Experimental Botany**, Oxford, v. 65, n. 15, p. 4451-4464, 2014.

GOSTI, F. *et al.* ABI1 protein phosphatase 2C is a negative regulator of abscisic acid signaling. **Plant Cell**, Cambridge, v. 11, p. 1897-1909, 1999.

HANADA, K. *et al.* Origin and evolution of genes related to ABA metabolism and its signaling pathways. **Journal of Plant Research**, Tokyo, v. 124, n. 4, p. 455-65, 2011.

HAO, Q. *et al.* The molecular basis of ABA-independent inhibition of PP2Cs by a subclass of PYL proteins. **Molecular Cell**, Cambridge, v. 42, n. 5, p. 662-672, 2011.

HARTUNG, W. Abscisic acid in the xylem: where does it come from, where does it go to? **Journal of Experimental Botany**, Oxford, v. 53, p. 27-32, 2002.

HATANAKA, T. *et al.* Transgenic plants of coffee *Coffea canephora f*rom embryogenic callus via *Agrobacterium tumefaciens*-mediated transformation. **Plant Cell Reports**, Berlin, v. 19, p. 106-110, 1999.

HATFIELD, J. L. Climate change: challenges for future crop adjustments. In: TUTEJA, N.; GILL, S. S. (Ed.). **Climate change and plant abiotic stress tolerance**. Weinheim: Wiley-VCH Verlag GmbH, 2013. p. 1-26.

HAUSER, F.; WAADT, R.; SCHROEDER, J. I. Evolution of abscisic acid synthesis and signaling mechanisms. **Current Biology**, London, v. 21, n. 9, p. R346-R355, 2011.

HAZARIKA, R. R.; MISHRA, V. K.; CHATURVEDI, R. **Crop improvement under adverse conditions**. New York: Springer-Verlag, 2013. 394 p.

HIGDON, J. V.; FREI, B. Coffee and health: a review of recent human research. **Critical Reviews in Food Science and Nutrition**, London, v. 46, n. 2, p. 101-123, 2006.

HINCHA, D. K.; HAGEMANN, M. Stabilization of model membranes during drying by compatible solutes involved in the stress tolerance of plants and microorganisms. **The Biochemical Journal**, London, v. 383, p. 277-283, 2004.

HUBBARD, K. E. *et al.* Early abscisic acid signal transduction mechanisms: newly discovered components and newly emerging questions. **Genes & Development**, Cold Spring Harbor, v. 24, n. 16, p. 1695-1708, 2010.

IKEGAMI, K. *et al.* Activation of abscisic acid biosynthesis in the leaves of *Arabidopsis thaliana* in response to water deficit. **Journal of Plant Research**, Tokyo, v. 122, p. 235-243, 2009.

INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE. Part A: global and sectoral aspects. contribution of working group II to the fifth assessment report of the intergovernmental panel on climate change. In: FIELD, C. B. *et al.* (Ed.). **Climate change 2014:** impacts, adaptation, and vulnerability. Cambridge: Cambridge University Press, 2014. p. 11-32.

INTERNATIONAL COFFEE ORGANIZATION. London, 2016. Available from: <http://www.ico.org/>. Access in: 10 Mar. 2016.

IYER, L. M.; KOONIN, E. V.; ARAVIND, L. Adaptations of the helix-grip fold for ligand binding and catalysis in the START domain superfamily. **Proteins**, New York, v. 43, n. 2, p. 134-144, 2001.

JARAMILLO, J. *et al.* Thermal tolerance of the coffee berry borer *hypothenemus hampei*: predictions of climate change impact on a tropical insect pest. **PLoS ONE**, San Francisco, v. 4, p. e6487, 2009.

JESCHKE, W. D. *et al.* Effects of P deficiency on assimilation and transport of nitrate and phosphate in intact plants of castor bean (*Ricinus communis* L.). **Journal of Experimental Botany**, Oxford, v. 48, p. 75-91, 1997.

JIANG, F.; HARTUNG, W. Long-distance signaling of abscisic acid (ABA): The factors regulating the intensity of the ABA signal. **Journal of Experimental Botany**, Oxford, v. 59, p. 37-43, 2008.

JIANG, Y.; JOYCE, D. C. ABA effects on ethylene production, PAL activity, anthocyanin and phenolic contents of strawberry fruit. **Plant Growth Regulation**, Dordrecht, v. 39, p. 171-174, 2003.

KANG, J. *et al.* PDR-type ABC transporter mediates cellular uptake of the phytohormone abscisic acid. **Proceedings of the National Academy of Sciences USA**, Washington, v. 107, n. 5, p. 2355-2360, 2010.

KANNO, Y. *ET AL.* Identification of an abscisic acid transporter by functional screening using the receptor complex as a sensor. **Proceedings of the National Academy of Sciences USA**, Washington, v. 109, n. 24, p. 9653-9658, 2012.

KERMODE, A. R. Role of abscisic acid in seed dormancy. **Journal of Plant Growth Regulation**, v. 24, n. 4, p. 319-344, 2005.

KLINGLER, T. *et al.* Molecular basis of the core regulatory network in ABA responses: sensing, signaling and transport. **Plant and Cell Physiology**, Oxford, v. 51, n. 11, p. 1821-1839, 2010.

KOBAYASHI, Y. *et al.* Differential activation of the rice sucrose nonfermenting1-related protein kinase2 family by hyperosmotic stress and Abscisic acid. **The Plant Cell**, Rockville, v. 16, p. 1163-1177, 2004.

KOMATSU, K. *et al.* Functional analyses of the ABI1-related protein phosphatase type 2C reveal evolutionarily conserved regulation of abscisic acid signaling between *Arabidopsis* and the moss *Physcomitrella patens*. **Plant Molecular Biology**, Dordrecht, v. 270, p. 341-357, 2009.

KRUG, C. A. **Ensayo mundial del cafe**. Rome: Agricultura de las Americas, 1965. 162 p.

KUROMORI, T. *et al.* ABC transporter AtABCG25 is involved in abscisic acid transport and responses. **Proceedings of the National Academy of Sciences USA**, Washington, v. 107, n. 5, p. 2361-2366, 2010.

LACKMAN, P. *et al.* Jasmonate signaling involves the abscisic acid receptor PYL4 to regulate metabolic reprogramming in Arabidopsis and tobacco. **Proceedings of the National Academy of**

Sciences, Washington, v. 108, n. 14, p. 5891-5896, 2011.

LAMBOT, C. *et al.* Evaluation of conilons for genetic diversity, cup quality and biochemical composition. In: INTERNATIONAL SCIENTIFIC COLLOQUIUM ON COFFEE, 22., 2008, Campinas. **Proceedings** Campinas, 2008. 1 CD-ROM.

LANG, D. *et al.* Genome-wide phylogenetic comparative analysis of plant transcriptional regulation: a timeline of loss, gain, expansion, and correlation with complexity. **Genome Biology and Evolution**, Oxford, v. 19, n. 2, p. 488-503, 2010.

LEE, K. H. *et al.* Activation of glucosidase via stress-induced polymerization rapidly increases active pools of abscisic acid. **Cell**, Cambridge, v. 126, p. 1109-1120, 2006.

LEONHARDT, N. *et al.* Microarray expression analyses of Arabidopsis guard cells and isolation of a recessive abscisic acid hypersensitive protein phosphatase 2C mutant. **Plant Cell**, Cambridge, v. 16**,** p. 596-615, 2004.

LEUNG, J. *et al.* Mécanisme moléculaire d'action de l'acide abscissique en réponse à la sécheresse chez les végétaux: the ABC of abscisic acid action in plant drought stress responses. **Biologie Aujourd'hui**, Les Ulis, v. 206, n. 4, p. 301-312, 2012.

LI, W. *et al.* Molecular basis for the selective and ABA-independent inhibition of PP2CA by PYL13. **Cell Research**, Shanghai, v. 23, p. 1369-1379, 2013.

LIMA, E. A. *et al.* The multi-resistant reaction of drought-tolerant Coffee 'Conilon Clone 14' to *Meloidogyne spp*. and late hypersensitive-like response in *Coffea canephora*. **Phytopathology**, Saint Paul, v. 105, n. 6, p. 805-814, 2015.

LINDSAY, J. Risk factors for Alzheimer's disease: a prospective analysis from the Canadian Study of Health and Aging. **American Journal of Epidemiology**, Oxford, v. 156, n. 5, p. 445-453, 2002.

LIU, X. *et al.* A G protein-coupled receptor is a plasma membrane receptor for the plant hormone abscisic acid. **Science**, New York, v. 315, p. 1712-1716, 2007.

LUAN, S. Protein phosphatases in plants. **Annual Review Plant Biology**, Palo Alto, v. 54, p. 63-92, 2003.

MA, Y. *et al.* Regulators of PP2C phosphatase activity function as abscisic acid sensors. **Science**, New York, v. 324, p. 1064-1068, 2009.

MARRACCINI, P. *et al.* Differentially expressed genes and proteins upon drought acclimation in tolerant and sensitive genotypes of *Coffea canephora*. **Journal of Experimental Botany**, Oxford, v. 63, p. 4191-4212, 2012.

MARRACCINI, P. *et al. RBCS1* expression in coffee: *coffea* orthologs, *Coffea arabica* homeologs, and expression variability between genotypes and under drought stress. **BMC Plant Biology**, London, v. 11, n. 1 p. 1471-2229, 2011.

MAZZAFERA, P. Estresse hídrico e consequências na composição das sementes de café e qualidade da bebida. In: SALVA, T. de J. G. (Ed.). **Cafés de qualidade:** aspectos tecnológicos, científicos e comerciais. Campinas: Instituto Agronômico de Campinas, 2007. p. 73-90.

MEDINA FILHO, H. P.; CARVALHO, A.; MEDINA, D. M. Germoplasma de *C. racemosa* e seu potencial no melhoramento do cafeeiro. **Bragantia**, Campinas, v. 36, p. 43-46, 1977.

MELCHER, K. *et al.* A gate-latch-lock mechanism for hormone signaling by abscisic acid receptors. **Nature**, London, v. 462, p. 602-608, 2009.

MERLOT, S. *et al.* The ABI1 and ABI2 protein phosphatases 2C act in a negative feedback regulatory loop of the abscisic acid signalling pathway. **Plant Journal**, Shannon, v. 25, p. 295-303, 2001.

MITTLER, R. Oxidative stress, antioxidants and stress tolerance. **Trends in Plant Science**, Oxford, v. 7, p. 405-410, 2002.

MIYAKAWA, T. *et al.* Structure and function of abscisic acid receptors. **Trends in Plant Science**, Oxford, v. 18, n. 5, p. 259-266, 2013.

MIYAZONO, K, I. *et al.* Structural basis of Abscisic acid signaling. **Nature**, London, v. 462, p. 609-614, 2009.

MOFATTO, L. S. *et al.* Identification of candidate genes for drought tolerance in coffee by high-throughput sequencing in the shoot apex of different *Coffea arabica* cultivars. **BMC Plant Biology**, London, v. 16, p. 94, 2016.

MONTAGNON, C.; CUBRY, P.; LEROY, T. Amélioration génétique du caféier *Coffea canephora* Pierre: connaissances acquises, stratégies et perspectives. **Cahiers Agricultures**, John Libbey, v. 21, n. 2/3, p. 143-153, 2012.

MONTAGNON, C.; LEROY, T. Réaction à la sécheresse de jeunes caféiers *Coffea canephora* de Côte-d'Ivoire appartenant à différents groupes génétiques. **Café Cacao Thé**, Paris, v. 37, p. 179-190, 1993.

MORICE, C. P. *et al.* Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: the HadCRUT4 data set. **Journal of Geophysical Research:** Atmospheres, Washington, v. 117, n. D8, 2012.

MOSQUNA, A. *et al.* Potent and selective activation of abscisic acid receptors in vivo by mutational stabilization of their agonist-bound conformation. **Proceedings of the National Academy of Sciences USA**, Washington, v. 108, n. 51, p. 20838-20843, 2011.

NAKAGAWA, M. *et al.* Mechanism of high-affinity abscisic acid binding to PYL9/RCAR1. **Genes to Cells**, New York, v. 19, n. 5, p. 386-404, 2014.

NAMBARA, E.; MARION-POLL, A. Abscisic acid biosynthesis and catabolism. **Annual Review of Plant Biology**, Palo Alto, v. 56, p. 165-185, 2005.

NG, L. M. *et al.* Abscisic acid perception and signaling: structural mechanisms and applications. **Acta Pharmacologica Sinica**, Amsterdam, v. 35, n. 5, p. 567-584, 2014.

NGUYEN, D. *et al.* How plants handle multiple stresses: hormonal interactions underlying responses to abiotic stress and insect herbivory. **Plant Molecular Biology**, Dordrecht, v. 91, n. 6, p. 1-14, 2016.

NISHIMURA, N. *et al.* PYR/PYL/RCAR family members are major in-vivo ABI1 protein phosphatase 2C-interacting proteins in *Arabidopsis*. **The Plant Journal:** for Cell and Molecular Biology, Oxford, v. 61, n. 2, p. 290-299, 2010.

OBER, E. S.; SHARP, R. E. Proline accumulation in maize (*Zea mays* L.) primary roots at low water potentials: I., requirement for increased levels of abscisic acid. **Plant Physiology**, Bethesda, v. 105, p. 981-987, 1994.

OHKUMA, K. *et al.* Abscisin II, an abscission-accelerating substance from young cotton fruit. **Science**, New York, v. 142, p. 1592-1593, 1963.

O'KEEFE, J. H. Effects of habitual coffee consumption on cardiometabolic disease, cardiovascular health, and all-cause mortality. **Journal of the American College of Cardiology**, New York,v. 62, n. 12, p. 1043-1051, 2013.

OSAKABE, Y. *et al.* ABA control of plant macroelement membrane transport systems in response to water deficit and high salinity. **New Phytologist**, Cambridge, v. 202, p. 35-49, 2014.

PANDEY, S.; NELSON, D. C.; ASSMANN, S. M. Two novel GPCR-Type G proteins are abscisic acid receptors in Arabidopsis. **Cell**, Cambridge, v. 136, p. 136-148, 2009.

PARK, S. Y. *et al.* Abscisic acid inhibits PP2Cs via the PYR/PYL family of ABA-binding START proteins. **Science**, New York, v. 324, p. 1068-1071, 2009.

PINHEIRO, H. A. *et al.* Drought tolerance is associated with rooting depth and stomatal control of water use in clones of *Coffea canephora*. **Annals of Botany**, London, v. 96, p. 101-108, 2005.

PINHEIRO, H. A.; FONTES, E. P. B.; LOUREIRO, M. E. Drought tolerance in relation to protection against oxidative stress in clones of *Coffea canephora* subjected to long-term drought. **Plant Science**, Shannon, v. 167, n. 6, p. 1307-1314, 2004.

PINTO, H.; ASSAD, E. **Global warming and the new geography of agricultural production in Brazil**. São Paulo: Ed. FAPESP, 2008. 23 p.

PINTO, H. S. *et al.* O aquecimento global e a cafeicultura brasileira. **Boletim da Sociedade Brasileira de Meteorologia**, Rio de Janeiro, v. 31, p. 65-72, 2007.

PONCET, V. *et al.* Development of genomic microsatellite markers in *Coffea canephora* and their transferability to other coffee species. **Genome**, Ottawa, v. 50, p. 1156-1161, 2007.

PONS, R.; CORNEJO, M. J.; SANZ, A. Is ABA involved in tolerance responses to salinity by affecting cytoplasm ion homeostasis in rice cell lines? **Plant Physiology and Biochemistry**, New York, v. 62, p. 88-94, 2013.

PRAXEDES, S. C. *et al.* Effects of long-term soil drought on photosynthesis and carbohydrate metabolism in mature robusta coffee (*Coffea canephora* Pierre var. kouillou) leaves. **Environmental and Experimental Botany**, Oxford, v. 56, n. 3, p. 263-273, 2005.

RADAUER, C.; LACKNER, P.; BREITENEDER, H. The Bet v 1 fold: an ancient, versatile scaffold for binding of large, hydrophobic ligands. **BMC Evolutionary Biology**, London, v. 8, p. 286, 2008.

RAMEGOWDA, V.; SENTHIL-KUMAR, M. The interactive effects of simultaneous biotic and abiotic stresses on plants: mechanistic understanding from drought and pathogen combination. **Journal of Plant Physiology**, Oxford, v. 176, p. 47-54, 2015.

RAMIREZ-VILLEGAS, J. *et al.* A way forward on adaptation to climate change in Colombian agriculture: perspectives towards 2050. **Climatic Change**, Dordrecht, v. 115, n. 3, p. 611-628,. 2012.

RAZEM, F. A. *et al.* The RNA-binding protein FCA is an abscisic acid receptor. **Nature**, London, v. 439, p. 290-294, 2006.

ROBBRECHT, E.; MANEN, J. The major evolutionary lineages of the coffee family (Rubiaceae, angiosperms): combined analysis (nDNA and cpDNA) to infer the position of *Coptosapelta* and *Luculia*, and supertree construction based on *rbcl*, *rps16*, *trnL-trnF* and *atpB-rbcL* data: a new class. **Systematics and Geography of Plants**, Oxford, v. 76, p. 85-145, 2006.

ROBERT-SEILANIANTZ, A.;GRANT, M.;JONES, J. D. Hormone crosstalk in plant disease and defense: more than just jasmonate-salicylate antagonism. **Annual Review Phytopathology**, Palo Alto, v. 49, p. 317-343, 2011.

RODRIGUEZ-GACIO, M. del C.; MATILLA-VÁZQUEZ, M. A.; MATILLA, A. J. Seed dormancy and ABA signaling. **Plant Signaling & Behavior**, Bethesda, v. 4, n. 11, p. 1035-1048, 2009.

ROMERO, P.; LAFUENTE, M. T.; RODRIGO, M. J. The Citrus ABA signalosome: identification and transcriptional regulation during sweet orange fruit ripening and leaf dehydration. **Journal of Experimental Botany**, Oxford, v. 63, n. 13, p. 4931-4945, 2012.

SAAVEDRA, X. *et al.* The Nuclear Interactor PYL8/RCAR3 of *Fagus sylvatica* FsPP2C1 is a positive regulator of abscisic acid signaling in seeds and stress. **Plant Physiology**, Bethesda, v. 152, n. 1, p. 133-150, 2010.

SAEZ, A. *et al.* Gain-of-function and loss-of-function phenotypes of the protein phosphatase 2C HAB1 reveal its role as negative regulator of abscisic acid signaling. **The Plant Journal**, Oxford, v. 37, p. 354-369, 2003.

SANDERSON, M. J. *et al.* Molecular evidence on plant divergence times. **American Journal of Botany**, Saint Louis, v. 91, n. 10, p. 1656-1665, 2004.

SANTIAGO, J. *et al.* The abscisic acid receptor PYR1 in complex with abscisic acid. **Nature**, London, v. 462, p. 665-668, 2009a.

SANTIAGO, J. *et al.* Modulation of drought resistance by the abscisic acid receptor PYL5 through inhibition of clade IIA PP2Cs. **The Plant Journal**, Oxford, v. 60, p. 575-588, 2009b.

SATO, S. *et al.* The tomato genome sequence provides insights into fleshy fruit evolution. **Nature**, London, v. 485, n. 7400, p. 635-641, 2012.

SAUTER, A.; DAVIES, W. J.; HARTUNG, W. The long-distance abscisic acid signal in the droughted plant: the fate of the hormone on its way from root to shoot. **Journal of Experimental Botany**, Oxford, v. 52, p. 1991-1997, 2001.

SCHROEDER, J. I.; KWAK, J. M.; ALLEN, G. J. Guard cell abscisic acid signaling and engineering drought hardiness in plants. **Nature**, London, v. 410, p. 327-330, 2001.

SCHWEIGHOFER, A.; HIRT, H.; MESKIENE, I. Plant PP2C phosphatases: emerging functions in stress signaling. **Trends in Plant Science**, Oxford, v. 9, p. 236-243, 2004.

SHARP, R. E. *et al.* Root growth maintenance during water deficits: physiology to functional genomics. **Journal of Experimental Botany**, Oxford, v. 55, n. 407, p. 2343-2351, 2004.

SHARP, R. E.; LENOBLE, M. E. ABA, ethylene and the control of shoot and root growth under water stress. **Journal of Experimental Botany**, Oxford, v. 53, p. 33-37, 2002.

SHEN, Y. Y. *et al.* The Mg-chelatase H subunit is an abscisic acid receptor. **Nature**, London, v. 443, p. 823-826, 2006.

SHI, Y. Serine/threonine phosphatases: mechanism through structure. **Cell**, Cambridge, v. 139, n. 3, p. 468-484, 2009.

SHIN, R. *et al.* The *Arabidopsis* transcription factor *MYB77* modulates auxin signal transduction. **The Plant Cell**, Rockville, v. 19, p. 2440-2453, 2007.

SHINKAWA, R. *et al.* Abscisic acid induced freezing tolerance in chilling-sensitive suspension cultures and seedlings of rice. **BMC Research Notes**, London, v. 6, p. 351, 2013.

SIVAMANI, E. *et al.* Improved biomass productivity and water use efficiency under water deficit conditions in transgenic wheat constitutively expressing the barley *HVA1* gene. **Plant Science**, Shannon, v. 155, p. 1-9, 2000.

SOON, F. *et al.* Molecular mimicry regulates ABA signaling by SnRK2 kinases and PP2C phosphatases. **Science**, New York, v. 335, n. 6064, p. 85-89, 2012.

SUN, D. *et al.* Crystal structures of the *Arabidopsis thaliana* abscisic acid receptor PYL10 and its complex with abscisic acid. **Biochemical and Biophysical Research Communications**, Orlando, v. 418, n. 1, p. 122-127, 2012.

SUN, L. *et al.* Transcriptional regulation of *SlPYL*, *SlPP2C*, and *SlSnRK2* gene families encoding ABA signal core components during tomato fruit development and drought stress. **Journal of Experimental Botany**, Oxford, v. 62, n. 15, p. 5659-5669, 2011.

TAKEUCHI, J. *et al.* Designed abscisic acid analogs as antagonists of PYL-PP2C receptor interactions. **Nature Chemical Biology**, London, v. 10, n. 6, p. 477-482, 2014.

THOMPSON, A. J. *et al.* Regulation and manipulation of ABA biosynthesis in roots. **Plant, Cell and Environment**, Oxford, v. 30, p. 67-78, 2007.

TODOROKI, Y.; HIRAI, N. Abscisic acid analogs for probing the mechanism of abscisic acid reception and inactivation. **Studies in Natural Products Chemistry**, Oxford, v. 27, p. 321-360, 2002.

TOSSI, V. *et al.* ABA says NO to UV-B: a universal response? **Trends in Plant Science**, Oxford, v. 17, n. 9, p. 510-517, 2012.

UMEZAWA, T. *et al.* Molecular basis of the core regulatory network in ABA responses: sensing, signaling and transport. **Plant & Cell Physiology**, Kyoto, v. 51, n. 11, p. 1821-1839, 2010.

UMEZAWA, T. *et al.* Type 2C protein phosphatases directly regulate Abscisic acid-activated protein kinases in *Arabidopsis*. **Proceedings of the National Academy of Sciences USA**, Washington, v. 106, p. 17588-17593, 2009.

UNITED STATES DEPARTMENT OF AGRICULTURA. Foreign Agricultural Service. **This report contains assessments of commodity and trade issues made by USDA staff and not necessarily statements of official U.S. government policy**. Washington, 2015. Available from: <http://gain.fas.usda.gov/Recent%20GAIN%20Publications/Coffee%20Annual_Sao%20Paulo%20ATO_Brazil_5-7-2015.pdf>. Access in: 10 Mar. 2016.

VAN DAM, R. M.; FESKENS, E. J. Coffee consumption and risk of type 2 Diabetes Mellitus. **Lancet**, London, v. 360, n. 9344, p. 1477-1478, 2002.

VERSLUES, P. E. *et al.* Methods and concepts in quantifying resistance to drought, salt and freezing, abiotic stresses that affect plant water status. **Plant Journal**, Oxford, v. 45, p. 523-539, 2006.

VIDAL, R. O. *et al.* A high-throughput data mining of single nucleotide polymorphisms in *Coffea* species expressed sequence tags suggests differential homeologous gene expression in the allotetraploid *Coffea arabica*. **Plant Physiology**, Bethesda, v. 154, n. 3, p. 1053-1066, 2010.

VIEIRA, L. G. E. *et al.* Brazilian coffee genome project: an EST-based genomic resource. **Brazilian Journal of Plant Physiology**, Piracicaba, v. 18, p. 95-108, 2006.

VIEIRA, N. G. *et al.* Different molecular mechanisms account for drought tolerance in *Coffea canephora* var. Conilon. **Tropical Plant Biology**, New York, v. 6, n. 4, p. 181-190, 2013.

VINECKY, F. *et al.* Controlled irrigation and NPK fertilization affect the quality of coffee (*Coffea arabica* L. cv. Rubi) beans through their biochemical composition. **Journal of Agricultural Science**, Cambridge, p.1-17, 2016 (online first DOI: 10.1017/S0021859616000988).

VLAD, F. *et al.* Protein Phosphatases 2C Regulate the Activation of the Snf1-Related Kinase OST1 by Abscisic Acid in Arabidopsis. **The Plant Cell**, Rockville, v. 21, p. 3170-3184, 2009.

WANG, W.; VINOCUR, B.; ALTMAN, A. Plant responses to drought, salinity and extreme temperatures: towards genetic engineering for stress tolerance. **Planta**, Berlin, v. 218, p. 1-14, 2003.

WANG, Y. *et al.* Comparative transcriptome analysis of tomato (*Solanum lycopersicum*) in response to exogenous abscisic acid. **BMC Genomics**, London, v. 14, Dec. 2013. Disponível em: <https://www.ncbi.nlm.nih.gov/pubmed/24289302>. Acesso em: 10 mar. 2016.

WASILEWSKA, A. *et al.* An update on abscisic acid signaling in plants and more. **Molecular Plant**, Saint Paul, v. 1, p. 198-217, 2008.

XU, Z. Y. *et al.* A vacuolar-glucosidase homolog that possesses glucose-conjugated abscisic acid hydrolyzing activity plays an important role in osmotic stress responses in Arabidopsis. **The Plant Cell**, Rockville, v. 24, p. 2184-2199, 2012.

XUE, T. *et al.* Genome-wide and expression analysis of protein phosphatase 2C in rice and Arabidopsis. **BMC Genomics**, London, v. 9, p. 550, 2008**.**

YIN, P. *et al.* Structural insights into the mechanism of Abscisic acid signaling by PYL proteins. **Nature Structural & Molecular Biology**, New York, v. 16, p. 1230-1236, 2009.

YOSHIDA, T. *et al.* ABA-Hypersensitive germination3encodes a protein Phosphatase 2C (AtPP2CA) that strongly regulates abscisic acid signaling during germination among Arabidopsis Protein Phosphatase 2Cs. **Plant Physiology**, Bethesda, v. 140, n. 1, p. 115-126, 2006.

YOSHIDA, T.; MOGAMI, J.; YAMAGUCHI-SHINOZAKI, K. Omics approaches toward defining the comprehensive abscisic acid signaling network in plants. **Plant and Cell Physiology**, Kyoto, v. 56, n. 6, p. 1043-1052, 2015.

ZHANG, J.; DAVIES, W. J. Abscisic acid produced in dehydrating roots may enable the plant to measure the water status of the soil. **Plant, Cell and Environment**, Oxford, v. 12, p. 73-81, 1989.

ZHANG, J. *et al.* Genetic analysis of osmotic adjustment in crop plants. **Journal of Experimental Botany**, Oxford, v. 50, p. 291-302, 1999.

ZHANG, X. *et al.* Complex structures of the abscisic acid receptor PYL3/RCAR13 reveal a unique regulatory mechanism. **Structure**, Philadelphia, v. 20, n. 5, p. 780-790, 2012.

ZHANG, X. *et al.* Structural insights into the abscisic acid stereospecificity by the ABA receptors PYR/PYL/RCAR. **PLoS ONE**, San Francisco, v. 8, n. 7, p. e67477, 2013.

Presentation of the PhD Project

Regarding the key roles of PYL/SnRK2/PP2C tripartite system in higher plants, the following scientific questions arisen concerning coffee:

- how many genes composed the PYL/SnRK2/PP2C tripartite system in coffee and how they are organized?
- are these genes expressed in the same manner in different coffee tissues and organs?
- are they differentially expressed in D^T and D^S clones and cultivars of coffee under drought and ABA?
- Does it exist different expression profiles of these genes in *C. arabica* and *C. canephora*?
- Is it possible to identify alleles for improving drought tolerance in *C. canephora* for use in breeding programs?
- is it possible to correlate the diversity of these genes with coffee evolution and adaptation?

In order to get the answers to these questions, the main objectives of this work were:

(i) to identify the candidates genes coding for the ABA proteins receptors (PYR/PYL/RCAR), the phosphatases (PP2C) and kinases (SnRK2) proteins involved in the first steps of ABA signalling pathways in *C. canephora*;

(ii) to characterize these *C. canephora* genes, comparing their families and structure with those described in model plants;

(iii) to identify the functional *C. canephora* PYL/PP2C/SnRK2 orthologs;

(iv) to characterize the expression profile of genes belonging to the tripartite system (PYR/PYL/RCAR-PP2C-SnRK2) in leaves and roots of D^T and D^S clones of *C. canephora* submitted or not to drought stress;

(v) to compare these expression profiles to those obtain in *silico* in different *C. canephora* tissues;

(vi) to study the effects of exogenous ABA on the gene expression of these genes;

(vii) The results obtained regarding these questions are presented in the following chapters.

CHAPTER 1

The PYL/PP2C/SnRK2 tripartite system in *C. canephora*

The increased availability of plant genome data is essential to perform comparative and functional genomic research with insights in plant evolution which can greatly expand the knowledge of the molecular basis of abiotic stress responses in *C. canephora.*

Comparative genomics studies has shown that ABA regulation in water-stress responses is functionally conserved throughout the land plant lineage, from the rise of bryophytes, around 450 million year ago (MYA), to angiosperms (usually represented by *A. thaliana*) that first appeared between 150 and 250 MYA (Doyle, 2012). An essential conservation of responses between the earliest lineages and the flowering plants is clear from studies of the consequences of ABA treatment, or the application of osmotic and drought-stress.

According to Ben-Ari (2012), the identification of orthologs using *A. thaliana* as reference is an excellent approach for functional studies and comparative genomics once *Arabidopsis* is the best studied model species for high plants. Besides phylogenetic considerations, Rubiaceae and Solanaceae are frequently considered as "sisters" plant families based on genetic similarities observed between *C. canephora* and *S. lycopersicum* (Guyot *et al.*, 2012), such as genome size (Noirot *et al.*, 2003; Van der Hoeven, 2002), the basic chromosome number, the cytogenetic chromosome architecture (Pinto-Maglio & Da Cruz, 1998; Hamon *et al.*, 2009; Yu *et al.*, 2011), the absence of polyploidization (Wu *et al.*, 2010) and expressed genes in the seed and cherry (Lin *et al.*, 2005). The structural relationships between *C. canephora*, *S. lycopersicum* and *V. vinifera* genomes were carried-out by Guyot *et al.* (2012) aiming to evaluate the genome conservation and evolution combining comparative mapping at the macro and micro-scale levels. These studies showed that *Solanaceae* microstructures appear much more different than the conservation between *C. canephora* and *V. vinifera* tree, suggesting a divergent and specific evolution of the locus in the *Solanaceae* prior to the separation with the *Rubiaceae*.

Recently, a high-quality draft genome of *C. canephora* was generated which displays a conserved chromosomal gene order among asterid angiosperms (Denoeud *et al.*, 2014). Although there is no sign of the whole-genome triplication as identified in *Solanaceae* species such tomato, the genome includes several species-specific gene family expansions.

In the last years, great efforts have been implemented in genomics to attempt to understand the genetic determinism of tolerance to environmental stresses, biotic and abiotic, especially in model species (Umezawa *et al.*, 2006; Ashraf, 2010). The same applies to coffee for which the recent progress in DNA sequencing methods, genetics and biotechnology permitted the identification of thousands EST sequences (Lin *et al.*, 2005; Poncet *et al.*, 2006; Vieira *et al.*, 2006; Vidal *et al.*, 2010; Mondego *et al.*, 2011), the recent complete genome sequence of *C. canephora* (Denoeud *et al.*, 2014). the construction of genetic maps (Lefebvre-Pautigny *et al.*, 2010, Leroy *et al.*, 2011) and the improvement of genetic transformation techniques (Ribas *et al.*, 2011) These scientific advances now paved the way to investigate the structure of complex gene families in this plant, as it is the case for the genes coding for the proteins of the PYL/PP2C/SnRK2 tripartite system.

Article

MOLECULAR MECHANISMS OF ABA-MEDIATED RESPONSE TO DROUGHT IN LEAVES AND ROOTS OF *COFFEA CANEPHORA*.

Running title: ABA-mediated response to drought in *Coffea canephora*

Michelle Guitton Cotta[1,2], Érica Cristina da Silva Rêgo[3], Stéphanie Sidibe-Bocs[4], Tatiana Santos Costa[3], Fernanda de Araújo Carneiro[1], Jean-François Dufayard[4], Dominique This[2], Pierre Marraccini[3,4], Alan Carvalho Andrade[3*#]

[1]Departamento de Química, Laboratório Central de Biologia Molecular (LCBM), UFLA, Lavras, Minas Gerais, Brazil

[2]Montpellier SupAgro, UMR AGAP, F-34398 Montpellier, France.

[3]Embrapa Recursos Genéticos e Biotecnologia (LGM), Parque EB, CP 02372, 70770-917 Brasília, DF, Brazil.

[4]CIRAD, UMR AGAP, F-34398 Montpellier, France.

E-mail addresses:

MGC: michellegcotta@gmail.com

ECSR: ericacristina.sr@gmail.com

SSB: stephanie.sidibe-bocs@cirad.fr

TSC: tatianaitase@gmail.com

FAC: fearca14@gmail.com

JDF: jean-francois.dufayard@cirad.fr

DM: dominique.this@supagro.fr

PM: marraccini@cirad.fr

ACA: alan.andrade@embrapa.br

present address: Embrapa Café, INOVACAFÉ, Campus UFLA, 37200-000 Lavras, MG, Brazil

*Corresponding author:

Alan Carvalho Andrade

Phone number: +55 35 38294587

ABSTRACT

Abscisic Acid (ABA) pathway is an ancient signaling universally conserved in land plants which coordinates several aspects of the plant response to water deficit such as root architecture, seed dormancy and stomatal regulation. A mechanism of ABA signal transduction has been proposed, evolving intracellular ABA receptors (PYR/PYL/RCARs) interacting with PP2Cs phosphatases and SnRK2 protein kinases. The goal of this study was to identify and characterize for the first time the orthologs of this tripartite system in *C. canephora*. For this purpose, protein sequences from Arabidopsis, citrus, rice, grape, tomato and potato were chosen as query to search orthologous genes in the Coffee Genome Hub (http://coffee-genome.org/). Differential expression in leaves, seeds, roots and floral organs was checked through *in silico* analyses. *In vivo* gene expression analyses were also performed by RT-qPCR in leaves and roots of drought-tolerant (D^T 14, 73 and 120) and -susceptible (D^S 22) *C. canephora* Conilon clones submitted to drought. This approach allowed the identification and characterization of 17 candidate genes (9 PYL/RCARs, 6 PP2Cs and 2 SnRK2s) in *C. canephora* genome. The protein motifs identified in predicted coffee sequences enabled to characterize these genes as family's members of receptors (PYL/RCARs), phosphatases (PP2Cs) or kinases (SnRK2s) of the ABA response pathway. These families were functionally annotated in the *C. canephora* genome. *In vivo* analyses revealed that eight genes are up-regulated under drought conditions in both leaves and roots tissues. Among them, three genes coding phosphatases were expressed in all clones therefore suggesting that they were activated as a general response to cope with drought stress. However, two other phosphatase coding genes were up-regulated only in the D^T clones, suggesting that they may constitute key-genes for drought tolerance in these clones. The D^T clones also showed differential gene expression profiles for five other genes therefore reinforcing the idea that multiple biological mechanisms are involved drought tolerance in *C. canephora*.

INTRODUCTION

The first bitter mouthful in the morning which gives daily energy to the planet is coffee, the major tropical commodity traded worldwide and source of income for many developing countries (Lashermes *et al.*, 2008). With about a third of the world production, Brazil is the first coffee producing country (ICO, 2016). Coffee production is subject to regular fluctuations mainly due to adverse climatic conditions, such as prolonged drought periods. Based on the last report of the Intergovernmental Panel on Climate Change (IPCC), the increase of temperature and drought periods would change the distribution of coffee production zones worldwide leading to environmental, economic and social problems (Davis *et al.*, 2012; Bunn *et al.*, 2015; Ovalle-Rivera *et al.*, 2015) as well as an increase in pests and diseases (Jaramillo *et al.*, 2009; Magrach & Ghazoul, 2015). Drought is a key factor affecting coffee plant development and production (DaMatta and Ramalho, 2006), bean biochemical composition (Vinecky *et al.*, 2016) and quality (Silva *et al.*, 2005).

Among the known 124 perennial species in the coffee genus (Davis *et al.*, 2011), the commercial coffee production concerns only two species, *Coffea canephora* and *C. arabica*. While *C. canephora* is allogamous and diploid (2n=2x=22), *C. arabica* is an autogamous allotetraploid species (2n=4x=44) coming from a natural hybridization between *C. canephora* and *C. eugenioides* ancestrors (Lashermes *et al.*, 1999). Concerning drought tolerance, it is well known that genetic variability exists within *C. canephora* species, the Guinean and SG1 sub-group of Congolese being more tolerant to drought than Congolese plants of SG2 sub-group (Montagnon & Leroy, 1993). Such diversity also exists in Conilon plants of *C. canephora* cultivated in Brazil that are closely related to the SG1 group (Montagnon *et al.*, 2012). Among the strategies commonly observed in coffee plants to cope with water limitation are leaf folding and inclination that reduce water loss and exposure to high irradiance. During the last decade, several drought-tolerant (D^T) and susceptible (D^S) clones of Conilon were identified and previously characterized physiologically (Lima *et al.*, 2002; DaMatta *et al.*, 2003; Pinheiro *et al.*, 2004; Praxedes *et al.*, 2005). At the molecular level, genes differentially expressed under drought were also identified

in leaves of D^T and D^S clones of *C. canephora* (Marraccini *et al.*, 2011, 2012; Vieira *et al.*, 2013), some of them (e.g. *RD29* and *DREB1D*) being linked to ABA-dependent pathways.

It is well known that abscisic acid (ABA) has a central role regulating the adaptive response to drought tolerance in plants (Gonzalez-Guzman *et al.*, 2014). Under stress conditions, this phytohormone, synthesized in roots and leaves during periods of water depletion (Thompson *et al.*, 2007), is perceived by PYR/PYL/RCAR receptors that are the first component of the ABA tripartite systems (Klingler *et al.*, 2010). Once formed, the PYL-ABA complex bind to the clade A phosphatase type 2C (PP2C) inactivating them (Hao *et al.*, 2011; Ma *et al.*, 2009; Park *et al.*, 2009). Then, the subclass III SNF1-related kinase (SnRK2) proteins are activated by dephosphorylation allowing expression of downstream stress responsive genes (Cutler *et al.*, 2010). In this system, SnRK2 and PP2C proteins function therefore as positive and negative regulators of the ABA pathway, respectively.

Using the recently published genome sequence of *C. canephora* (Denoeud *et al.*, 2014), the main objective of this work was (i) to identify the orthologous genes belonging to the tripartite system (*PYL-PP2C-SnRK2*) of ABA in *C. canephora,* , (ii) to characterize these orthologs according to gene structure, protein functional domains, phylogeny, synteny and (iii) to evaluate the expression profile of those genes in leaves and roots of contrasting (D^T and D^S) clones *C. canephora* submitted or not to drought conditions.

MATERIAL AND METHODS

Plant material

Drought-tolerant (D^T: 14, 73 and 120) and -susceptible (D^S: 22) clones of *C. canephora* Conilon were grown in greenhouse conditions (under controlled temperature 25°C, relative humidity of 70% and photosynthetic flux PPF 900 $\mu mol^{-2}s^{-1}$) at UFV (University of Viçosa-UFV, Minas Gerais, Brazil). At 6 months old, drought stress was applied to the plants by water withdrawal (NI: non-irrigated) to reach a predawn leaf water potential (Ψ_{pd}) of around -3.0 MPa. From each clone, biological triplicate samples

(leaves and roots) were collected in both irrigated (I: control) and NI conditions, immediately frozen in liquid nitrogen and stored at -80°C for RNA extractions and ABA quantification.

Genomic data

Genomic data from a double haploid accession of C. canephora available in Coffee Genome Database (http://coffee-genome.org/, Dereeper *et al.* [2015]) were used as reference sequences.

In silico identification and characterisation of candidate genes of the PYR/PYL/RCAR-PP2C-SnRK2 tripartite system

PYR/PYL/RCAR-PP2C-SnRK2 orthologs genes from *C. canephora*, and their orthologs from *Arabidopsis thaliana*, *Solanum lycopersicum*, *Solanum tuberosum*, *Vitis vinifera*, *Citrus sinensis* and *Oryza sativa* were identified in the following databases: NCBI (http://www.ncbi.nlm.nih.gov/), TAIR (http://www.arabidopsis.org/), AtGDB (http://www.plantgdb.org/AtGDB/), Phytozome (http://www.phytozome.net/), Sol Genomics Network (http://solgenomics.net/), SlGDB (http://www.plantgdb.org/SlGDB/), GreenPhyl (http://www.phytozome.net/), Grape Genome Database (http://www.genoscope.cns.fr/externe/), Gramene Database (http://www.gramene.org/), Plant Genome Database (http://www.plantgdb.org/), Citrus Genome Database (http://www.citrusgenomedb.org/) and Rice Genome Annotation (http://rice.plantbiology.msu.edu). BLAST searches were carried out using these sequences as query against the Coffee Genome Database and Rubiaceae ESTs database (e-value $< e^{-10}$) to isolate coffee genes that were further translated to compare their corresponding proteins with proteins of other species using the MAFFT program (Katoh & Toh, 2008) available at South Green Platform (http://www.southgreen.fr/). The conserved amino acids were identified using the GeneDoc program (http://www.nrbsc.org/gfx/genedoc/). Genes that did not contain specific domains were removed. Phylogenetic analyses were performed with orthologous sequences that were filtered with Gblocks (Castresana, 2000) and used to construct the phylogenetic trees using PhyML algorithm (Guindon *et al.*, 2010). To compare gene and species, reconciled trees were constructed using the RAP-

Green algorithm (Dufayard *et al.*, 2005) and the reference tree provided by the NCBI taxonomic database (http://www.ncbi.nlm.nih.gov/sites/entrez?db=taxonomy). The HMM (Hiden Markov Model) was used to build and validate the phylogenetic analyses that were visualized using the Dendroscope software (Huson *et al.*, 2007). All candidate genes were functionally annotated in the Coffee Genome Database using Artemis software (Carver *et al.*, 2012). Gene structures were predicted using the Gene Structure Display Server (http://gsds.cbi.pku.edu.cn/). The transcriptomic data available in the Coffee Genome Database were used to perform in silico expression analyses that were normalized using RPKM (Fig. S1). The gene duplication patterns were generated using the MCScanX software (Wang *et al.*, 2012) and were formatted by Circos (http://circos.ca/) for graphical representation.

RNA extraction and real-time qPCR assays

Total RNAs were extracted from leaves and roots of C. canephora as previously described (Marraccini *et al.*, 2011). Contaminant genomic DNA was eliminated from purified RNAs by RQ1 RNase-free DNase (Promega) treatment according to the fabricant. RNA integrity was verified by agarose gel electrophoresis with ethidium bromide staining. Synthesis of the first-strand cDNA was done by treating 2.4µg of total RNA with the ImProm-II Reverse Transcription System and oligo (dT15) according to the manufacturer's recommendations (Promega). Real-time qPCR assays were carried out with the synthesized single-stranded cDNA using the protocol recommended for 7500 Fast Real-Time PCR Systems (Applied Biosystems, Foster City, CA, USA). cDNA preparations were diluted (1/20) and tested by qPCR using primer pairs (Table S1) designed using the Primer Express software (Applied Biosystems) and preliminarily tested for their specificity and efficiency against a cDNA mix from roots and leaves. The qPCR was performed with 1µl of diluted single-stranded cDNA and 0.2 µM (final concentration) of each primer in a final volume of 10µl with 1x SYBR green fluorochrome (SYBRGreenqPCR Mix-UDG/ROX, Invitrogen). The reaction was incubated for 2 min at 50°C and 5 min at 95°C (UDG step), followed by 40 amplification cycles of 3 s at 95°C, 30 s at 60°C. Data were analysed using the SDS 2.1 software (Applied Biosystems) to determine the cycle threshold (Ct) values.

Specificity of the PCR products generated for each set of primers was verified by analysing the Tm (dissociation) of amplified products. Gene expression levels were normalized to expression level of ubiquitin (CcUBQ10) as a constitutive reference (Barsalobres-Cavallari *et al.*, 2009). Expression was expressed as relative quantification by applying the formula $(1+E)^{-\Delta\Delta Ct}$, where $\Delta Ct_{target} = Ct_{target\ gene} - Ct_{reference\ gene}$ and $\Delta\Delta Ct = \Delta Ct_{target} - \Delta Ct_{internal\ calibrator}$.

ABA extraction and quantification

ABA was extracted from leaves and roots tissues of C. canephora clones stored at -80°C as previously mentioned (see plant material section). Initially, samples were lyophilised and ground to a power in liquid nitrogen. ABA was extracted (Berry & Bewley, 1992) and quantified by ELISA using the Phytodetek ABA test kit (Agdia, Elkhart, IN, USA).

Statistical analyses

The statistical analyses were performed using GraphPad Prism software (GraphPad Software Inc., La Jolla, CA, USA).

RESULTS

Identification of CcPYLs-PP2Cs-SnRK2s orthologs in C. canephora

The protein sequences of ABA receptors, phosphatases and kinases from *A. thaliana*, *C. sinensis*, *V. vinifera*, *S. lycopersicum*, *S. tuberosum* and *O. sativa* were used as query to identify orthologous genes in *C. canephora* through BLASTP (Table S2-S4), leading to the identification of 17 putative coffee proteins according to the analysis of their functional domains. Nine proteins were homologous to the PYR/PYL/RCAR (Fig. 1a), six to clade-A PP2C (Fig. 1b) and seven putative coffee SnRK2 kinases belonging to subclass I and II (Fig.1c). Two additional SnRK2s of subclass III were also identified (Fig. 1d). These genes were named according to the results of phylogenetic analyses and

sequence homology, as follows: *CcPYR1* (*Cc08_g02750*), *CcPYL2* (*Cc08_g10450*), *CcPYL4* (*Cc02_g05990*), *CcPYL7a* (*Cc00_g17440*), *CcPYL7b* (*Cc00_g23730*), *CcPYL8a* (*Cc02_g01800*), *CcPYL8b* (*Cc08_g15960*), *CcPYL9* (*Cc02_g39180*), *CcPYL13* (*Cc02_g15060*), *CcABI1* (*Cc08_g11010*), *CcABI2* (*Cc06_g11740*), *CcAHG2* (*Cc08_g16010*), *CcAHG3* (*Cc02_g07430*), *CcHAB* (*Cc04_g01620*), *CcHAI* (*Cc01_g13400*), *CcSnRK2.1* (*Cc00_g19320*), *CcSnRK2.2* (*Cc07_g05710*), *CcSnRK2.6* (*Cc02_g18420*), *CcSnRK2.8* (*Cc10_g06790*), *CcSnRK2.8* (*Cc07_g14700*), *CcSnRK2.10* (*Cc02_g22790*), *CcSnRK2.11* (*Cc08_g11200*), *CcSnRK2.12* (*Cc00_g35430*) and *CcSnRK2.13* (*Cc00_g07830*).

Most of *CcPYLs-PP2Cs-SnRK2s* genes were found in chromosome 2 of *C. canephora* (Fig. 2a). Regarding *PYR/PYL/RCAR* gene family, the *CcPYL7a* and *CcPYL7b* genes were located on the chromosome 0. The seven others *CcPYLs* genes were on the chromosomes 2 (*CcPYL4*, *CcPYL8a*, *CcPYL9* and *CcPYL13*) and 8 (*CcPYR1*, *CcPYL2* and, *CcPYL8b*). The six clade-A *PP2Cs* genes were positioned on five different chromosomes: *CcHAI* in chr1, *CcAHG3* in chr2, *CcHAB* in chr4, *CcABI2* in chr6 and *CcABI1* and CcAHG2 in chr8. The *CcSnRK2.6* and *CcSnRK2.2* of subclass III were located on the chr2 and chr7, respectively. For the seven *SnRK2* genes of subclasses I and II, *CcSnRK2.1*, *CcSnRK2.12* and *CcSnRK2.13* were located on the chr0 whereas *CcSnRK2.*10, *CcSnRK2.*8, *CcSnRK2.*11 and *CcSnRK2.*8 were in chr2, chr7, chr8 and chr10, respectively.

Functional annotation of CcPYL-PP2C-SnRK2 genes

The 24 genes of the coffee tripartite system were functionally annotated on *C. canephora* genome (Fig. 2b). The occurrence of duplication events in the *CcPYL-PP2C-SnRK2* gene families was investigated through analyses of the paralogous regions. These analyses showed that CcPYL proteins shared high identity with ABA receptors from grape, while the CcPP2Cs were closely related to tomato and potato phosphatases, and CcSnRK2s with citrus kinases (Table S2-S4). Except the CcSnRK2.12 and CcSnRK2.13 proteins which not contained all domains (Fig. 1c), the lengths of CcPYL, CcPP2C and CcSnRK2 protein sequences were between 174-231, 418-546 and 336-363 amino acids, respectively (Fig. 1). The phylogenetic trees showed that PYL receptors and SnRK2 proteins were distributed in the

three main subfamilies (Fig. 3a and 3b, Fig. S2 and S4). The putative protein sequences of *CcPYL7a*, *CcPYL7b*, *CcSnRK2.1*, *CcSnRK2.12-13* coding-genes located on chr0 were not showed on the resumed phylo-analyses, however, they are represented in the complete ones (Fig. S2 and S4).

ABA (PYR/PYL/RCAR) receptors

Among the nine PYR/PYL/RCAR proteins, CcPYR1, CcPYL8a and CcPYL9 showed high sequence identity (72%, 83% and 84%) with tomato sequences while CcPYL2 and CcPYL4 shared 84% and 74% of identity with the potato proteins, and CcPYL7a and CcPYL7b had 54% and 53% of identity with the same grape locus while the CcPYL8a and CcPYL8b proteins shared 82% of identity. Finally, the CcPYL13 showed 62% of sequence identity with grape GSVIVG01013161001 protein. BLASTP results showed that CcPYL4, CcPYL7a and CcPYL7b proteins were highly homologous to AtPYL6 from *A. thaliana*, CsPYL5 from *C. sinensis* and VvRCAR6 from *V. vinifera*, respectively (Table S2). The CcPYL4, CcPYL7a and CcPYL7b proteins also shared high identity respectively with the Solyc10g076410, Solyc10g085310 and Solyc03g095780 proteins of *S. lycopersicum*. All these coffee PYR/PYL/RCAR proteins (including in CcPYL7a and CcPYL7b located on chr0), contained key amino acid residues involved of both gate and latch loops conserved in ABA receptors (Fig. 1a).

The seven mapped *CcPYL* genes were identified on different ancestral blocks of the seven eudicot chromosomes such as the G2 (*CcPYR1*, *CcPYL8a*, *CcPYL8b* and *CcPYL9*), G4 (*CcPYL2*), G6 (*CcPYL4*) or G7 (*CcPYL13*) groups. All *CcPYLs* genes identified on the G2 ancestral block were located at the edges of their respective chromosomes. In addition of being located on the same chromosome (chr8) and G2 ancestral block, the *CcPYR1* and *CcPYL8b* genes also belonged to the same paralogous region (Fig. 2a). A different situation was observed for the *CcPYL8a* and *CcPYL9* genes that derived from different paralogous regions.

Manual curation of *CcPYL* genes revealed that *CcPYR1*, *CcPYL2*, *CcPYL7a*, *CcPYL7b* and *CcPYL13* did not contain introns, while one intron was found in *CcPYL4* and two in *CcPYL8a*, *CcPYL8b* and *CcPYL9* genes (Fig. 2b). No evidence of 5' UTRs regions was found for *CcPYL2*, *CcPYL7a* and *CcPYL7b* genes. *CcPYL8a* was the only gene presenting an intron (of 316 bp length) on the 5' UTR

region. For *CcPYL8b*, a 3'UTR extension was based on sequence alignments with a corresponding EST of *C. arabica* (GR997267) expressed in leaf, fruit, flower, root and calli tissues. The *CcPYL8b* gene was also extended in its 5' UTR using the similarities found in *PYL9* genes of tomato (LOC101258886) and potato (LOC102591194) (Table S2). In the same way, an extension was also found in *CcPYL13* 5'UTR region based on the GT013431 EST sequence of *C. arabica* expressed in fruits. Phylogenetic analyses revealed that the ABA receptors CcPYL8a, CcPYL8b and CcPYL9 belong to the subfamily I together with AtPYL7-10 from *A. thaliana*, CsPYL8-9 from *C. sinensis* and VvPYL8-9 from *V. vinefera* (Fig. 3a, Fig. S2). The CcPYL4 protein was located in the subfamily II together with AtPYL4-6 and CsPYL4-5 while CcPYL13 was closely related to AtPYL11-13. Finally, the subfamily III contained the CcPYR1 and CcPYL2 proteins, the first being related to the AtPYR1, AtPYL1 and CsPYR1 proteins, and the second to AtPYL2-3 and CsPYL2 proteins.

Phosphatase type 2C (PP2C) proteins

The majority of coffee PP2Cs were identical to phosphatases proteins from *Solanaceae* (Table S3). Among them, the CcABI1, CcABI2, CcHAB and CcHAI were highly similar to potato proteins while CcABI2, CcAHG3 and CcHAB were related to tomato sequences. On the other hand, the CcAHG2 protein presented 57% of identity with a grape sequence. The catalytic domain of PP2Cs composed of 11 conserved motifs with Mg^{2+}/Mn^{2+} [xxD] and [DG] (D: aspartic acid, G: glycine) motifs, was highly conserved throughout the six coffee PP2Cs (Fig. 1b).

Regarding genome localization, these proteins evolved from G2 (*CcAHG2*), G4 (*CcHAB*, *CcABI1*, *CcABI2*) and G6 (*CcHAI* and *CcAHG3*) ancestral blocks (Fig. 2a). Even thought *CcHAB*, *CcABI1* and *CcABI2* in one hand, and *CcHAI* and *CcAHG3* in another, evolved from the same ancestral block genes, all these genes belonged to a different paralogous region. *CcAHG2* and *CcAHG3* contained three introns while four were observed in *CcABI1*, *CcABI2*, *CcHAI* and *CcHAB* genes (Fig. 2b). Only *CcABI1* and *CcABI2* genes contained introns in their corresponding 5'UTR regions.

The phylogenetic analyses revealed that CcABI1-2 and CcHAB proteins evolved together with AtABI1-2, AtHAB1-2, CsABI1 and CsHAB1 while CcAHG2, CcAHG3 and CcHAI were grouped with

AtAHG1, AtAHG3, AtHAI1-3 and CsAHG3 (Fig. 3b). The members of the ABA-hypersensitive germination (AHG) subfamily in *C. canephora* were represented by *CcAHG2* and *CcAHG3* genes. Because *CcAHG2* had no ortholog in *A. thaliana* and presented low homology with *AtAHG1*, *AtAHG3*, *AtHAI1-3* and *AtABI1-2* genes, it clustered separately from these genes (Fig.3b, Fig. S3). On the other hand, *CcAHG3* was orthologous to *AtAHG3*, *CsAHG3* and *VvPP2C8*. Finally, the coffee *CcHAI* appeared homologous to ABA-induced genes *AtHAI1*, *AtHAI2* and *AtHAI3*.

SNF1-related (SnRK2) protein kinases

Nine putative SnRK2 protein kinases were identified in *C. canephora*. CcSnRK2.1, CcSnRK2.8, CcSnRK2.12 and CcSnRK2.13 shared high identity (84%, 83%, 92% and 84%) with their respective proteins of *C. sinensis* (Table S4). On the other hand, CcSnRK2.2, CcSnRK2.6 and CcSnRK2.11 had 86%, 93% and 83% of identity with tomato proteins while CcSnRK2.2 and CcSnRK2.10 proteins shared 86% and 91% with potato relatives. Excepted CcSnRK2.12 and CcSnRK2.13, all other coffee SnRK2s contained in their N-terminal region the GXGXXG kinase (ATP binding) domain and the highly acidic ABA box domain (motif I) important for their interactions with PP2Cs in their C-terminal region (Fig. 1c). In addition to these domains, CcSnRK2.2 and CcSnRK2.6 also contained the C-terminal domains I and II (Fig. 1d) responsible of SnRK2 activation by osmotic stress in ABA-independent and ABA-dependent manners, respectively (Yoshida *et al.*, 2006).

At the gene level, *CcSnRK2.1* had orthologous genes in tomato (*SlSnRK2.1*), grape (*VviSnRK2.12*), and *Arabidopsis* (*AtSnRK2.1* and *AtSnRK2*.5). On the other hand, *CcSnRK2.10* was orthologous to *VvSnRK2.11* and homologous to *AtSnRK2.10* and *AtSnRK2.4* of Arabidopsis. The *CcSnRK2.7* had an ortholog in Arabidopsis *(AtSnRK2.7*), a co-ortholog in tomato (*SlSnRK2C*) and two homologs in grape (*VvSnRK2.7a* and *VvSnRK2.7b*). *CcSnRK2.8* had two orthologs in Arabidopsis (*AtSnRK2.8*) and grape (*VviSnRK2.8*). According to the classification of Kobayashi *et al.* (2004), *CcSnRK2s* were divided into three subclasses which differed by their activation in response to ABA (Fig. 3c). The *CcSnRK2.1* and *CcSnRK2.10* clustered in the subclass I corresponding to genes not activated in the presence of ABA. The *CcSnRK2.7* and *CcSnRK2.8* belong to the subclass II activated

to a lesser extent by ABA. Finally, the subclass III was composed by *CcSnRK2.2* and *CcSnRK2.6* genes strongly activated by ABA. Interestingly, the coffee *CcSnRK2.11* gene did not clustered in any of these subclasses (Fig. S4).

Concerning genome localization, *CcSnRK2.11* gene belong to the G2 ancestral block while *CcSnRK2.2*, *CcSnRK2.6 CcSnRK2.8*, *CcSnRK2.7* and *CcSnRK2.10* genes were identified on the G3 block (Fig. 2a). Among them, *CcSnRK2.8* and *CcSnRK2.10*, as well as *CcSnRK2.2* and *CcSnRK2.7* genes, evolved from the same ancestral block and paralogous regions. Excepted *CcSnRK2.12* and *CcSnRK2.13* genes that contained four and five introns, respectively, other *SnRK2* genes contained eight introns (Fig. 2b). Because *CcSnRK2.12* and *CcSnRK2.13* genes also missed a stop codon, they were considered as uncompleted sequences and were not further analyzed.

Expression profiles of CcPYLs-PP2Cs-SnRK2 genes in leaves and roots of C. canephora submitted to drought conditions

Expression of PYL/PYR/RCAR-PP2C-SnRK2 genes was analyzed in leaves and roots of the D^T and D^S clones of *C. canephora* grown under I (irrigated) and NI (non-irrigated) (Fig. 4). Whatever the primer pairs designed for *CcPYL7a*, *CcPYL7b* and *CcPYL13* (Table S1) and irrigation conditions, no expression was detected in leaves and roots (data not shown). For *CcPYR1* and *CcPYL4*, expression was observed in leaves of all clones under control condition and decreased under drought. A similar pattern was observed for *CcPYL2*, except that this gene was expressed under irrigation only in leaves in D^T clones 14 and 120. Expression levels of *CcPYL8a* gene did not changed significantly from I to NI conditions in leaves of D^T clones 14, 73 and 120 but decreased significantly under drought in D^S clone 22. Whatever the clones, *CcPYL8b* and *CcPYL9* were the most expressed genes in leaves of *C. canephora* plants under irrigation. However, expression of *CcPYL8b* and *CcPYL9* genes increased significantly under drought in leaves of clones 22, 73 and 120, and in those of D^T clones 14 and 73, respectively.

In roots, expression of *CcPYR1*, and *CcPYL4* decreased under drought in all clones of *C. canephora* (Fig. 4). On the other hand *CcPYL8b* gene expression was significantly induced by drought

in the D^T clones 73 and 120. Up-regulated expression of *CcPYL9* was also noticed in D^T clone 73 under drought.

For the *PP2C* genes, expression of *CcABI2*, *CcAHG3* and *CcHAI* genes were significantly up-regulated under drought in leaves and roots of both D^T and D^S clones (Fig. 4). Drought-induced expression of*CcAHG2* and *CcHAB* was also observed but only in leaves of the D^T clones 14, 73 and 120. Different expression profiles were observed for *CcABI1* in leaves and roots with water conditions. For example, *CcABI1* expression was up-regulated under drought only in D^T clone 73, but down-regulated by drought in D^S clone 22. On the other hand, while *CcABI1* gene expression decreased under drought in the D^T clone 14, it was highly up-regulated by drought in the D^T clone 120. Compared to other *PP2Cs*, *CcHAI* was the most expressed in both leaves and roots under drought. Expression of *CcAHG2* was also greatly up-regulated by drought in leaves of all D^T clones but not in those of D^S clone 22. However, no detectable expression of this gene was observed in roots (data not shown). In roots, the D^T clone 120 stands out other clones by the fact that it presented high up-regulated expression under drought of *CcABI1*, *CcABI2* and *CcAHG3*, as well as of *CcSnRK2.2*, *CcSnRK2.6* and *CcSnRK2.8* genes. The expression of subclass III *CcSnRK2.2* gene was also up-regulated by drought in leaves of the D^T clones 14 and 73 but also in roots of all D^T clones. On the other hand, *CcSnRK2.6* gene expression was unaffected by water condition in leaves but increased under NI conditions only in roots of D^T clone 73.

The expression of *CcSnRK2.7* gene increased under NI conditions in leaves in the D^T clone 73 and mainly in roots of D^T clone 120 (Fig. 4). On the other hand, expression of *CcSnRK2.8* was up-regulated by drought only in leaves of D^T clone 73 as well as in roots of D^T clone 73 and D^S clone 22. No significant differences of expression profiles were observed for *CcSnRK2.10* and *CcSnRK2.11* in leaves of all coffee clones. However, *CcSnRK2.10* gene expression was down-regulated under drought in roots of D^T clones 14 and 120, but unaffected in clones 73 and 22. While *CcSnRK2.11* expression was not detected in roots (data not shown), expression levels detected in leaves were not significantly affected by water treatments. For *CcSnRK2.1*, as well as *CcSnRK2.12* and *CcSnRK2.13*, expression was

undetectable in drought-stressed roots and leaves of all coffee clones with the tested primer pairs (data not shown).

ABA quantification

In leaves, ABA was detected in all clones under both I and NI conditions, ranging from 2 to 8 pmol.g^{-1} of DW (Fig. 5). A significant increase of ABA content under drought was observed in leaves of D^T clone 120 while ABA contents were considered as relatively stable in other clones whatever the irrigation conditions. In roots, ABA contents were similar (around 4 pmol.g^{-1} of DW) in all clones under irrigated conditions. If these contents tended to decrease under drought in all clones, this reduction was significant only in roots of the D^T clone 14.

DISCUSSION

For the first time, the orthologous genes coding for proteins of the PYL/PYR/RCAR-PP2C-SnRK2 tripartite system involved in the first steps of ABA perception and signal transduction were identified and thoroughly characterized in *C. canephora*. Based on sequence similarity with other plant genes, nine *CcPYL*-type genes, six *PP2C*-type and nine *SnRK2*-type genes divided in three subclasses were found.

PYR/PYL/RCAR gene family in C. canephora

Nine *PYR/PYL/RCAR* genes were found in the *C. canephora* genome. This number was similar to *PYLs* found in *C. sinensis* (Romero *et al.*, 2012) and *V. vinifera* (Boneh *et al.*, 2012b), but smaller than *PYLs* in Arabidopsis (Ma *et al.*, 2009; Park *et al.*, 2009), tomato (Gonzalez-Guzman *et al.*, 2014) and rice (Kim *et al.*, 2012). Interestingly, *C. canephora* contained duplicated genes of *PYL7* (*CcPYL7a* and *CcPYL7b*) and *PYL8* (*CcPYL8a* and *CcPYL8b*) (Fig. 6). The duplicated *CcPYL7s* were located into the chr0 corresponding to unmapped scaffolds grouped arbitrary in a pseudomolecule (Denoeud *et al.*, 2014) and not expressed in leaves or roots of *C. canephora*. These results are in accordance with *in silico* data deduced from the Coffee Genome Database (Fig. S1). However, since *CcPYL7a* and

CcPYL7b were expressed in developing beans of *C. arabica* (data not shown), it can be concluded that they correspond to functional genes like *CcPYL8a* (chr2) and *CcPYL8b* (chr8).

Denoeud *et al.* (2014) recently reported that the coffee genome contained several species-specific gene families that probably occurred by segmental and tandem gene duplication, as well as transposition events. Despite the fact that *CcPYL8a*, *CcPYL8b* and *CcPYL9* harboured different chromosome localizations, their chromosome position, origin, similar gene structure and expression profiles suggested that they underwent duplications (Fig. 6). This hypothesis is supported by the fact that these genomic fragments harboured other duplicated genes (e.g. lipid transfer protein, zinc finger DOF protein, heat shock protein, Dehydration-responsive element-binding protein 1D) (data not shown) previously shown to be important in responses of *C. canephora* (Marraccini *et al.*, 2012; Vieira *et al.*, 2013) and *C. arabica* (Mofatto *et al.*, 2016) to drought.

In the present work, *CcPYL8a* and *CcPYL8b* paralogs showed different expression profiles in roots under drought. Such differences could be explained by the presence of the 316 bp intron in the 5' UTR region of *CcPYL8a* affecting expression of this gene. This hypothesis is reinforced by the presence of two LTR *copia* retrotransposons in *CcPYL8a*, one located in its promoter region (2 kb) and the other in its first intron. TEs located near host genes are known to impact gene expression and to play a role in the genome adaptation to environmental changes (Casacuberta & González, 2013), as suggested in coffee where high TEs expression was observed in *C. canephora* and *C. arabica* submitted to drought (Lopes *et al.*, 2013).

To our knowledge, the results presented here are the first reporting functional duplication of *PYL8* gene. They demonstrated that *CcPYL8b* and *CcPYL9* were the genes mostly expressed in roots and leaves of *C. canephora* indicating their probable key role to cope with drought in coffee, as also suggested in Arabidopsis (Ma *et al.*, 2009; Zhao *et al.*, 2014). However, expression of *CcPYL8a, CcPYR1, CcPYL2* and *CcPYL4* was unaffected by drought, suggesting that these genes played a limited role in the response of *C. canephora* to water limitation.

Coffee PP2C gene family

Six *CcPP2Cs* were identified in the *C. canephora* genome. This gene number is higher to that found in *C. sinensis* (Romero *et al.*, 2012), but lower to that of Arabidopsis (Ma *et al.*, 2009; Park *et al.*, 2009), grape (Boneh *et al.*, 2012a), tomato (Sun *et al.*, 2011) and rice (Xue *et al.*, 2008). Expression analyses revealed that these coffee genes were functional since they were all expressed in leaves particularly in drought stressed coffee. Among them, *CcHAI* retained attention since its expression was low under unstressed conditions but highly induced under drought in all clones. In roots, this gene was highly up-regulated under drought in D^T clone 73 and D^S clone 22, while the increase was much more reduced in D^T clones 14 and 120. In Arabidopsis, *hai* mutants exhibited inhibition of root growth and induction of many ABA-regulated genes such as dehydrins, late embryogenesis abundant proteins, *NCED3* and *NACs* (Bhaskara *et al.*, 2012). Here, *CcHAI* was the gene mostly up-regulated under drought in leaves and roots of all *C. canephora* clones, suggesting its key role in coffee responses to drought.

Several studies already reported induced expression of the *PP2C* genes under abiotic stress (Tähtiharju & Palva, 2001), as observed for *ABI1*, *ABI2* and *HAB1* in leaves of Arabidopsis early during drought treatment (Harb *et al.*, 2010). *ABI1* is a key gene of ABA signaling in the guard cells where ABI1 inhibition after ABA perception stimulates stomatal closure (Saez *et al.*, 2006). Such a role is not expected in roots where expression of *CcABI1* and *CcABI2* was highly up-regulated under drought, particularly in D^T clone 120. The fact that these two genes exhibited similar expression profiles could be explained by their overlapping roles in controlling ABA action (Leung *et al.*, 1997; Merlot *et al.*, 2001). The up-regulated expression of *CcAHG3* also observed in parallel to the accumulation of *CcABI1* and *CcABI2* transcripts might be related with the function of this gene in ABA response pathway (Nishimura *et al.*, 2004). As previous studies shown that high concentrations of ABA inhibit root growth (Beaudoin *et al.*, 2000), it is possible that these *PP2C* genes could act together on the development of coffee root system under drought.

Another interesting result concerned *CcAHG2* whose expression was significantly up-regulated in leaves under drought specifically in D^T clones and undetected in roots. These expression profiles are not contradictory to those of *in silico* (Fig. S1) that did not detected *CcAHG2* expression in leaves since

RNA-seq libraries were generated from unstressed coffee plants (Dereeper *et al.*, 2015). The fact that *CcAHG2* was expressed in drought-stressed leaves of D^T clones but not in those of D^S clone 22, highly suggests a key function of this gene in leaves of *C. canephora* D^T clones submitted drought.

SnRK2 gene family in C. canephora

In this work, nine putative *SnRK2* genes were identified in the *C. canephora* genome. Expression studies revealed that *SnRK2.2* was up-regulated upon drought in leaves of D^T clones 14 and 73. For other *SnRK2* genes, expression levels can be considered as relatively stable and poorly affected by drought in all *C. canephora* clones. An opposite situation was observed in roots in which the expression profiles of *CcSnRK2.2, CcSnRK2.6* and *CcSnRK2.7* genes were highly up-regulated upon drought, mainly in D^T clone 120.

Among SnRK2 proteins, those of subgroup III (e.g. *SnRK2.2* and *SnRK2.6*) play important roles in ABA-induced stomatal closure (Cutler *et al.*, 2010). Phosphorylated forms of SnRK2.2 and 2.6 were also reported to activate the ABA-responsive structural gene *RD29B* (Yoshida *et al.*, 2010). Zheng *et al.* (2010) also reported the role of *SnRK2.6* in increasing carbon supply and stimulating plant growth. Even though some functional redundancy had been postulated between SnRK2.2 and SnRK2.6 (Fujii & Zhu, 2009), our results clearly suggest a key role of these kinases in response to drought, mainly in roots of *C. canephora*.

In contrast to subgroup III, the main targets of subgroup II SnRK2s are stress-responsive genes coding transcription factors (Kulik *et al.*, 2011). For example, Zhang *et al.* (2010) showed that over-expression of wheat *SnRK2.8* in Arabidopsis enhanced tolerance to drought, salt and cold stresses by up-regulating the expression of genes involved in ABA biosynthesis and signaling. On the other hand, *A. thaliana* over-expressing *SnRK2.7* from wheat showed enhanced photosystem II activity and root growth (Zhang *et al.*, 2011). Even though SnRK2.7 and SnRK2.8 might be functionally redundant, *SnRK2*.7 was shown to be expressed in roots, leaves and flowers of Arabidopsis while *SnRK2.8* was mainly expressed in roots, indicating different tissue specificities of these two kinases (Mizoguchi *et al.*,

2010). The up-regulated expression of *CcSnRK2.7* and *CcSnRK2.8* in roots of drought-stressed *C. canephora*, led us to propose key functions of both kinases in coffee roots.

How the tripartite system PYL-PP2C-SNRK2 of ABA perception could explain D^T and D^S phenotypes of C. canephora clones?

In higher plants, ABA content is rigorously controlled by the rate of biosynthesis, catabolism, compartmentalization and transport, increasing in both roots and leaves in response to water deficit. Here, we showed that D^T clone 120 was the only one presenting significant increase of leaf ABA content under drought. Whatever the *C. canephora* clone, no significant differences of ABA contents were observed in roots, therefore indicating that D^T and D^S phenotypes were probably due to altered ABA signalling pathway rather than deficiencies of ABA synthesis.

Previous studies revealed that transport rate/CO_2 assimilation (*ETR*/*A*) ratio was significantly higher under drought in D^T clone 73 compared to D^T clones 14 and 120, therefore suggesting the participation of an alternative electron sink protecting the photosynthetic apparatus against photoinhibition by limiting electron accumulation and ROS formation in clone 73. Interestingly, drought-induced up-regulated expression of genes encoding for ascorbate peroxidase (*CcAPX1*), a prephenate-dehydrogenase like protein (*CcPDH1*) and a non-symbiotic haemoglobin (*CcNSH1*) was already reported in this clone, suggesting its protection involved strong induction of antioxidant and osmoprotection systems (Vieira *et al.*, 2013). The up-regulated expression of *SnRK2.2*, *SnRK2.7* and *SnRK2.8* upon drought in its leaves could participate in activating such pathways.

Another interesting result concerned *CcAHG2* that was expressed only in leaves of all D^T clones but not in those of D^S clone 22. Because *CcAHG2* lacks ortholog in *A. thaliana*, its biochemical function is unknown. Despite this, there are two CcAHG2 orthologous in Solanum species (Solyc08g082260 in *S. lycorpersicum* and PGSCOOO3DMP400066209 in *S. tuberosum*). Further research is therefore needed to know if *CcAHG2* could be used as a molecular marker of drought tolerance in coffee.

Compared to D^S clones of *C. canephora*, it was already reported that D^T clone 120 had a deeper root system that should allow greater access to soil water (Pinheiro *et al.*, 2005). Interestingly, *CcPYL8b*,

but also of *SnRK2* (*CcSnRK2.2*, *CcSnRK2.6* and *CcSnRK2.7*) and *PP2C* (*CcABI1*, *CcABI2* and *CcAHG3*) genes were highly up-regulated under drought in roots, indicating a key role of root system in responses to drought in this clone. Even though D^T clone 14 also had a root depth similar to clone 120 (Pinheiro *et al.*, 2005), it did not showed up-regulate expression of *PYL*, *SnRK2* and *PP2C* genes in roots, indicating the existence of different mechanisms amongst the D^T coffee clones regarding water deficit (Vieira *et al.*, 2012). Whatever it is, the differences observed for *SnRK2* (mainly of subclass III) gene expression profiles clearly indicated the involvement of the ABA-dependent signalling pathway in the response to drought, at least in D^T clones. Of course, this does not preclude the involvement of other hormonal regulatory pathways in the establisment of drought tolerance phenotypes in coffee. For example, up-regulated expression of subclass II *SnRK2* genes by salicylic acid, ethylene, and jasmonates, has already been reported (Kulik *et al.*, 2011). The occurence of such entangled crosstalks between biotic and abiotic pathways might exist in coffee, as suggested by the fact that the D^T clone 14 was also recently identified as resistant to multiple races of root-knot nematodes *Meloidogyne* (Lima *et al.*, 2015).

ACKNOWLEDGEMENTS

This work was carried out under the Embrapa-CIRAD scientific cooperation project entitled "Genetic determinism of drought tolerance in coffee". The authors acknowledge financial support from the Brazilian Consortium on Coffee R&D, Instituto Nacional de Ciência e Tecnologia – Café (Conselho Nacional de Desenvolvimento Cientifico e Tecnologico/Fundação de Amparo à Pesquisa do Estado de Minas Gerais) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior/Comité Français d'Évaluation de la Coopération Universitaire et Scientifique avec le Brésil (Project 407-2012/University of Lavras and Montpellier SupAgro). The authors wish also to thank Peter Biggins (CIRAD-DGDRS) for English revision of the manuscript.

REFERENCES

Barsalobres-Cavallari CF, Severino FE, Maluf MP, Maia IG. 2009. Identification of suitable internal control genes for expression studies in *Coffea arabica* under different experimental conditions. *BMC Molecular Biology* **10:** 1.

Beaudoin N, Serizet C, Gosti F, Giraudat J. 2000. Interactions between abscisic acid and ethylene signaling cascades. *Plant Cell* **12:** 1103–1116.

Berry T, Bewley JD. 1992. A role for the surrounding fruit tissues in preventing the germination of tomato (*Lycopersicon esculentum*) seeds: a consideration of the osmotic environment and abscisic acid. *Plant Physiology* **100:** 951–957.

Bhaskara GB, Nguyen TT, Verslues PE. 2012. Unique drought resistance functions of the highly ABA-induced clade A protein phosphatase 2Cs. *Plant Physiology* 160: 379-395.

Boneh U, Biton I, Schwartz A, Ben-Ari G. 2012a. Characterization of the ABA signal transduction pathway in *Vitis vinifera*. *Plant Science* **187:** 89–96.

Boneh U, Biton I, Zheng C, Schwartz A, Ben-Ari G. 2012b. Characterization of potential ABA receptors in *Vitis vinifera*. *Plant Cell Reports* **31:** 311–321.

Bunn C, Läderach P, Pérez Jimenez JG, Montagnon C, Schilling T. 2015. Multiclass classification of agro-ecological zones for Arabica coffee: an improved understanding of the impacts of climate change. *PLoS ONE* **10:** e0140490.

Carver T, Harris SR, Berriman M, Parkhill J, McQuillan JA. 2012. Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics* **28:** 464–469.

Casacuberta E, González J. 2013. The impact of transposable elements in environmental adaptation. *Molecular Ecology* **22:** 1503–1517.

Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution* **17:** 540–552.

Cutler SR, Rodriguez PL, Finkelstein RR, Abrams SR. 2010. Abscisic acid: emergence of a core signaling network. *Annual Review of Plant Biology* **61:** 651–679.

DaMatta FM, Chaves ARM, Pinheiro HA, Ducatti C, Loureiro ME. 2003. Drought tolerance of two field-grown clones of *Coffea canephora*. *Plant Science* **164:** 111–117.

DaMatta FM, Ramalho JDC. 2006. Impacts of drought and temperature stress on coffee physiology and production: a review. *Brazilian Journal of Plant Physiology* **18:** 55–81.

Davis AP, Tosh J, Ruch N, Fay MF. 2011. Growing coffee: Psilanthus (Rubiaceae) subsumed on the basis of molecular and morphological data; implications for the size, morphology, distribution and evolutionary history of *Coffea*. *Botanical Journal of the Linnean Society* **167:** 357–377.

Davis AP, Gole TW, Baena S, Moat J. 2012. The impact of climate change on indigenous Arabica coffee (*Coffea arabica*): predicting future trends and identifying priorities. *PLoS ONE* **7:** e47981.

Denoeud F, Carretero-Paulet L, Dereeper A, Droc G, Guyot R, Pietrella M, Zheng C, Alberti A, Anthony F, Aprea G, *et al.* 2014. The coffee genome provides insight into the convergent evolution of caffeine

biosynthesis. *Science* **345:** 1181–1184.

Dereeper A, Bocs S, Rouard M, Guignon V, Ravel S, Tranchant-Dubreuil C, Poncet V, Garsmeur O, Lashermes P, Droc G. 2015. The coffee genome hub: a resource for coffee genomes. *Nucleic Acids Research* **43:** D1028–D1035.

Dufayard JF, Duret L, Penel S, Gouy M, Rechenmann F, Perriere G. 2005. Tree pattern matching in phylogenetic trees: automatic search for orthologs or paralogs in homologous gene sequence databases. *Bioinformatics* **21:** 2596–2603.

Fujii H, Zhu JK. 2009. Arabidopsis mutant deficient in 3 abscisic acid-activated protein kinases reveals critical roles in growth, reproduction, and stress. *Proceedings of the National Academy of Sciences, USA* **106:** 8380–8385.

Gonzalez-Guzman M, Rodriguez L, Lorenzo-Orts L, Pons C, Sarrion-Perdigones A, Fernandez MA., Peirats-Llobet M, Forment J, Moreno-Alvero M, Cutler SR, *et al.* 2014. Tomato PYR/PYL/RCAR abscisic acid receptors show high expression in root, differential sensitivity to the abscisic acid agonist quinabactin, and the capability to enhance plant drought resistance. *Journal of Experimental Botany* **65:** 4451–4464.

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology* **59:** 307–321.

Hamanishi ET, Campbell MM. 2011. Genome-wide responses to drought in forest trees. *Forestry* **84:** 273–283.

Hao Q, Yin P, Li W, Wang L, Yan C, Lin Z, Wu JZ, Wang J, Yan SF, Yan N. 2011. The molecular basis of ABA-independent inhibition of PP2Cs by a subclass of PYL proteins. *Molecular Cell* **42:** 662–672.

Harb A, Krishnan A, Ambavaram MMR, Pereira A. 2010. Molecular and physiological analysis of drought stress in Arabidopsis reveals early responses leading to acclimation in plant growth. *Plant Physiology* **154:** 1254–1271.

Hauser F, Waadt R, Schroeder JI. 2011. Evolution of abscisic acid synthesis and signaling mechanisms. *Current Biology* **21:** R346–R355.

Huai J, Wang M, He J, Zheng J, Dong Z, Lv H, Zhao J, Wang G. 2008. Cloning and characterization of the *SnRK2* gene family from *Zea mays*. *Plant Cell Reports* **27:** 1861–1868.

Hubbard KE, Nishimura N, Hitomi K, Getzoff ED, Schroeder JI. 2010. Early abscisic acid signal transduction mechanisms: newly discovered components and newly emerging questions. *Genes & Development* **24:** 1695–1708.

Huson DH, Richter DC, Rausch C, Dezulian T, Franz M, Rupp R. 2007. Dendroscope: an interactive viewer for large phylogenetic trees. *BMC Bioinformatics* **8:** 460.

Jaillon O, Aury JM, Noel B, Policriti A, Clepet C, Casagrande A, Choisne N, Aubourg S, Vitulo N, Jubin C, *et al.* 2007. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature* **449:** 463–467.

Jaramillo J, Chabi-Olaye A, Kamonjo A, Vega FE, Poehling H-M, Borgemeister C. 2009. Thermal tolerance of the Coffee Berry Borer *Hypothenemus hampei*: predictions of climate change impact on a tropical insect pest. *PLoS ONE* **4:** e6487.

Katoh K, Toh H. 2008. Recent developments in the MAFFT multiple sequence alignment program. *Briefings in Bioinformatics* **9:** 286–298.

Kim H, Hwang H, Hong JW, Lee YN, Ahn IP, Yoon IS, Yoo SD, Lee S, Lee SC, Kim BG. **2012.** A rice orthologue of the ABA receptor, OsPYL/RCAR5, is a positive regulator of the ABA signal transduction pathway in seed germination and early seedling growth. *Journal of Experimental Botany* **63:** 1013–1024.

Klingler JP, Batelli G, Zhu JK. **2010.** ABA receptors: the START of a new paradigm in phytohormone signalling. *Journal of Experimental Botany* **61:** 3199–3210.

Kobayashi Y, Yamamoto S, Minami H, Kagaya Y, Hattori T. **2004.** Differential activation of the rice sucrose nonfermenting1–related protein kinase2 family by hyperosmotic stress and abscisic acid. *Plant Cell* **16:** 1163–1177.

Kulik A, Wawer I, Krzywińska E, Bucholc M, Dobrowolska G. 2011. SnRK2 protein kinases-key regulators of plant response to abiotic stresses. *OMICS A Journal of Integrative Biology* **12:** 859–872.

Lashermes P, Combes MC, Robert J, Trouslot P, D'Hont A, Anthony F, Charrier A. **1999**. Molecular characterisation and origin of the *Coffea arabic*a L. genome. *Molecular and General Genetics* **261:** 259–266.

Lashermes P, Andrade AC, Etienne H. 2008. Genomics of coffee, one of the world's largest traded commodities. In: Moore H, Ming R, eds. *Genomics of tropical crop plants*. Berlin, D: Springer, 203–226.

Leung J, Merlot S, Giraudat J. **1997.** The Arabidopsis ABSCISIC ACID-INSENSITIVE2 (ABI2) and ABI1 genes encode homologous protein phosphatases 2C involved in abscisic acid signal transduction. *Plant Cell* **9:** 759–771.

Lima ALS, DaMatta FM, Pinheiro HA, Totola MR, Loureiro ME. 2002. Photochemical responses and oxidative stress in two clones of *Coffea canephora* under water stress. *Environmental and Experimental Botany* **47:** 239–247.

Lima EA, Furlanetto C, Nicole M, Gomes AC, Almeida MR, Jorge-Júnior A, Correa VR, Salgado SM, Ferrão MA, Carneiro RM. 2015. The multi-resistant reaction of drought-tolerant coffee 'Conilon clone 14' to *Meloidogyne* spp. and late hypersensitive-like response in *Coffea canephora*. *Phytopathology* **105:** 805–814.

Lopes FR, Jjingo D, da Silva CRM, Andrade AC, Marraccini P, Teixeira JB, Carazzolle MF, Pereira GAG, Pereira LFP, Vanzela ALL, *et al.* 2013. Transcriptional activity, chromosomal distribution and expression effects of transposable elements in *Coffea* genomes. *PLoS ONE* **8:** e78931.

Ma Y, Szostkiewicz I, Korte A, Moes D, Yang Y, Christmann A, Grill E. **2009.** Regulators of PP2C phosphatase activity function as abscisic acid sensors. *Science* **324:** 1064–1068.

Magrach A, Ghazoul J. 2015. Climate and pest-driven geographic shifts in global coffee production: implications for forest cover, biodiversity and carbon storage. *PLoS ONE* **10:** e0133071.

Marraccini P, Freire LP, Alves GSC, Vieira NG, Vinecky F, Elbelt S, Ramos HJO, Montagnon C, Vieira LGE, Leroy T, Pot D, Silva VA, Rodrigues GC, Andrade AC. **2011.** *RBCS1* expression in coffee: *Coffea* orthologs, *Coffea arabica* homeologs , and expression variability between genotypes and under drought stress. *BMC Plant Biology* **11:** 85.

Marraccini P, Vinecky F, Alves GSC, Ramos HJO, Elbelt S, Vieira NG, Carneiro FA, Sujii PS, Alekcevetch JC, Silva VA, *et al.* 2012. Differentially expressed genes and proteins upon drought acclimation in tolerant and sensitive genotypes of *Coffea canephora*. *Journal of Experimental Botany* **63:** 4191–4212.

Merlot S, Gosti F, Guerrier D, Vavasseur A, Giraudat J. **2001.** The ABI1 and ABI2 protein phosphatases 2C act in a negative feedback regulatory loop of the Abscisic acid signalling pathway. *Plant Journal* **25:** 295–303.

Mizoguchi M, Umezawa T, Nakashima K, Kidokoro S, Takasaki H, Fujita Y, Yamaguchi-Shinozaki K, Shinozaki K. 2010. Two closely related subclass II SnRK2 protein kinases cooperatively regulate drought inducible gene expression. *Plant and Cell Physiology* **51:** 842–847.

Mofatto LS, Carneiro F A, Vieira NG, Duarte KE, Vidal RO, Alekcevetch JC, Cotta MG, Verdeil JL, Lapeyre-Montes F, Lartaud M, *et al.* 2016. Identification of candidate genes for drought tolerance in coffee by high-throughput sequencing in the shoot apex of different *Coffea arabica* cultivars. *BMC Plant Biology* **16:** 94.

Montagnon C, Cubry P, Leroy T. 2012. Amélioration génétique du caféier *Coffea canephora* Pierre: connaissances acquises, stratégies et perspectives. *Cahiers Agricultures* **21:** 143–153.

Montagnon C, Leroy T. 1993. Réaction à la sécheresse de jeunes caféiers *Coffea canephora* de Côte-d'Ivoire appartenant à différents groupes génétiques. *Café Cacao Thé* **37:** 179–190.

Nishimura N, Yoshida T, Murayama M, Asami T, Shinozaki K, Hirayama T. 2004. Isolation and characterization of novel mutants affecting the abscisic acid sensitivity of Arabidopsis germination and seedling growth. *Plant and Cell Physiology* **45:** 1485–1499.

Ovalle-Rivera O, Läderach P, Bunn C, Obersteiner M, Schroth G. 2015. Projected shifts in *Coffea arabica* suitability among major global producing regions due to climate change. *PLoS ONE* **10:** e0124155.

Park SY, Fung P, Nishimura N, Jensen DR, Fujii H, Zhao Y, Lumba S, Santiago J, Rodrigues A, Chow TF, *et al.* 2009. Abscisic acid inhibits PP2Cs via the PYR/PYL family of ABA-binding START proteins. *Science* **324:** 1068–1071.

Pinheiro HA, DaMatta FM, Chaves ARM, Loureiro ME, Ducatti C. 2005. Drought tolerance is associated with rooting depth and stomatal control of water use in clones of *Coffea canephora*. *Annals of Botany* **96:** 101–108.

Pinheiro HA, Fontes EPB, Loureiro ME. 2004. Drought tolerance in relation to protection against oxidative stress in clones of *Coffea canephora* subjected to long-term drought. *Plant Science* **167**: 1307–1314.

Praxedes SC, DaMatta FM, Loureiro ME, Ferrão MAG, Cordeiro AT. 2005. Effects of long-term soil drought on photosynthesis and carbohydrate metabolism in mature robusta coffee (*Coffea canephora* Pierre var . *kouillou*) leaves. *Environmental and Experimental Botany* **56:** 263–273.

Romero P, Lafuente MT, Rodrigo MJ. 2012. The Citrus ABA signalosome: Identification and transcriptional regulation during sweet orange fruit ripening and leaf dehydration. *Journal of Experimental Botany* **63:** 4931–4945.

Saez A, Robert N, Maktabi MH, Schroeder JI, Serrano R, Rodriguez PL. 2006. Enhancement of abscisic acid sensitivity and reduction of water consumption in Arabidopsis by combined inactivation of the protein phosphatases type 2C ABI1 and HAB1. *Plant Physiology* **141:** 1389–1399.

Silva EA, Mazzafera P, Brunini O, Sakai E, Arruda FB, Mattoso LHC, Carvalho CRL, Pires RCM. 2005. The influence of water management and environmental conditions on the chemical composition and beverage quality of coffee beans. *Brazilian Journal of Plant Physiology* **17:** 229–238.

Sun L, Wang YP, Chen P, Ren J, Ji K, Li Q, Li P, Dai SJ, Leng P. 2011. Transcriptional regulation of *SlPYL*, *SlPP2C*, and *SlSnRK2* gene families encoding ABA signal core components during tomato fruit development and drought stress. *Journal of Experimental Botany* **62:** 5659–5669.

Tähtiharju S, Palva T. 2001. Antisense inhibition of protein phosphatase 2C accelerates cold acclimation in

Arabidopsis thaliana. *Plant Journal* **26:** 461–470.

Thompson AJ, Mulholland BJ, Jackson AC, Mckee JMT, Hilton HW, Symonds RC, Sonneveld T, Burbidge A, Stevenson P, Taylor IB. **2007.** Regulation and manipulation of ABA biosynthesis in roots. *Plant, Cell and Environment* **30:** 67–78.

Vieira NG, Carneiro FA, Sujii PS, Alekcevetch JC, Freire LP, Vinecky F, Elbelt S, Silva VA, DaMatta FM, Ferrão MAG, *et al.* 2013. Different molecular mechanisms account for drought tolerance in *Coffea canephora* var. Conilon. *Tropical Plant Biology* **6:** 181–190.

Vinecky F, Davrieux F, Mera AC, Alves GSC, Lavagnini V, Leroy T, Bonnot F, Rocha OC, Bartholo G, Guerra A, *et al.* 2016. Controlled irrigation and NPK fertilization affect the quality of coffee (*Coffea arabica* L. cv. Rubi) beans through their biochemical composition. *Journal of Agricultural Science* in press.

Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X, Lee T, Jin H, Marler B, Guo H, *et al.* 2012. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Research* **40:** e49.

Xue T, Wang D, Zhang S, Ehlting J, Ni F, Jakab S, Zheng C, Zhong Y. **2008.** Genome-wide and expression analysis of protein phosphatase 2C in rice and Arabidopsis. *BMC Genomics* **9:** 550.

Yoshida R, Umezawa T, Mizoguchi T, Takahashi S, Takahashi F, Shinozaki K. **2006.** The regulatory domain of SRK2E/OST1/SnRK2.6 interacts with ABI1 and integrates abscisic acid (ABA) and osmotic stress signals controlling stomatal closure in *Arabidopsis*. *Journal of Biological Chemistry* **281:** 5310–5318.

Yoshida T, Fujita Y, Sayama H, Kidokoro S, Maruyama K, Mizoi J, Shinozaki K, Yamaguchi-Shinozaki K. 2010. AREB1, AREB2, and ABF3 are master transcription factors that cooperatively regulate ABRE-dependent ABA signaling involved in drought stress tolerance and require ABA for full activation. *Plant Journal* **61:** 672–685.

Zhang H, Mao X, Wang C, Jing R. **2010.** Overexpression of a common wheat gene *TaSnRK2.8* enhances tolerance to drought, salt and low temperature in Arabidopsis. *PLoS ONE* **5:** e16041.

Zhang H, Mao X, Jing R, Chang X, Xie H. 2011. Characterization of a common wheat (*Triticum aestivum* L.) *TaSnRK2.7* gene involved in abiotic stress reponses. *Journal of Experimental Botany* **62:** 975–988.

Zhao Y, Xing L, Wang X, Hou YJ, Gao J, Wang P, Duan CG, Zhu X, Zhu JK. **2014.** The ABA receptor PYL8 promotes lateral root growth by enhancing MYB77-dependent transcription of auxin-responsive genes. *Science Signaling* **7:** ra53–ra53.

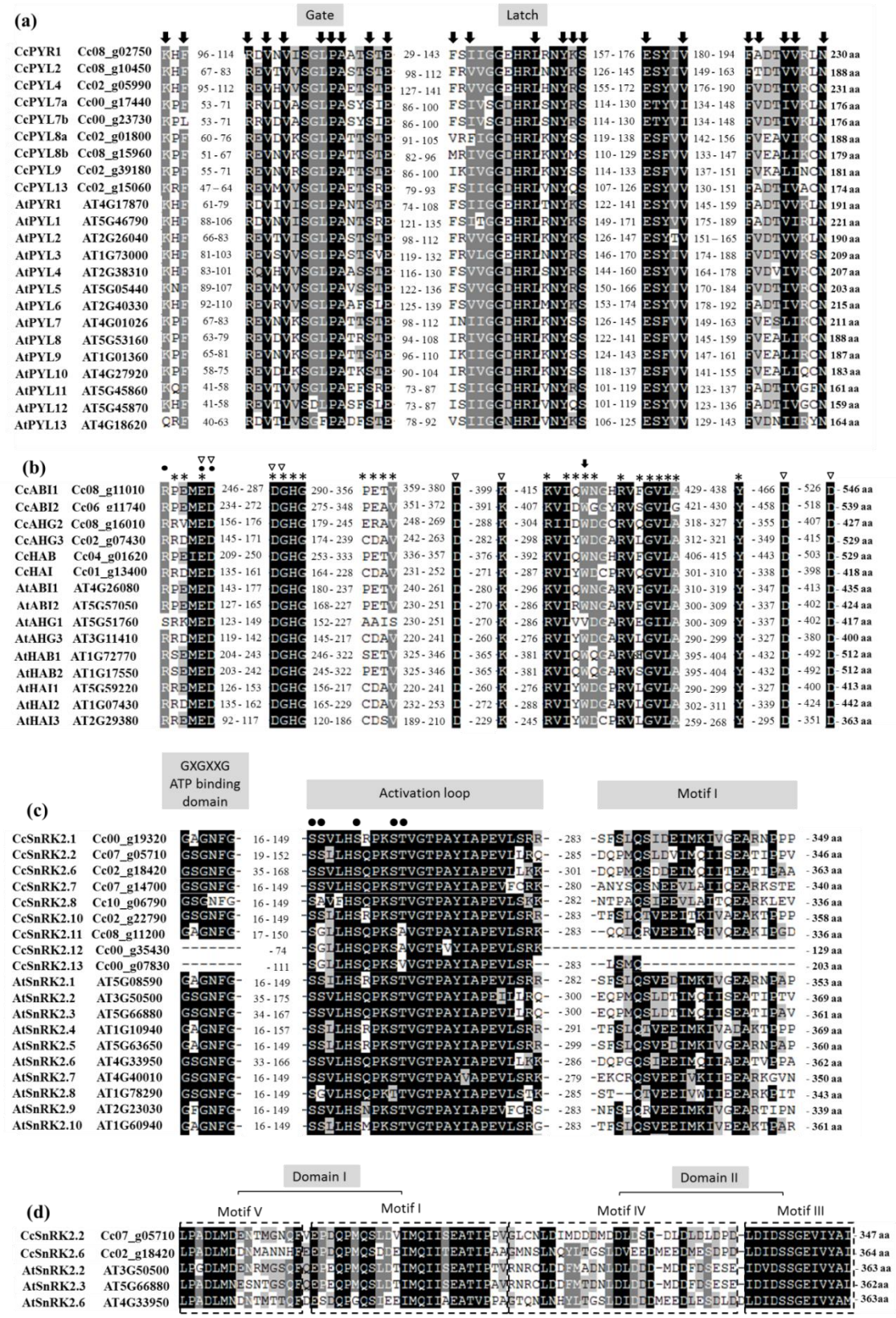


Figure 1 Sequence alignments of the PYL, PP2C and SnRK2 putative proteins. Amino acid sequences are shown only for functional residues and conserved domains. For each protein, total length is indicated in amino acids (aa). Conserved residues are marked with black or grey shading. (a): sequence alignment of the PYL proteins. Residues forming the ligand-binding pocket are marked by black arrows. The gate and latch domains are indicated. (b):

sequence alignment of the PP2C proteins. Residues interacting with ABA, PYLs and Mn^{2+}/Mg^{2+} ions are marked by black arrows, asterisks, and white triangles, respectively. Phosphatase sites are marked with black points. (c): sequence alignment of C-terminal regions of subclass III SnRK2s. Functional domains (ATP binding site, activation loop and motif I) are indicated. (d): sequence alignment of C-terminal regions of subclass III SnRK2s. Functional domains (domains I and II with their corresponding motifs) are indicated.

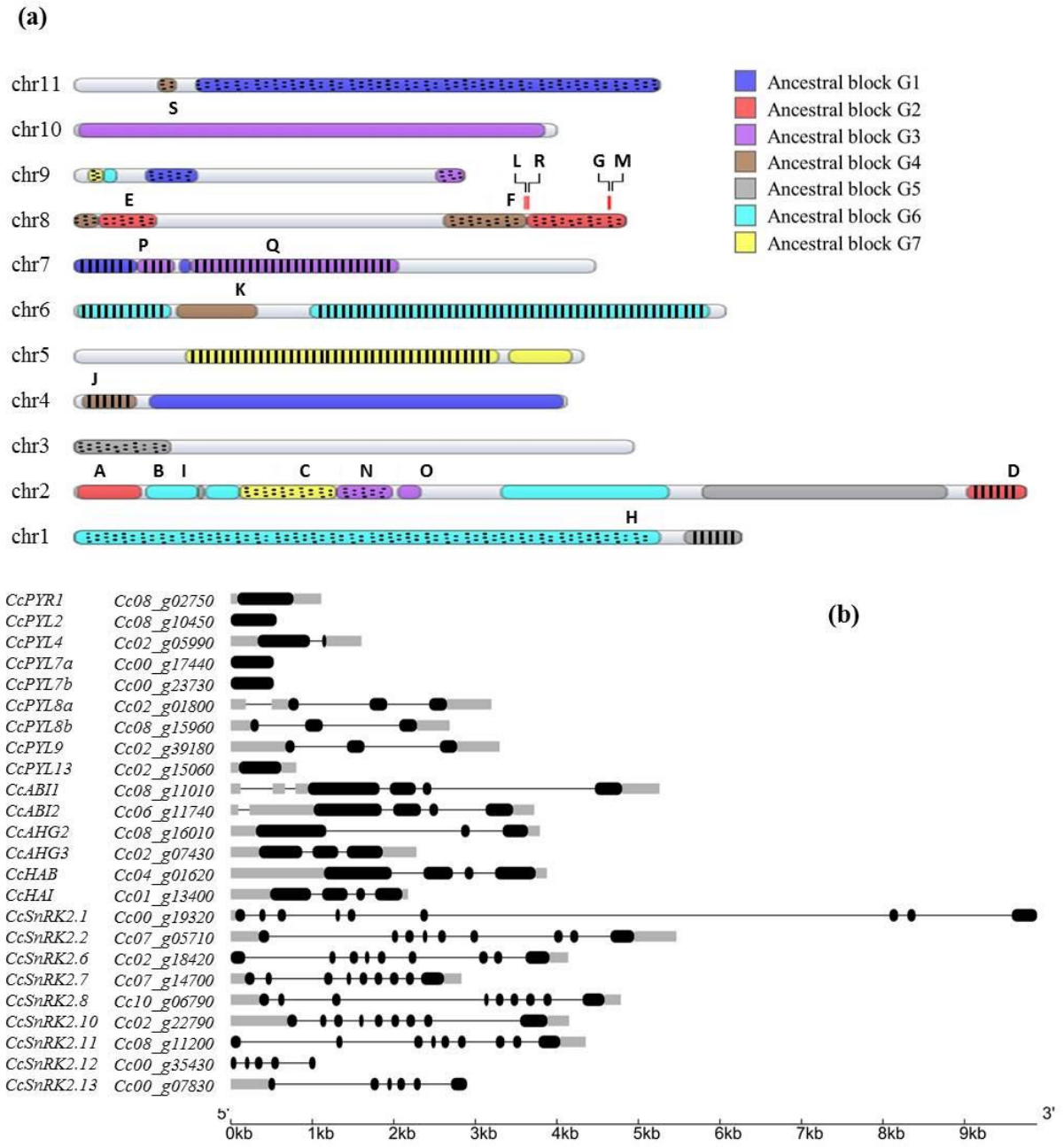


Figure 2 Localization and structure of *PYR/PYL/RCAR*, *PP2C* and *SnRK2* genes. (a): localization of genes in *C. canephora* chromosomes. *CcPYL* genes: *CcPYL8a* (A), *CcPYL4* (B), *CcPYL13* (C), *CcPYL9* (D), *CcPYR1* (E), *CcPYL2* (F) and *CcPYL8b* (G). *CcPP2C* genes: *CcHAI* (H), *CcAHG3* (I), *CcHAB* (J), *CcABI2* (K), *CcABI1* (L) and *CcAHG2* (M). *CcSnRK2* genes: *CcSnRK2.6* (N), *CcSnRK2.10* (O), *CcSnRK2.2* (P), *CcSnRK2.7* (Q), *CcSnRK2.11* (R) and *CcSnRK2.8* (S). The *PYLs* (*CcPYL7a* and *CcPYL7b*, and *SnRK2 (CcSnRK2.13)*, *CcSnRK2.1* and *CcSnRK2.12* genes unanchored in the chromosome 0 and are not indicated. The coloured regions represent the ancestral blocks of the 7 core eudicot chromosomes (adapted from Denoeud *et al.* [2014]). (b): structure of *CcPYL*, *CcPP2C* and *CcSnRK2* genes. The black blocks represent exons, the gray blocks the upstream and downstream transcribed and untranslated regions (UTRs) and the lines the introns. The structure of genes located in the chromosome 0 is not represented. For the *CcPYL2*, *CcPYL7a* and *CcPYL7b* genes, no 5' and 3'UTRs were found.

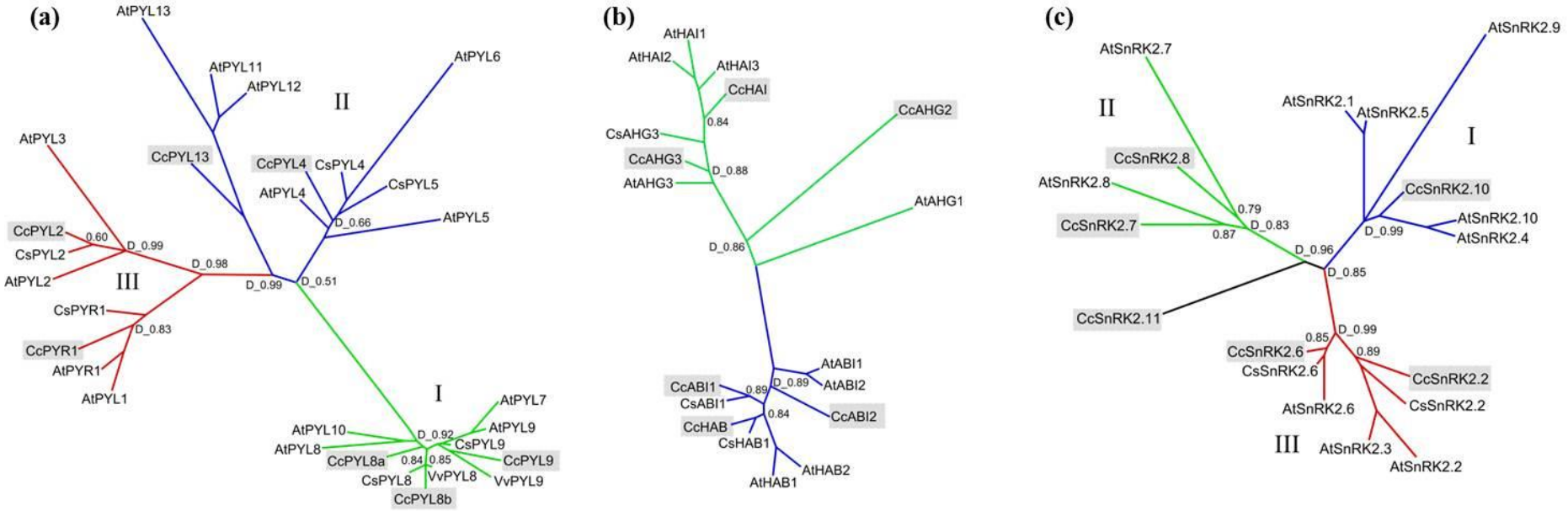


Figure 3 Phylogenetic analyses of *C. canephora* PYR/PYL (a), clade-A PP2C (b) and SnRK2 (c) proteins. Trees were constructed using amino proteins of *C. canephora* and orthologous proteins from *A. thaliana* (At), *C. sinensis* (Cs) and *V. vinifera* (Vv) (see Tables S2-S4 and Fig. S2-S4). The coffee proteins are highlighted in gray. The proteins coded by genes located in the chromosome 0 are not included. For PYR/PYL and SnRK2 trees, protein subclasses are also indicated.

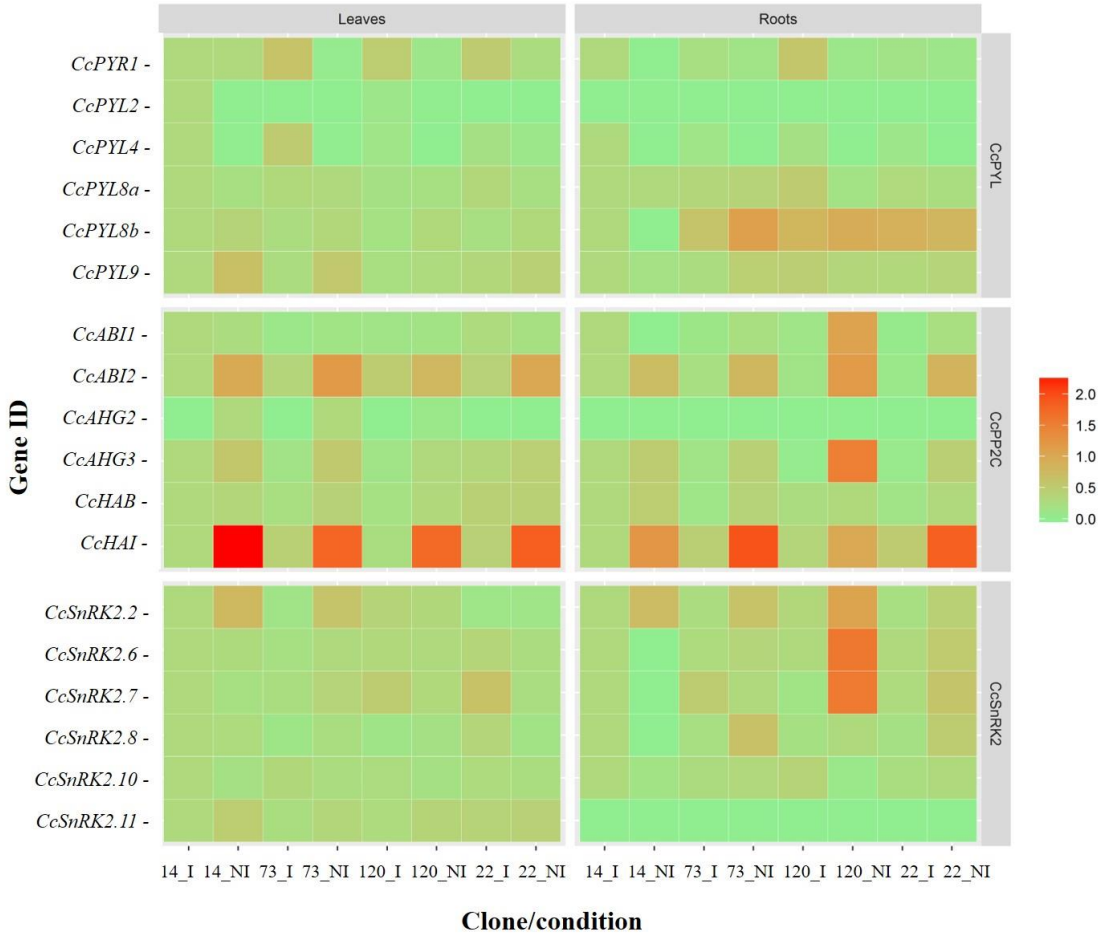


Figure 4 Expression profiles of *PYR/PYL*, *PP2C* and *SnRK2* genes in leaves and roots of D^T (14, 73 and 120) and D^S (22) clones of *C. canephora* subjected (NI) or not (I) to drought. The gene names are indicated in the heatmap. Values are the mean of at least three technical repetitions ± SD which are standardized independently with *CcUBQ10* (ubiquitin) as reference gene. Results are expressed using 14I as an internal calibrator (RE=1), except for *CcAHG2* gene where 14NI was used. Higher expression for each gene was presented in red, otherwise, green was used.

.

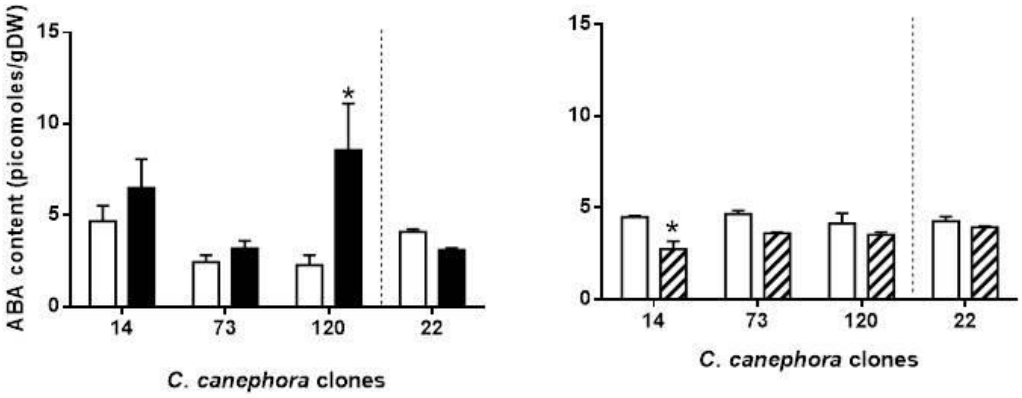


Figure 5 ABA content of leaves and roots of D^T (14, 73 and 120) and D^S (22) clones of *C. canephora* subjected (NI) or not (I: white isobars) to drought. Black and striped isobars corresponded to drought conditions in leaves and roots, respectively**.** For the statistical analysis, significant differences ($P \leq 0.05$) between the treatments were evaluated using 2way ANOVA test (non-parametric test) and are indicated by an asterisk.

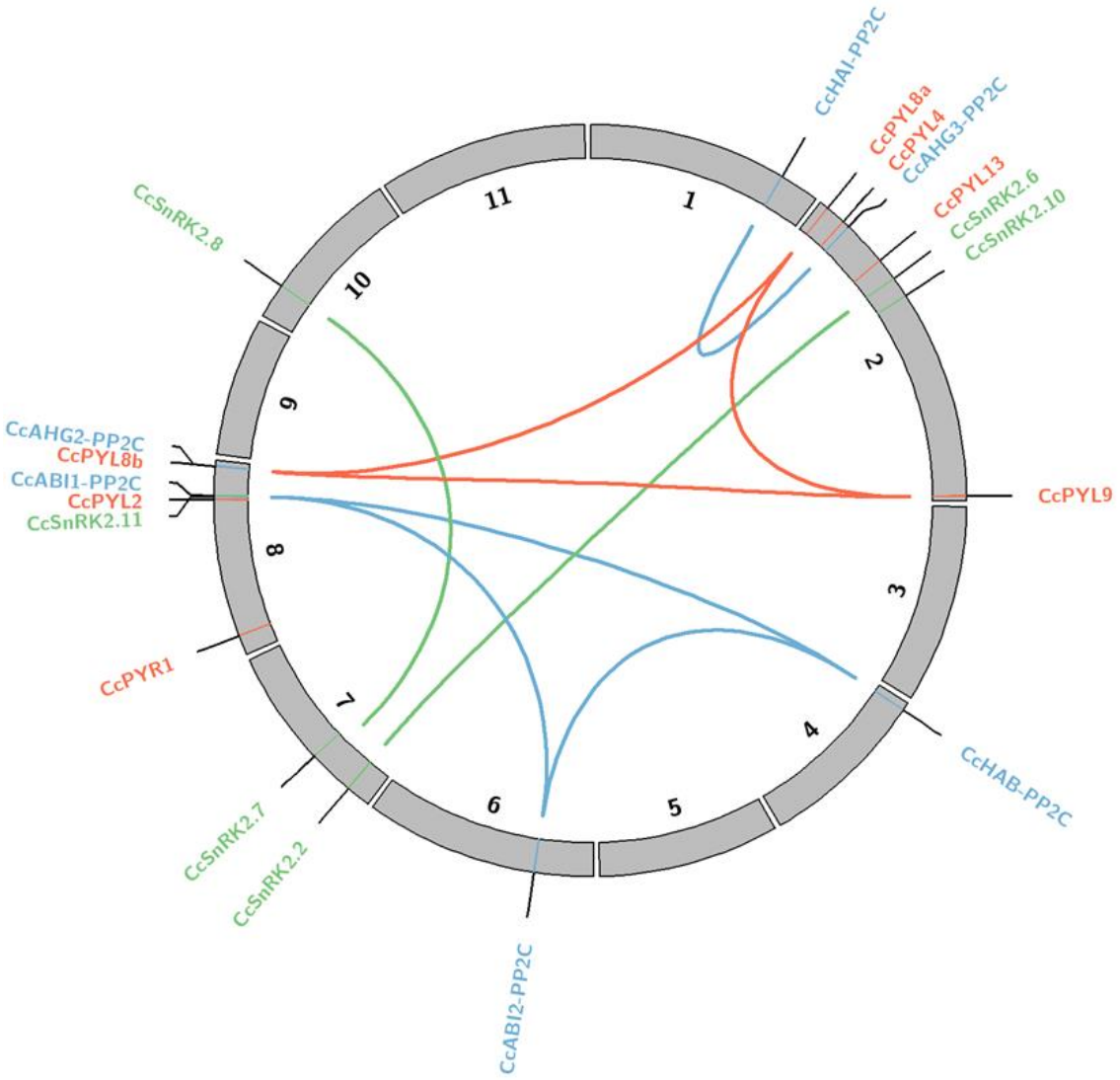
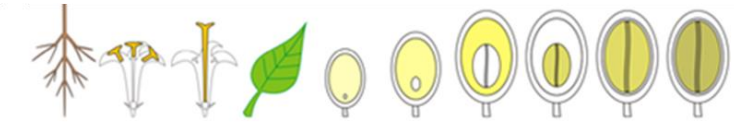


Figure 6 Graphical representation of the *CcPYL-CcPP2C-CcSnRK2* duplicated genes on *C. canephora* chromosomes (indicated by numbers, from 1 to 11). The *CcPYL*, *CcPP2C* and *CcSnRK2* duplications genes are indicated by with red, blue and green lines, respectively. The *CcPYL8a*, *CcPYL8b*, *CcPYL9*, *CcABI1*, *CcABI2* and *CcHAB* as well as *CcSnRK2.2* and *CcSnRK2.6*, evolved through proximal duplications. The genes located on the chromosome 0 are not showed.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information tab for this article:



(a)	***CcPYLs***									
CcPYR1	*Cc08_g02750*	71	2	12	9	1	2	3	9	10	82
CcPYL2	*Cc08_g10450*	0	0	5	0	0	1	1	30	14	6
CcPYL4	*Cc02_g05990*	36	103	61	4	128	171	168	21	16	223
CcPYL7a	*Cc00_g17440*	0	1	0	0	0	0	0	0	0	1
CcPYL7b	*Cc00_g23730*	0	19	0	0	0	0	0	1	0	1
CcPYL8a	*Cc02_g01800*	63	13	56	31	28	30	81	71	54	169
CcPYL8b	*Cc08_g15960*	123	21	71	103	152	93	122	40	61	116
CcPYL9	*Cc02_g39180*	243	28	84	142	169	89	140	41	42	65
CcPYL13	*Cc02_g15060*	0	0	2	0	3	6	18	36	67	129
(b)	***CcPP2Cs***										
CcABI1	*Cc08_g11010*	195	19	46	214	36	37	14	14	24	21
CcABI2	*Cc06_g11740*	112	39	48	100	28	35	22	17	45	62
CcAHG2	*Cc08_g16010*	4	2	1	6	0	1	2	61	82	164
CcAHG3	*Cc02_g07430*	261	74	214	227	17	9	6	82	124	208
CcHAB	*Cc04_g01620*	82	12	63	78	23	23	17	11	32	47
CcHAI	*Cc01_g13400*	103	16	151	156	39	58	40	108	150	135
(c)	***CcSnRK2s***										
CcSnRK2.1	*Cc00_g19320*	0	139	0	0	0	0	0	0	0	0
CcSnRK2.2	*Cc07_g05710*	219	28	16	18	58	44	47	27	25	136
CcSnRK2.6	*Cc02_g18420*	86	16	65	52	33	24	15	12	51	33
CcSnRK2.7	*Cc07_g14700*	85	97	10	93	32	23	18	6	3	2
CcSnRK2.8	*Cc10_g06790*	38	13	9	34	28	15	9	22	37	36
CcSnRK2.10	*Cc02_g22790*	93	38	82	74	80	89	45	70	63	46
CcSnRK2.11	*Cc08_g11200*	12	0	1	2	3	2	2	22	33	41
CcSnRK2.12	*Cc00_g35430*	0	0	2	0	3	2	1	1	2	1
CcSnRK2.13	*Cc00_g07830*	3	3	8	2	11	6	3	4	6	4

Figure S1 Heat map visualization of the *CcPYR/PYL* (a), *CcPP2C* (b) and *CcSnRK2* (c) gene families. From left to right, the libraries correspond to root, stamen, pistil, leaf, perisperm (120, 150 and 180 days after pollination-DAP) and endosperm (180, 260 and 320 DAP) from *C. canephora* RNA-Seq data. Transcript abundance was normalized with RPKM and the level of gene expression is indicated with a colour scale, from white (weakly expressed) to red (strongly expressed) (adapted from: http://www.coffee-genome.org/).

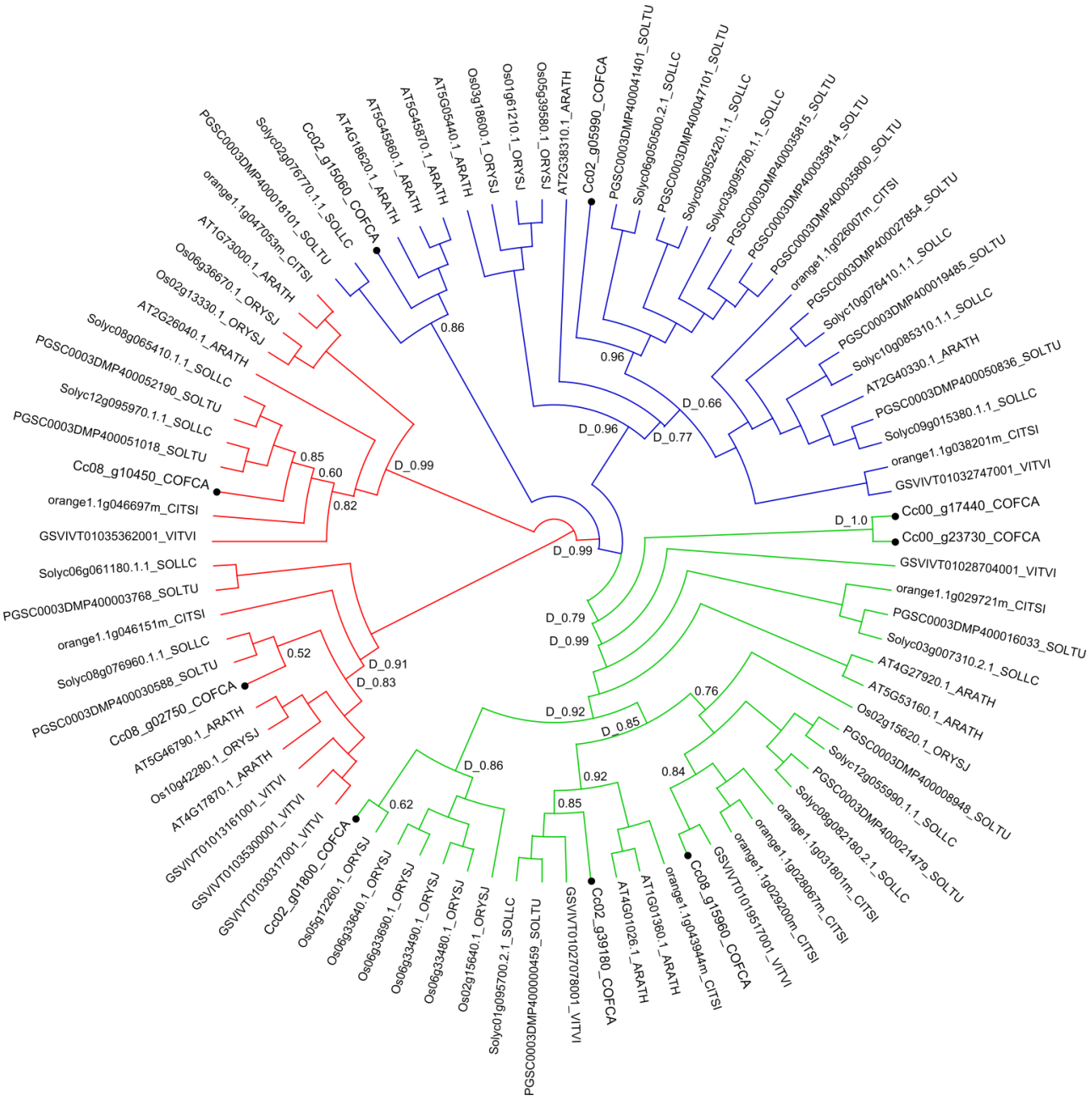


Figure S2 Phylogenetic analysis of CcPYL protein sequences with orthologous proteins of *A. thaliana* (ARATH), *C. sinensis* (CITSI), *O. sativa* (ORYSJ), *S. lycopersicum* (SOLLC), *S. tuberosum* (SOLTU) and *V. vinifera* (VITVI). The phylo-HMM approach was based on NNI (Nearest Neighbor Interchange) topology. Subfamilies I (green), II (blue) and III (red) are indicated. Main bootstraps values are indicated.

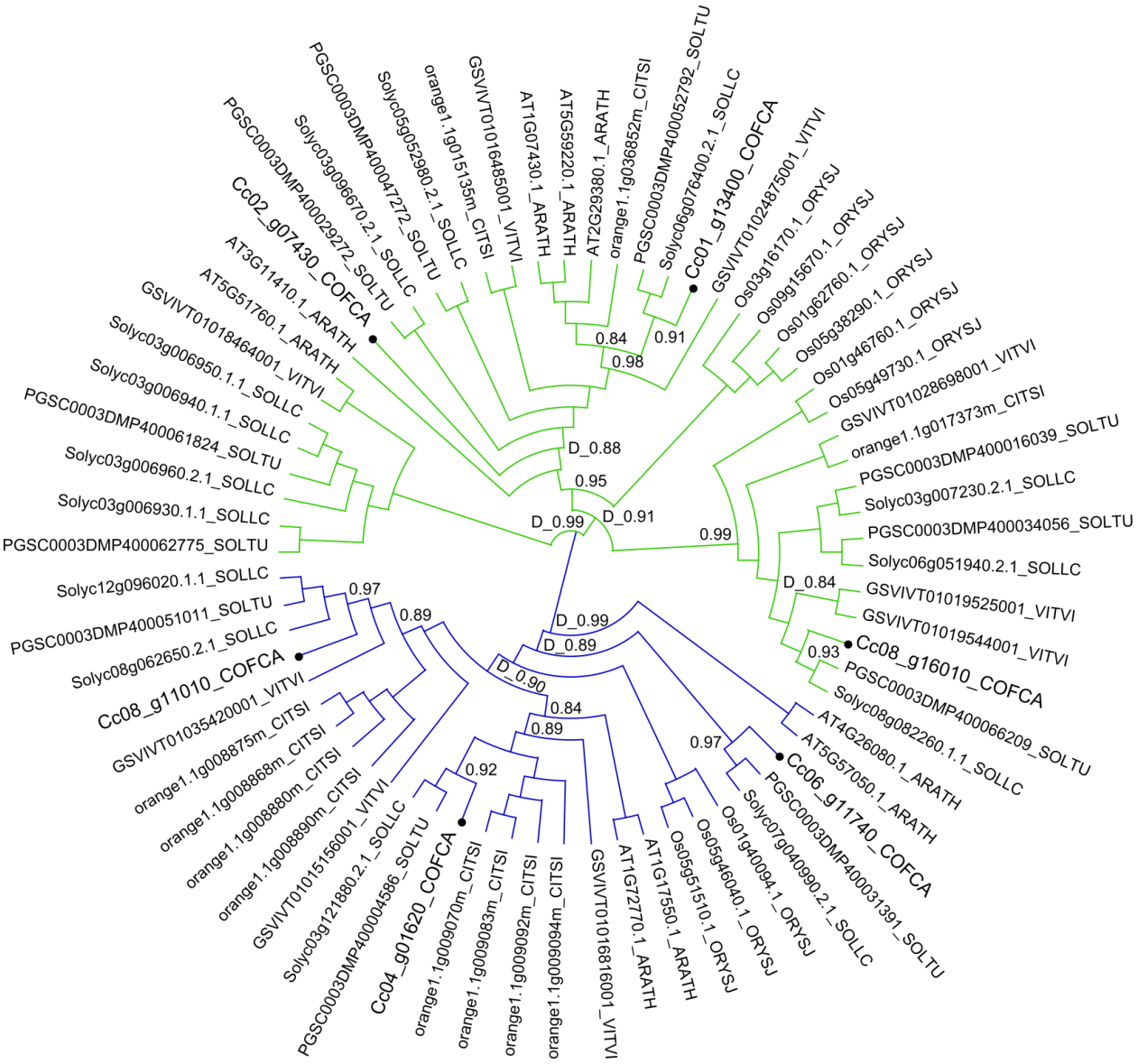


Figure S3 Phylogenetic analyses of PP2C protein sequences with orthologous proteins of *A. thaliana* (ARATH), *C. sinensis* (CITSI), *O. sativa* (ORYSJ), *S. lycopersicum* (SOLLC), *S. tuberosum* (SOLTU) *and V. vinifera* (VITVI). The phylo-HMM approach was based on NNI (Nearest Neighbor Interchange) topology. Subfamilies are indicated with different colours. Main bootstraps values are indicated.

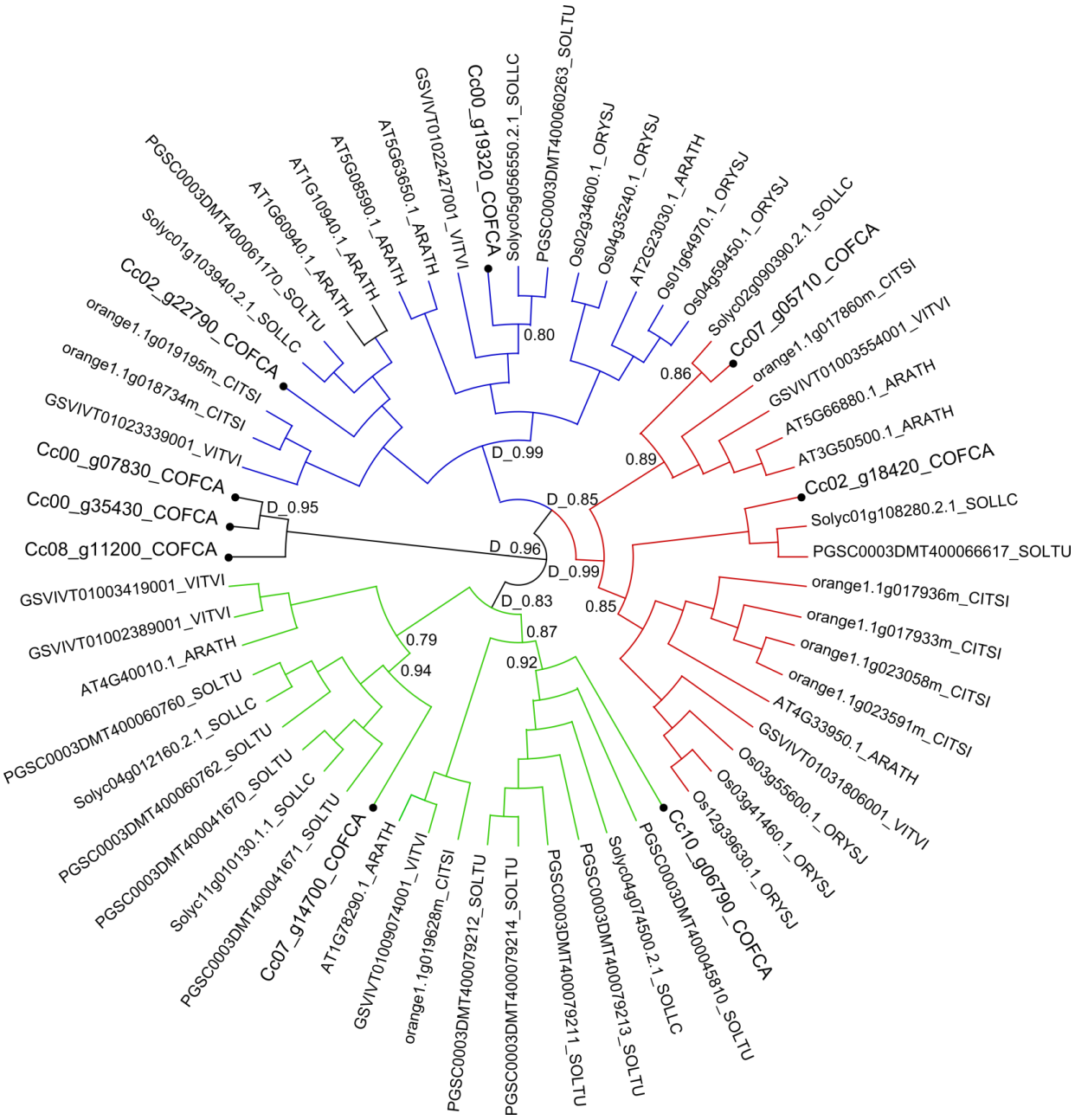


Figure S4 Phylogenetic analyses of SnRK2 protein sequences with orthologous proteins of *A. thaliana* (ARATH), *C. sinensis* (CITSI), *O. sativa* (ORYSJ), *S. lycopersicum* (SOLLC), *S. tuberosum* (SOLTU) *and V. vinifera* (VITVI. The phylo-HMM approach was based on NNI (Nearest Neighbor Interchange) topology. Subfamilies I (green), II (blue) and III (red) are indicated. Main bootstraps values are indicated.

Table 1 Candidate genes and corresponding primers used for qPCR experiments. Pairs of primers were designed for each gene using the Primer Express software (Applied Biosystems). The primers select to qPCR experiments F (Forward) and R (Reverse) are indicated. For *Cc02_g05990* and *Cc10_g06790* genes two different pairs of primers were used in each tissue, F1R1 (leaves) and F2R2 (roots). The *CcUBQ10*F and R primer par was used for the ubiquitin (UBI) as reference gene.

Gene Family	Gene Name	Gene ID	Primer sequence 5'-3'
PYR/PYL/RCAR	*CcPYR1*	*Cc08_g02750*	F: CGGTGACGACTGTCCATGAG
			R: TCCGGCACGTCAACGATATA
	CcPYL2	*Cc08_g10450*	F: AAAAGTGGTGTGGCCATTCG
			R: CTTCCATCCCCTGTCATGTTG
	CcPYL4	*Cc02_g05990*	F1: CCTATGCCTTCGTCCCTTCA
			R1: CGCGGAATTGGTGGTTGTAG
			F2: TACCATTGTGCGGTGCAACT
			R2: TTCTGTTCTGGGCTTCCATGA
	CcPYL7a	*Cc00_g17400*	F: GAGCGGCTCGAGACTCTTGA
			R: GCCGCTGACAATGCTGAAC
	CcPYL7b	*Cc00_g23730*	F: GCGGCTGGAGATTCTTGATC
			R: CGCCTGGCCATCTATGATTC
	CcPYL8a	*Cc02_g01800*	F: GGTTTGATCAGCCCCAGAAAT
			R: CCACTTCCCTAAGGCTTCCAA
	CcPYL8b	*Cc08_g15960*	F: GCCAGAGGGAAATACCAAGGA
			R: CAGCTAGGCGCTCTGAGACA
	CcPYL9	*Cc02_g39180*	F: CACCCGTGCTCTTCCTCTGT
			R: TCCTCACCAGTGACCAAACG
	CcPYL13	*Cc02_g15060*	F: TCCCAAACCAGTGCACTTCA
			R: TTGTCGAATTGACGGACCAA
PP2C	*CcABI1*	*Cc08_g11010*	F: TGCTGAGGTTGGAGGGAAAA
			R: CGAACAAACAAGGGCAACAA
	CcABI2	*Cc06_g11740*	F: TACGGCTGTGGTTGGCATTA
			R: CTGCCCTTGAATCACCACAA
	CcAHG2	*Cc08_g16010*	F: AGAGCCCTGCTCCTGGTGAT
			R: GGTCATGCTACCGCGATCTT
	CcAHG3	*Cc02_g07430*	F: ACCGGAGGTGACGATAATCG
			R: CCCACAAGCTGTGTCATTGG
	CcHAB	*Cc04_g01620*	F: TGGCTTGTGGGATGTCATGA
			R: CGTTCTTCTTGTGCCAAAGCA
	CcHAI	*Cc01_g13400*	F: CATCGACGCTGCTTGTCAAT
			R: CCACCGCGTCTTCCATATCT
SnRK2	*CcSnRK2.1*	*Cc00_g19320*	F: TAGCCCCCGAGGTTCTCTCT
			R: TCACTCCGCAAGACCACACA
	CcSnRK2.2	*Cc07_g05710*	F: CGAGGATGAGGCTCGTTTTT
			R: GCTGGGCTTCCGTCTAACAA
	CcSnRK2.6	*Cc02_g18420*	F: GCATATATTGCGCCCGAAGT
			R: AAAGGGTATGCCCCCACAAG
	CcSnRK2.7	*Cc07_g14700*	F: AAGCCCAGAACCACGTCTCA
			R: GATTTGGGTTGGGAATGCAA
	CcSnRK2.8	*Cc10_g06790*	F1: AACATGTGCAGCGGGAGATT
			R1: CTCCTGCCGCATACTCCATT
			F2: CCGCTTCAAAGAGGTCTTGCT
			R2: TTCTCCTCCTGCCGCATACT
	CcSnRK2.10	*Cc02_g22790*	F: TCGATTCAAGGAGGTGGTGTT
			R: TTCCCCTCCAGCTGCATACT
	CcSnRK2.11	*Cc08_g11200*	F: AGGAACCTGACCCTCACCAA
			R: CCTGGGATTTTTGCCTCTTG
	CcSnRK2.12	*Cc00_g35430*	F: ACTTGAAGTTGGAAAACACATTTTTG
			R: GTCAAGGAAGGAATATGATGGGAAG
	CcSnRK2.13	Cc00_g07830	F: GGTGTTAGTTACTGTCATTCAATGGAA
			R: ACTTGAAGCTGGAAAACACACTTTT
	CcUBQ10	*Cc05_g13290*	F: AAGACAGCTTCAACAGAGTACAGCAT
			R: GGCAGGACCTTGGCTGACTATA

Table S1 Comparison of CcPYL protein sequences with orthologous sequences from *A. thaliana* (At), *C. sinensis* (Cs), *O. sativa* (Os), *S. lycopersicum* (Sl), *S. tuberosum* (St) and *V. vinifera* (Vv) plant species (Sp). NCBI accession numbers (ID), I (introns), aa (amino acid length), Ident. (Identity), Align. (Match/Aligned), QC (Query Cover), e-value and function information were obtained through BLASTp results. Coffee IDs were identified in Coffee Genome Hub (http://www.coffee-genome.org/) and Gene IDs in Phytozome 10.3 (http://phytozome.jgi.doe.gov/pz/portal.html).

Coffee ID	NCBI ID	*Gene ID*	Sp	I	aa	Ident.	Align.	QC	e-value	Function
	AAD25950.1	*AT2G40330.1*	*At*	0	175	47%	69%	91%	2,21E-47	hypothetical protein
	KDO68852.1	*orange1.1g038201m*	*Cs*	0	201	49%	69%	97%	2,53E-49	hypothetical protein CISIN
	EEE62745.1	*LOC_Os05g12260.1*	*Os*	0	196	48%	69%	84%	9,3E-42	hypothetical protein OsJ_17548
Cc00_g17440	XP_004249065.1	*Solyc10g076410.1.1*	*Sl*	0	203	52%	71%	92%	4,02E-50	PREDICTED: abscisic acid receptor PYL4-like
	XP_006359557.1	*PGSC0003DMG400029194*	*St*	0	214	52%	72%	92%	7,21E-51	PREDICTED: abscisic acid receptor PYL4-like
	XP_002264158.1	GSVIVG01032747001	*Vv*	0	227	54%	72%	95%	7,95E-54	PREDICTED: abscisic acid receptor PYL4
	AAD25950.1	*AT2G40330.1*	*At*	0	175	46%	68%	93%	6,88E-48	hypothetical protein
	KDO68852.1	*orange1.1g038201m*	*Cs*	0	201	48%	68%	97%	3,42E-48	hypothetical protein CISIN
	EEE62745.1	*LOC_Os05g12260.1*	*Os*	0	196	47%	68%	84%	1,57E-39	hypothetical protein OsJ_17548
Cc00_g23730	XP_004249671.1	*Solyc10g085310.1.1*	*Sl*	0	213	50%	71%	93%	3,9E-49	PREDICTED: abscisic acid receptor PYL4
	XP_006359557.1	*PGSC0003DMG400029194*	*St*	0	214	51%	71%	92%	5,34E-50	PREDICTED: abscisic acid receptor PYL4-like
	XP_002264158.1	GSVIVG01032747001	*Vv*	0	227	53%	71%	94%	3,01E-52	PREDICTED: abscisic acid receptor PYL4
	BAF00266.1	*AT5G53160.2*	*At*	2	188	82%	92%	90%	5E-100	hypothetical protein
	XP_006476396.1	orange1.1g028067m	*Cs*	2	197	78%	90%	98%	1,2E-103	PREDICTED: abscisic acid receptor PYL8-like
	NP_001046464.1	*LOC_Os02g15640.1*	*Os*	2	204	79%	92%	94%	2,71E-98	Os02g0255500
Cc02_g01800	XP_004234175.1	*Solyc03g007310.2.1*	*Sl*	2	185	83%	94%	98%	8,7E-111	PREDICTED: abscisic acid receptor PYL8
	NP_001275025.1	*PGSC0003DMG400009108*	*St*	2	185	82%	94%	98%	1,1E-109	abscisic acid receptor PYL8-like
	XP_002270037.3	*GSVIVG01028704001*	*Vv*	2	185	85%	94%	96%	2,1E-109	PREDICTED: abscisic acid receptor PYL8
	NP_565928.1	*AT2G40330.1*	*At*	0	215	57%	71%	91%	9,06E-74	abscisic acid receptor PYL6
	KDO68852.1	*orange1.1g038201m*	*Cs*	0	201	77%	87%	72%	5,75E-86	hypothetical protein CISIN
	NP_001055819.1	*LOC_Os05g39580.1*	*Os*	0	216	65%	75%	71%	4,62E-62	Os05g0473000
Cc02_g05990	XP_004235232.1	*Solyc03g095780.1.1*	*Sl*	0	201	75%	82%	88%	1,4E-99	PREDICTED: abscisic acid receptor PYL4-like
	XP_006353422.1	*PGSC0003DMG400023949*	*St*	0	218	74%	81%	91%	1,3E-104	PREDICTED: abscisic acid receptor PYL4-like
	XP_002264158.1	GSVIVG01032747001	*Vv*	0	227	68%	77%	90%	2,59E-91	PREDICTED: abscisic acid receptor PYL4
	AAD25950.1	*AT2G40330.1*	*At*	0	175	56%	79%	87%	5,89E-54	hypothetical protein
	KDO68852.1	*orange1.1g038201m*	*Cs*	0	201	60%	75%	93%	1,17E-60	hypothetical protein CISIN
	NP_001049838.1	*LOC_Os03g18600.1*	*Os*	0	229	55%	77%	93%	5,45E-52	Os03g0297600
Cc02_g15060	XP_004249671.1	*Solyc10g085310.1.1*	*Sl*	0	213	53%	79%	94%	3,56E-61	PREDICTED: abscisic acid receptor PYL4
	XP_006359557.1	*PGSC0003DMG400029194*	*St*	0	214	57%	79%	93%	1,09E-61	PREDICTED: abscisic acid receptor PYL4-like
	CAN72620.1	GSVIVG01013161001	*Vv*	0	172	62%	76%	97%	4,64E-70	hypothetical protein VITISV_004947

Table S1Continue… for legend see the previous page.

Coffee ID	NCBI ID	*Gene ID*	Sp	I	aa	Ident.	Align.	QC	e-value	Function
	3OQU	*AT1G01360.1*	*At*	2	205	75%	87%	96%	5,82E-87	Abscisic Acid Receptor Pyl9
	XP_006476396.1	*orange1.1g029200m*	*Cs*	2	197	75%	88%	96%	8,42E-90	PREDICTED: abscisic acid receptor PYL8-like
	NP_001054923.1	*LOC_Os05g12260.1*	*Os*	2	209	76%	88%	90%	5,84E-84	Os05g0213500
Cc02_g39180	XP_004231210.1	*Solyc01g095700.2.1*	*Sl*	2	186	83%	95%	91%	4,15E-98	PREDICTED: abscisic acid receptor PYL8
	NP_001284557.1	*PGSC0003DMG400000215*	*St*	2	186	81%	95%	91%	5,95E-96	abscisic acid receptor PYL8-like
	XP_010659134.1	*GSVIVG01019517001*	*Vv*	2	189	76%	89%	96%	6,96E-90	PREDICTED: abscisic acid receptor PYL8-like
	NP_193521.1	*AT4G17870.1*	*At*	0	191	68%	82%	83%	2,37E-87	abscisic acid receptor PYR1
	XP_006491739	*orange1.1g046151m*	*Cs*	0	187	55%	72%	78%	4,97E-66	PREDICTED: abscisic acid receptor PYL1-like
	NP_001065470.1	*LOC_Os10g42280.1*	*Os*	0	212	62%	73%	75%	2,29E-64	Os10g0573400
Cc08_g02750	XP_004245893.1	*Solyc08g076960.1.1*	*Sl*	0	231	72%	80%	97%	3,7E-110	PREDICTED: abscisic acid receptor PYR1-like
	NP_001284559.1	*PGSC0003DMG400017514*	*St*	0	231	70%	80%	97%	1,3E-107	abscisic acid receptor PYL1-like
	XP_002280361.1	*GSVIVG01013161001*	*Vv*	1	214	76%	85%	78%	7,3E-99	PREDICTED: abscisic acid receptor PYR1-like
	NP_180174.1	*AT2G26040.1*	*At*	0	190	72%	82%	95%	7,97E-85	abscisic acid receptor PYL2
	KDO80051.1	*orange1.1g046697m*	*Cs*	0	187	82%	88%	99%	1,4E-106	hypothetical protein CISIN
	NP_001172865.1	*LOC_Os02g13330.1*	*Os*	0	207	61%	74%	90%	2,27E-60	Os02g0226801
Cc08_g10450	XP_004253195.1	*Solyc12g095970.1.1*	*Sl*	1	190	80%	88%	99%	1,1E-103	PREDICTED: abscisic acid receptor PYL2-like
	XP_006360983.1	*PGSC0003DMG400029952*	*St*	1	188	84%	89%	98%	9,3E-107	PREDICTED: abscisic acid receptor PYL2-like
	XP_010648333.1	*GSVIVG01035362001*	*Vv*	0	185	82%	88%	95%	2,8E-99	PREDICTED: abscisic acid receptor PYL2
	NP_200128.1	*AT5G53160.2*	*At*	2	188	80%	91%	96%	2,32E-95	regulatory component of ABA receptor 3
	XP_006476396.1	*orange1.1g029200m*	*Cs*	2	197	87%	95%	97%	2,9E-106	PREDICTED: abscisic acid receptor PYL8-like
	NP_001046464.1	*LOC_Os02g15640.1*	*Os*	2	204	80%	90%	97%	1,92E-95	Os02g0255500
Cc08_g15960	XP_004245523.1	*Solyc08g082180.2.1*	*Sl*	2	189	85%	95%	97%	3E-105	PREDICTED: abscisic acid receptor PYL9
	XP_006343869.1	*PGSC0003DMG400012155*	*St*	2	189	85%	95%	97%	3E-105	PREDICTED: abscisic acid receptor PYL9-like
	XP_010659134.1	*GSVIVG01019517001*	*Vv*	3	189	90%	95%	97%	7,7E-111	PREDICTED: abscisic acid receptor PYL8-like

Table S2 Comparison of CcPP2C protein sequences with orthologous sequences from *A. thaliana* (At), *C. sinensis* (Cs), *O. sativa* (Os), *S. lycopersicum* (Sl), *S. tuberosum* (St) and *V. vinifera* (Vv) plant species (Sp). NCBI accession numbers (ID), I (introns), aa (amino acid length), Ident. (Identity), Align. (Match/Aligned), QC (Query Cover), e-value and function information were obtained through BLASTp results. Coffee IDs were identified in Coffee Genome Hub (http://www.coffee-genome.org/) and Gene IDs in Phytozome 10.3 (http://phytozome.jgi.doe.gov/pz/portal.html).

Coffee ID	NCBI ID	Gene ID	Sp	I	Len.aa	Ident.	Align.	Query cov.	e-value	Function inferred by BLAST
Cc01_g13400	NP_180499.1	*AT2G29380.1*	*At*	2	362	71%	83%	69%	2.3E-139	highly ABA-induced PP2C protein 3
	XP_006488392.1	*orange1.1g036852m*	*Cs*	2	429	64%	75%	98%	1.7E-159	PREDICTED: probable protein phosphatase 2C 78-like
	NP_001044788.1	*LOC_Os01g62760.1*	*Os*	3	414	55%	68%	70%	8.7E-108	Os01g0846300
	XP_004241211.1	*Solyc06g076400.2.1*	*Sl*	3	410	67%	76%	98%	7.2E-171	PREDICTED: probable protein phosphatase 2C 24
	XP_006350789.1	*PGSC0003DMG400030332*	*St*	3	410	67%	78%	97%	3E-173	PREDICTED: protein phosphatase 2C 37-like
	XP_002282608.1	*GSVIVT01024875001*	*Vv*	3	408	65%	74%	98%	2.3E-158	PREDICTED: probable protein phosphatase 2C 24
Cc02_g07430	NP_172223.1	*AT1G07430.1*	*At*	2	442	63%	77%	69%	1.3E-134	protein phosphatase 2C 3
	XP_006488392.1	*orange1.1g036852m*	*Cs*	2	429	59%	72%	99%	1.3E-150	PREDICTED: probable protein phosphatase 2C 78-like
	NP_001044788.1	*LOC_Os01g62760.1*	*Os*	3	414	58%	71%	73%	1.6E-121	Os01g0846300
	XP_004239911.1	*Solyc05g052980.2.1*	*Sl*	3	409	68%	78%	100%	0	PREDICTED: protein phosphatase 2C 37
	XP_006355694.1	*PGSC0003DMG400027196*	*St*	3	418	67%	77%	100%	0	PREDICTED: protein phosphatase 2C 37-like
	XP_002282703.1	*GSVIVT01016485001*	*Vv*	3	400	63%	73%	99%	7.9E-163	PREDICTED: protein phosphatase 2C 37-like
Cc04_g01620	NP_177421.1	*AT1G72770.1*	*At*	4	511	56%	74%	98%	0	protein phosphatase 2C 16
	XP_006465975.1	*orange1.1g009094m*	*Cs*	4	544	61%	74%	98%	0	PREDICTED: protein phosphatase 2C 16-like
	EEE54872.1	*LOC_Os01g40094.1*	*Os*	3	352	66%	77%	64%	4.8E-161	hypothetical protein OsJ_02363
	BAI39595.1	*Solyc03g121880.2.1*	*Sl*	4	544	73%	82%	98%	0	protein phosphatase 2C ABI2 homolog
	XP_006342955.1	*PGSC0003DMG400002573*	*St*	4	545	73%	83%	98%	0	PREDICTED: protein phosphatase 2C 16-like
	XP_002278167.2	*GSVIVT01016816001*	*Vv*	3	548	62%	74%	98%	0	PREDICTED: protein phosphatase 2C 16
Cc06_g11740	NP_177421.1	*AT1G72770.1*	*At*	4	511	63%	74%	61%	1.1E-142	protein phosphatase 2C 16
	KDO73536.1	*orange1.1g008890m*	*Cs*	4	550	49%	65%	100%	4.8E-161	hypothetical protein CISIN_1g008880mg
	NP_001046464.1	*LOC_Os01g40094.1*	*Os*	3	396	63%	75%	61%	6.9E-137	Os01g0583100, partial
	XP_004243737.1	*Solyc07g040990.2.1*	*Sl*	3	543	52%	68%	99%	0	PREDICTED: probable protein phosphatase 2C 50
	XP_006342333.1	*PGSC0003DMG400018004*	*St*	4	543	52%	68%	99%	0	PREDICTED: probable protein phosphatase 2C 6-like
	XP_002279140.1	*GSVIVT01015156001*	*Vv*	3	550	50%	66%	100%	5E-171	PREDICTED: protein phosphatase 2C 77
Cc08_g11010	NP_177421.1	*AT1G72770.1*	*At*	4	511	49%	63%	100%	2.6E-158	protein phosphatase 2C 16
	KDO73536.1	*orange1.1g008890m*	*Cs*	4	550	65%	78%	99%	0	hypothetical protein CISIN_1g008880mg
	NP_001065470.1	*LOC_Os01g40094.1*	*Os*	3	396	69%	80%	59%	2E-155	Os01g0583100, partial
	XP_004253091.1	Solyc12g096020.1	*Sl*	3	540	66%	79%	100%	0	PREDICTED: probable protein phosphatase 2C 53
	XP_006342498.1	*PGSC0003DMG400029297*	*St*	3	536	67%	79%	100%	0	PREDICTED: probable protein phosphatase 2C 6-like
	XP_010648365.1	*GSVIVT01035420001*	*Vv*	4	551	66%	78%	100%	0	PREDICTED: probable protein phosphatase 2C 53
Cc08_g16010	NP_172223.1	*AT1G07430.1*	*At*	2	442	47%	62%	64%	1.1E-80	protein phosphatase 2C 3
	KDO76517.1	*orange1.1g023178m*	*Cs*	4	286	57%	71%	65%	2.59E-95	hypothetical protein CISIN_1g023178mg
	NP_001043754.1	*LOC_Os01g46760.1*	*Os*	2	403	55%	70%	64%	2.77E-98	Os01g0656200
	XP_004240955.1	*Solyc06g051940.2.1*	*Sl*	2	442	57%	73%	67%	4.6E-103	PREDICTED: probable protein phosphatase 2C 51
	XP_006350568.1	*PGSC0003DMG400009112*	*St*	2	399	54%	70%	64%	1.88E-94	PREDICTED: probable protein phosphatase 2C 51-like
	XP_002266149.1	*GSVIVT01019525001*	*Vv*	2	393	57%	75%	70%	2E-107	PREDICTED: probable protein phosphatase 2C 51

Table S3 Comparison of CcSnRK2 protein sequences with orthologous sequences from *A. thaliana* (At), *C. sinensis* (Cs), *O. sativa* (Os), *S. lycopersicum* (Sl), *S. tuberosum* (St) and *V. vinifera* (Vv) plant species (Sp). NCBI accession numbers (ID), I (introns), aa (amino acid length), Ident. (Identity), Align. (Match/Aligned), QC (Query Cover), e-value and function information were obtained through BLASTp results. Coffee IDs were identified in Coffee Genome Hub (http://www.coffee-genome.org/) and Gene IDs in Phytozome 10.3 (http://phytozome.jgi.doe.gov/pz/portal.html).

Coffee ID	NCBI ID	*Gene ID*	Sp	I	Len.aa	Ident.	Align.	Query cov.	e-value	Function inferred by BLAST
Cc00_g07830	NP_172563.1	*AT1G10940.1*	*At*	9	363	80%	88%	70%	4.32E-76	serine/threonine-protein kinase SRK2A
	XP_006477070.1	*orange1.1g019433m*	*Cs*	8	341	84%	90%	70%	2.21E-77	PREDICTED: serine/threonine-protein kinase SAPK3-like
	NP_001050274.1	*LOC_Os03g27280.1*	*Os*	8	342	82%	88%	70%	2.78E-75	Os03g0390200
	XP_004245833.1	*Solyc08g077780.2.1*	*Sl*	8	339	82%	90%	70%	9.46E-77	PREDICTED: serine/threonine-protein kinase SAPK3
	XP_006359207.1	*PGSC0003DMG400026211*	*St*	8	339	82%	90%	70%	9.46E-77	PREDICTED: serine/threonine-protein kinase SAPK3-like
	XP_002262726.1	GSVIVT01004839001	*Vv*	8	340	83%	88%	70%	5.14E-76	PREDICTED: serine/threonine-protein kinase SAPK3
Cc00_g19320	NP_196476.1	*AT5G08590.1*	*At*	8	353	85%	93%	90%	0	serine/threonine-protein kinase SRK2G
	XP_006466196.1	orange1.1g018734m	*Cs*	8	354	84%	92%	99%	0	PREDICTED: serine/threonine-protein kinase SRK2A-like
	NP_001052827.1	*LOC_Os04g35240.1*	*Os*	8	359	76%	85%	99%	0	Os04g0432000
	XP_004239628.1	*Solyc05g056550.2.1*	*Sl*	6	356	85%	91%	99%	0	PREDICTED: serine/threonine-protein kinase SAPK7-like
	NP_001274892.1	*PGSC0003DMG400023803*	*St*	9	360	82%	90%	98%	0	serine/threonine-protein kinase SRK2B-like
	XP_002267922.1	*GSVIVT01022427001*	*Vv*	8	355	85%	90%	99%	0	PREDICTED: serine/threonine-protein kinase SRK2A
Cc00_g35430	NP_172563.1	*AT1G10940.1*	*At*	9	363	87%	93%	94%	8.76E-71	serine/threonine-protein kinase SRK2A
	KDO57025.1	*orange1.1g024336m*	*Cs*	8	269	92%	95%	94%	7.14E-75	hypothetical protein CISIN
	BAT12097.1	*LOC_Os10g41490.1*	*Os*	8	289	88%	94%	94%	1.8E-72	Os10g0564500
	XP_004245833.1	*Solyc08g077780.2.1*	*Sl*	8	339	92%	95%	94%	5.72E-74	PREDICTED: serine/threonine-protein kinase SAPK3
	XP_006359207.1	*PGSC0003DMG400026211*	*St*	8	339	92%	95%	94%	7.26E-74	PREDICTED: serine/threonine-protein kinase SAPK3-like
	XP_002262726.1	*GSVIVT01004839001*	*Vv*	8	340	92%	94%	94%	3.02E-73	PREDICTED: serine/threonine-protein kinase SAPK3
Cc02_g18420	NP_567945.1	*AT4G33950.1*	*At*	9	362	87%	93%	99%	0	calcium-independent ABA-activated protein kinase
	KDO49166.1	*orange1.1g017933m*	*Cs*	7	363	91%	97%	99%	0	hypothetical protein CISIN
	NP_001050653.1	*LOC_Os03g41460.1*	*Os*	6	362	87%	92%	99%	0	Os03g0610900
	XP_004230794.1	*Solyc01g108280.2.1*	*Sl*	9	362	93%	98%	99%	0	PREDICTED: serine/threonine-protein kinase SRK2E
	NP_001275318.1	*PGSC0003DMG400025895*	*St*	10	362	93%	97%	99%	0	serine/threonine-protein kinase SRK2E-like
	XP_002284959.1	*GSVIVT01031806001*	*Vv*	8	363	90%	96%	99%	0	PREDICTED: serine/threonine-protein kinase SAPK10

Table S3 Continue…for legend see the previous page.

Coffee ID	NCBI ID	*Gene ID*	Sp	I	Len.aa	Ident.	Align.	Query cov.	e-value	Function inferred by BLAST
Cc02_g22790	AAM67112.1	*AT1G60940.1*	*At*	8	361	84%	91%	99%	0	putative serine/threonine-protein kinase
	XP_006471015.1	*orange1.1g018734m*	*Cs*	8	351	87%	92%	99%	0	PREDICTED: serine/threonine-protein kinase SRK2A-like
	NP_001052827.1	*LOC_Os04g35240.1*	*Os*	8	359	83%	93%	99%	0	Os04g0432000
	XP_004230475.1	*Solyc01g103940.2.1*	*Sl*	9	361	88%	94%	99%	0	PREDICTED: serine/threonine-protein kinase SRK2B
	NP_001274892.1	*PGSC0003DMG400023803*	*St*	9	360	91%	96%	99%	0	serine/threonine-protein kinase SRK2B-like
	XP_002269221.1	*GSVIVT01023339001*	*Vv*	8	356	91%	95%	99%	0	PREDICTED: serine/threonine-protein kinase SRK2A
Cc07_g05710	NP_201489.1	*AT5G66880.1*	*At*	8	361	83%	91%	99%	0	serine/threonine-protein kinase SRK2I
	KDO49166.1	*orange1.1g017933m*	*Cs*	7	363	84%	93%	99%	0	hypothetical protein CISIN
	NP_001050653.1	*LOC_Os03g41460.1*	*Os*	6	362	81%	91%	99%	0	Os03g0610900
	XP_004232055.1	*Solyc02g090390.2.1*	*Sl*	8	352	86%	95%	99%	0	PREDICTED: serine/threonine-protein kinase SRK2I
	XP_006338224.1	*PGSC0003DMG400025895*	*St*	8	352	86%	95%	99%	0	PREDICTED: serine/threonine-protein kinase SRK2I-like
	XP_002284959.1	*GSVIVT01031806001*	*Vv*	8	363	82%	91%	99%	0	PREDICTED: serine/threonine-protein kinase SAPK10
Cc07_g14700	NP_567945.1	*AT4G33950.1*	*At*	9	362	73%	87%	93%	1.2E-173	calcium-independent ABA-activated protein kinase
	XP_006466260.1	orange1.1g019628m	*Cs*	8	341	82%	90%	99%	0	PREDICTED: serine/threonine-protein kinase SAPK2-like
	NP_001060312.1	*LOC_Os07g42940.1*	*Os*	8	339	79%	88%	99%	0	Os07g0622000
	XP_010312635.1	Solyc04g012160.2.1	*Sl*	8	345	80%	87%	99%	0	PREDICTED: serine/threonine-protein kinase SAPK2
	NP_001274912.1	*PGSC0003DMG400023636*	*St*	8	344	79%	87%	99%	0	serine/threonine-protein kinase SAPK2-like
	XP_003632469.1	*GSVIVT01003419001*	*Vv*	8	338	83%	91%	99%	0	PREDICTED: serine/threonine-protein kinase SAPK2
Cc08_g11200	NP_567945.1	*AT4G33950.1*	*At*	9	362	72%	87%	95%	9.9E-180	calcium-independent ABA-activated protein kinase
	XP_006477070.1	*orange1.1g019433m*	*Cs*	8	341	84%	91%	99%	0	PREDICTED: serine/threonine-protein kinase SAPK3-like
	BAD17999.1	*LOC_Os10g41490.1*	*Os*	9	334	85%	93%	90%	0	serine/threonine protein kinase SAPK3
	XP_004245833.1	*Solyc08g077780.2.1*	*Sl*	8	339	83%	91%	100%	0	PREDICTED: serine/threonine-protein kinase SAPK3
	XP_006359207.1	*PGSC0003DMG400026211*	*St*	8	339	82%	91%	100%	0	PREDICTED: serine/threonine-protein kinase SAPK3-like
	XP_002262726.1	*GSVIVT01004839001*	*Vv*	8	340	84%	90%	99%	0	PREDICTED: serine/threonine-protein kinase SAPK3
Cc10_g06790	NP_974170.1	*AT1G78290.3*	*At*	5	343	83%	91%	88%	0	serine/threonine-protein kinase SRK2C
	KDO81023.1	*orange1.1g019628m*	*Cs*	8	338	83%	90%	99%	0	hypothetical protein CISIN
	NP_001050274.1	*LOC_Os03g27280.1*	*Os*	8	342	77%	87%	99%	0	Os03g0390200
	XP_004237936.1	*Solyc04g074500.2.1*	*Sl*	8	336	80%	90%	98%	0	PREDICTED: serine/threonine-protein kinase SAPK2
	NP_001275016.1	*PGSC0003DMG400030830*	*St*	8	335	80%	90%	98%	0	serine/threonine-protein kinase SAPK2-like
	XP_003634478.1	*GSVIVT01009074001*	*Vv*	8	335	81%	90%	99%	0	PREDICTED: serine/threonine-protein kinase SAPK2

CHAPTER 2

Gene expression profiles in *Coffea arabica* and *Coffea canephora* leaves revealed transcriptional regulations of key genes involved in ABA signaling

GENE EXPRESSION PROFILES IN *COFFEA ARABICA* AND *COFFEA CANEPHORA* LEAVES REVEALED TRANSCRIPTIONAL REGULATIONS OF KEY GENES INVOLVED IN ABA SIGNALING.

INTRODUCTION

Stomatal guard cells are functionally specialized epidermal cells usually located on plant aerial organs which control gas exchanges between plant and the surrounding atmosphere. These guard cells have developed mechanisms to sense and respond to various endogenous and environmental stimuli (Hetherington and Woodward, 2003; Gray, 2005; Masle *et al.*, 2005).

The role of ABA in guard cell regulation after drought response has been extensively studied since a long time (Schroeder *et al.*, 2001*a*; Nilson and Assmann, 2007; Sirichandra *et al.*, 2009). For example, applications of exogenous ABA was show to stimulate stomatal closure in the wilty tomato *flacca* mutant deficient in ABA (Imber and Tal, 1970; Tal *et al.*, 1970), as well as in *Xanthium* (Jones and Mansfield, 1970). The opening and closing of the stomatal pore are regulated by osmotic pressure of guard cells envolving dynamic changes in the intracellular concentrations of inorganic ions and sugars (Sirichandra *et al.*, 2009).

It is well known that the ABA PYR/PYL/RCAR receptors play a key role for the whole-plant stomatal adjustments and responses to low humidity, darkness, and elevated CO_2, for example (Merilo *et al.*, 2013). Under drought, some plant species maintain leaf water potential (isohydric behavior) while other favor stomatal conductance to maintain CO_2 assimilation (anisohydric behaviour). The first mechanism results of the enhancement of the ABA effect on stomatal conductance (g_s) by low Ψ_{leaf} (Tardieu and Simonneau, 1998). ABA production induced by low Ψ_{leaf} is thought to prevent stomata to reach their maximal opening by a transduction network involving ABI1 and ABI2 protein phosphatases 2C and the OST2 and SLAC1 effectors (Kim *et al.*, 2010). On the other hand, vascular ABA decreases K_{leaf} putatively by inactivating aquaporins such as the plasm membrane intrinsic proteins (PIPs) (Shatil-Cohen *et al.*, 2011), through a transduction pathway distinct from the network already described. This conceptual model for the dual action of ABA on stomata closure has been recently proposed (Pantin *et al.*, 2013).

Regarding the key roles of tripartite system in higher plants, the following scientific questions arisen concerning coffee:

- how the *PYL*/*PP2C*/*SnRK2* genes are expressed in leaves of coffee plants in response to exogenous ABA?
- are they differentially expressed in D^T and D^S clones?
- does it exist different expression profiles of these genes in *C. arabica* and *C. canephora*?

- is it possible to correlate the expression profiles of the genes with stomatal responses in the D^T and D^S clones of *C. canephora* and *C. arabica*?
- Is it possible to correlate the expression profiles of these genes with those observed under drought conditions for the *C. canephora* plants (chapter I)?

Aiming to get the answers to these questions, the main objectives of this work were:

(i) to cultivate in hydroponic conditions *C. arabica* and *C. canephora* plants;

(ii) to characterize the expression profile in time-course of genes belonging to the tripartite system (*PYR -PP2C-SnRK2*) in leaves of D^T and D^S clones of *C. canephora* and *C. arabica* submitted to exogenous ABA treatment;

(iii) to study the effects of exogenous ABA on stomatal aperture in *C. canephora* and *C. arabica* plants;

MATERIAL AND METHODS

Plant material

D^T (14, 73 and 120) and D^S (22) clones of *C. canephora* corresponded to those previously described in the chapter I were grown in greenhouse conditions (under controlled temperature 25°C, relative humidity of 70% and photosynthetic flux PPF 900 $\mu mol^{-2}s^{-1}$) in small containers at UFV (University of Viçosa-UFV, Minas Gerais, Brazil) and used for stem cuttings to generate *C. canephora* plantlets to be tested in hydroponic conditions. Plants of the D^T (IAPAR59) and D^S (Rubi) cultivars of *C. arabica* were obtained from seeds harvested in the experimental fields of Embrapa Cerrados that were germinated in deionized water.

Hydroponic condition for ABA experiment

For both D^T and D^S genotypes of *C. canephora* and *C. arabica*, 2 plants were used as biological repetitions. The plants were hydroponically grown in culture room with 150-200 µmol photon/m^2/s light intensity, 12/12 dark/light hours, 70% relative humidity at 24±1°C in pH 5.5 adjusted Hoagland solution (Hoagland, D.R.; Arnon, 1950) ¼ strength. For hydroponic assay, *C. canephora* and *C. arabica* plants of 6 and 3 months-old, respectively, were transferred from the greenhouse to culture room in individual pots (300 mL) immersed with nutritive solution that was renewed weekly. ABA assays were performed one month after plants acclimation in hydroponic conditions by adding ABA to a final concentration of 500 µM in the nutritive Hoagland solution.

RNA extraction

RNAs were extracted as previously described (Marraccini *et al.*, 2012) from the first pair of leaf of coffee plants grown in hydroponic conditions where they were submitted to ABA treatment during 3

days. The samples were collected at 11:30 am in control (Hoagland ¼ strength w/o ABA) and under ABA (500 µM) conditions at the first and third days. All purified RNAs were quantified using a NanoDrop 1000 Spectrophotometer (Waltham, MA, USA). Contaminant genomic DNA was eliminated from purified RNAs by RQ1 RNase-free DNase (Promega) treatment according to the fabricant. RNA integrity was verified by agarose gel electrophoresis with ethidium bromide staining. Synthesis of the first-strand cDNA was done by treating 2.4µg of total RNA with the ImProm-II Reverse Transcription System and oligo (dT15) according to the manufacturer's recommendations (Promega).

Real time qPCR assays

Genomic DNA was eliminated from purified RNAs by RQ1 RNase-free DNase (Promega) treatment according to the fabricant. RNA integrity was verified by agarose gel electrophoresis with ethidium bromide staining. Synthesis of the first-strand cDNA was done by treating 2.4µg of total RNA with the ImProm-II Reverse Transcription System and oligo (dT15) according to the manufacturer's recommendations (Promega). Real-time qPCR assays were carried out with the synthesized single-stranded cDNA described above and using the protocol recommended for 7500 Fast Real-Time PCR Systems (Applied Biosystems, Foster City, CA, USA). cDNA preparations were diluted (1/20) and tested by qPCR using 48 primer pairs designed for the 24 candidate genes of the tripartite systems. Primer pairs were designed using the Primer Express software (Applied Biosystems) and preliminarily tested for their specificity with a cDNA mix from roots. The qPCR was performed with 1µl of diluted single-stranded cDNA and 0.2 µM (final concentration) of each primer in a final volume of 10µl with 1x SYBR green fluorochrome (SYBRGreenqPCR Mix-UDG/ROX, Invitrogen). The reaction mixture was incubated for 2 min at 50°C and 5 min at 95°C (UDG step), followed by 40 amplification cycles of 3 s at 95°C, 30 s at 60°C. Data were analyzed using the SDS 2.1 software (Applied Biosystems) to determine the cycle threshold (Ct) values. Specificity of the PCR products generated for each set of primers was verified by analyzing the Tm (dissociation) of amplified products. Gene expression levels were normalized to expression level of ubiquitin (*CcUBQ10*) as a constitutive reference (Barsalobres-Cavallari *et al.*, 2009). Expression was expressed as relative quantification by applying the formula $(1+E)^{-\Delta\Delta Ct}$, where $\Delta Ct_{\text{target}} = Ct_{\text{target gene}} - Ct_{\text{reference gene}}$ and $\Delta\Delta Ct = \Delta Ct_{\text{target}} - \Delta Ct_{\text{internal calibrator}}$, the internal calibrator always being the 14I sample with relative quantification equal to 1. Data are presented as the mean ± standard error of the mean. Graphs are generated and analyzed using GraphPad Prism ©.

Microscopic analyses

For each genotype, the first pair of leaf from two different plants was used for transversal sections. Two different areas of the leaves were collected twice at mid-day at 11:30 am before ABA treatment (control) and at the same time in each one of the three days of assay. Additional sample was collected in the third day at 6 pm. Immediately after harvest, the material was fixed in FAA 50%

(formaldehyde, acetic acid and ethanol) solution for both scanning electronic (SEM) or optical microscopy. After 24 hours of incubation, samples were dehydrated through a graded series of ethanol until 70% and then cleared in sodium hypochlorite 2,5% over 2 hours before to be analyzed by microscopy (Leica DM 750 microscope). For optical analyses, images were treated using the Leica Application Suite 3.0 LasEz software and stomatal densities were determined using ImageJ software (National Institutes of Health, Bethesda, MD, USA). For calculation of stomatal aperture, the length of one hundred stomatal guard cells was measured in leaves of each genotype under control and ABA treatments.

RESULTS

Expression profiles of CcPYL-PP2C-SnRK2 genes in hydroponic-grown plants treated with exogenous ABA

The expression profiles of the tripartite system *CcPYL-PP2C-SnRK2* genes were analyzed in coffee plants growing under hydroponic condition and submitted to ABA treatment (Figure). For this purpose, plants of *C. canephora* and *C. arabica* were incubated during three days in nutritive solution containing 500 µM of ABA. Leaf samples were collected for all plants (*C. canephora*, D^T: clone 14 and D^S: clones 22; *C. arabica* D^T I59 and D^S Rubi) before assay (control, without ABA) and after one (24 hours) and three days (72 hours) under ABA treatment. These samples were used for qPCR (Figure 3 and Figure 4) experiments and microscopy analyses (Figure 5).

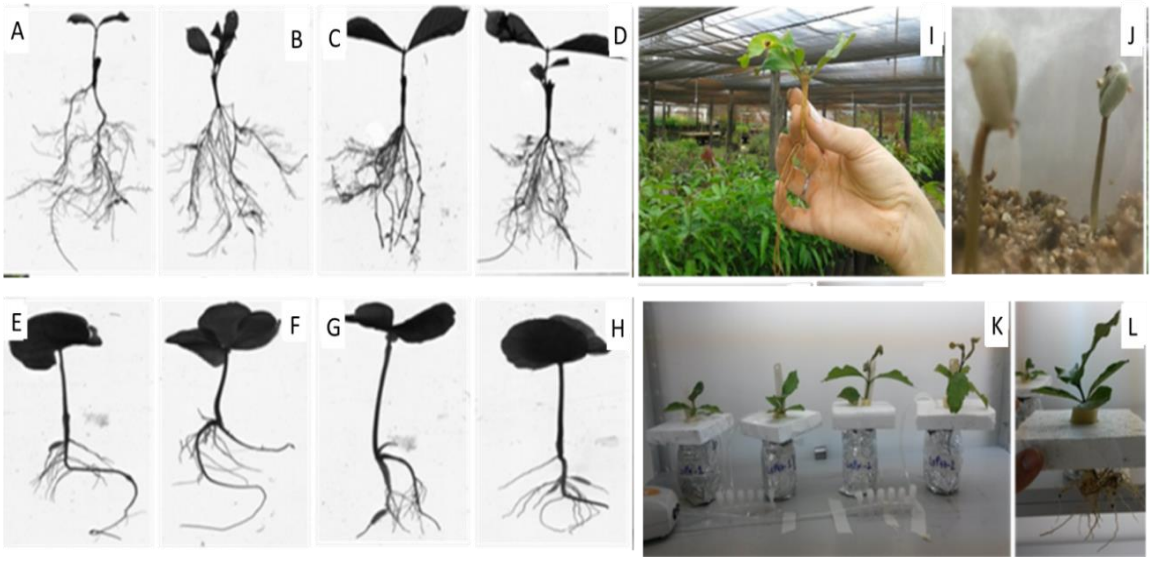


Figure 1 Experimental condition for hydroponic assays. C. canephora D^T clone 14 (A-B) and D^S clone 22 (C-D) were originated from stem cuttings (I). The plantlets of D^T cultivar IAPAR59 (E-F) and D^S cultivar Rubi (G-H) of *C. arabica* were originated from germinated seeds (J). Images of individual plantlets (A-H) were generated using the WinRhizo software prior ABA treatment. All plants were hydroponically grown (K-L) in controlled growth chamber.

Expression of PYL genes

Among the nine *CcPYL* genes previously identified (chapter I), *PYL2*, *CcPYL7a*, *PYL7b* and *PYL13* were not expressed in leaves of *C. canephora* and *C. arabica* genotypes either under control or ABA treatments. However, the *PYR1*, *PYL4*, *PYL8a*, *PYL8b* and *PYL9* genes were expressed in low level in leaves of all coffee genotypes grown under hydroponic conditions without ABA (Figure 2).

In *C. canephora*, up-regulated expression of *CcPYR1* and *CcPYL8b* genes was clearly observed after 24 hours of ABA treatment specifically in leaves of the D^T clone 14 but not in those of D^S clone

22. For both clones, leaf expression of *PYR1*, *PYL4*, *PYL8a*, *PYL8b* and *PYL9* genes decrease hereafter to be undetectable at 72h of ABA treatment.

In *C. arabica*, *CaPYL8a* was the only gene showing up-regulated expression under at 24h of ABA treatment in I59. At 72h of ABA treatment, leaf expression of *CaPYR1*, *CaPYL4*, *CaPYL8a* and *CaPYL9* genes was no more detected in I59 but observed in Rubi. In both genotypes, expression of *CaPYL9* gene was undetected in control and at 24h of ABA treatment and considered as low at 72h (Figure 2).

Altogether, this study clearly highlighted the existence of different *PYL* expression profiles between D^T and D^S clones in each coffee species but also between *C. canephora* and *C. arabica* plantlets, mainly regarding the time-course of *PYL* expression upon ABA treatments.

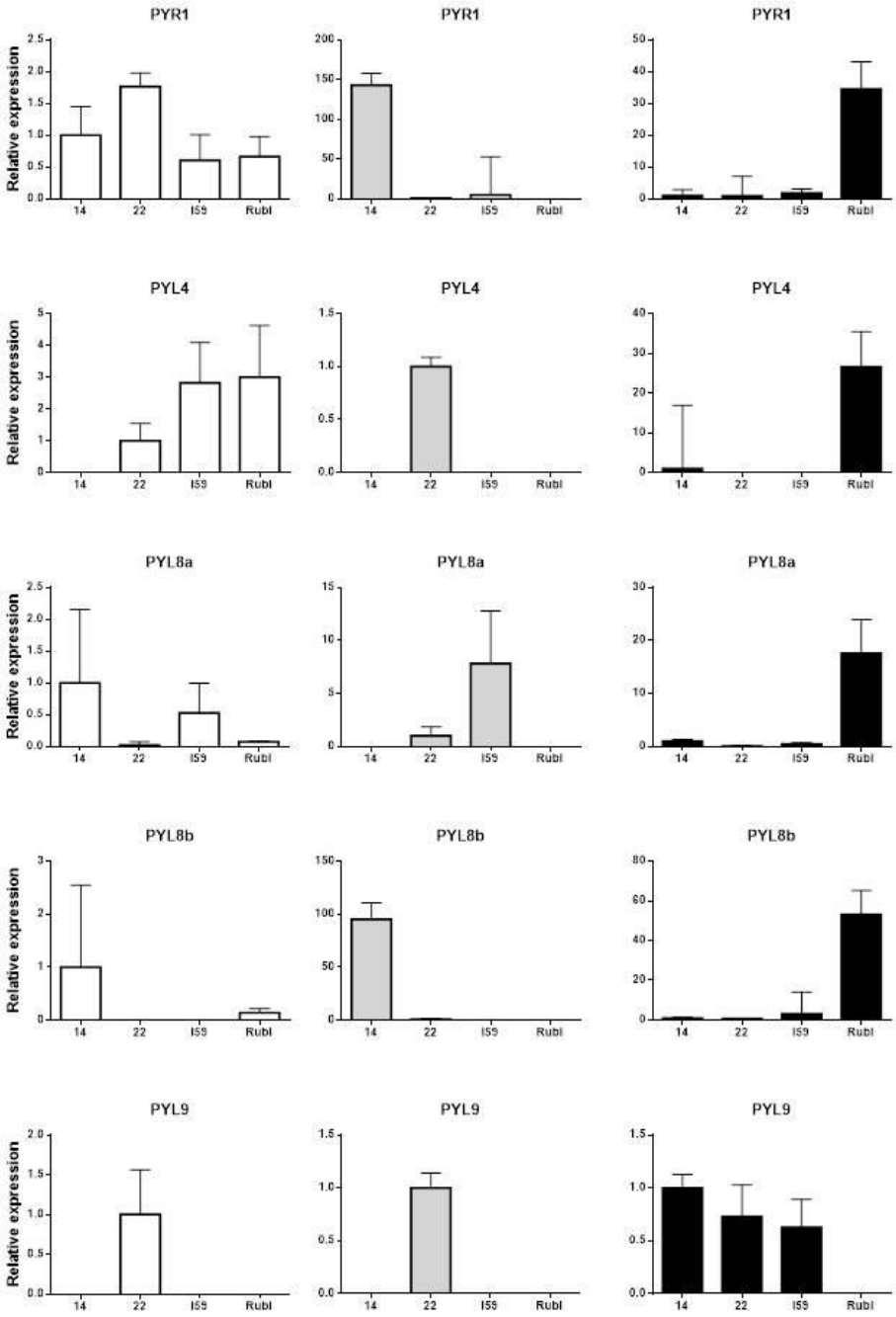


Figure 2 Expression profiles of *PYL* genes in leaves of *C. canephora* D^T (clone 14) and D^S (clone 22) and *C. arabica* D^T (I59) and D^S (Rubi) plants in response to exogenous ABA. RNA were extracted from leaves of coffee plantlets without exogenous ABA (control, white bars) as well as after 24 (grey bars) or 72 hours (black bars) under ABA treatment (500 µM), *PYL* genes studied corresponded to *PYR1*, *PYL4*, *PYL8a*, *PYL8b* and *PYL9* genes. Values are the mean of at least three technical repetitions ± SD which are standardized independently with *UBQ10* (ubiquitin) as reference gene. The clone 14 was choose as preferential internal calibrator (RE=1).

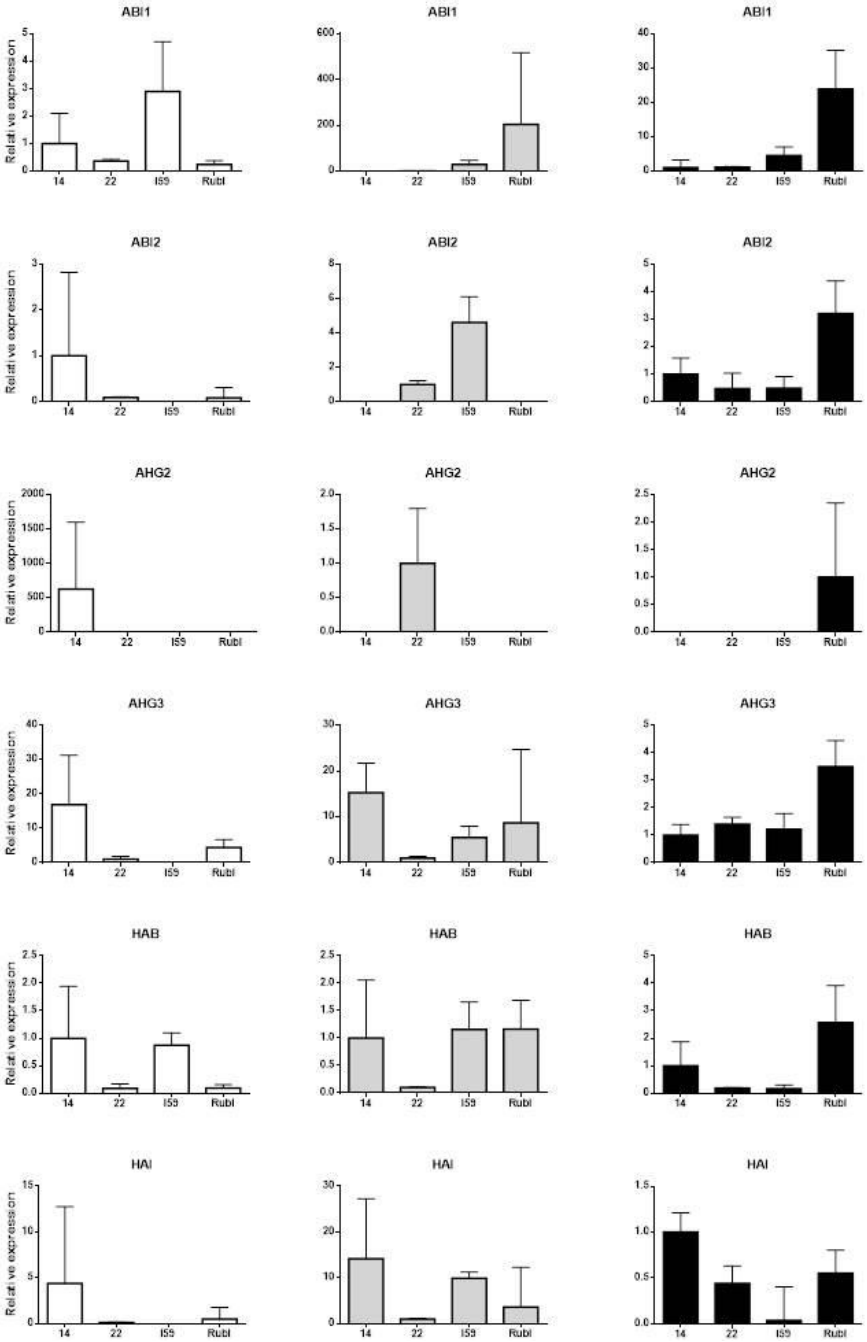


Figure 3 Expression profiles of *PP2C* genes in leaves of *C. canephora* D^T (clone 14) and D^S (clone 22) and *C. arabica* D^T (I59) and D^S (Rubi) plants in response to exogenous ABA. RNA were extracted from leaves of coffee plantlets without exogenous ABA (control, white bars) as well as after 24 (grey bars) or 72 hours (black bars) under ABA treatment (500 µM), *PP2C* genes studied corresponded to *ABI1-2*, *AGH2-3*, *HAB*, *HAI* genes. Values are the mean of at least three technical repetitions ± SD which are standardized independently with *CcUBQ10* (ubiquitin) as reference gene. The clone 14 was choose as preferential internal calibrator (RE=1).

Expression of PP2C genes

Among the six *CcPP2C* genes previously identified (chapter I), and except *AHG2 in C. arabica*, all (*CcABI1* and *2*, *CcAHG3*, *CcHAB* and *CcHAI*) were expressed in leaves of both *C. canephora* and *C. arabica* plantlets in hydroponic prior to ABA treatment (Figure 3).

In *C. canephora*, it is worth noting the higher expression level of *CcAHG2*, *CcAHG3* and *CcHAI* genes in D^T clone 14 compared to D^S clone 22 under unstressed conditions. After 24h of ABA treatment, *CcAHG2* leaf expression decreased significantly in D^T clone 14. However, expression profiles of all other genes were similar to those observed in the control condition, and continued to be low at 72h of ABA.

In *C. arabica* and whatever the genotype, expression levels of *PP2C* genes were considered as low under control condition. After 24h of ABA treatment, the main changes in expression profiles were observed for *ABI1* gene that was highly up-regulated in cultivar Rubi but not in I59. Even though, *ABI1* expression levels decreased hereafter, to be lower than those measured at 24h, *ABI1* expression continued to be higher in Rubi than in I59 at 72h of ABA treatment. Interestingly, *AHG2* expression was not detected in leaf of both cultivars under control condition and after 24h of ABA, but was detectable at 72h of ABA treatment only in leaves of Rubi D^S cultivar. For other *PP2Cs*, ABA treatments did not modify significantly gene expression profiles that were considered as low and relatively stable in both cultivars.

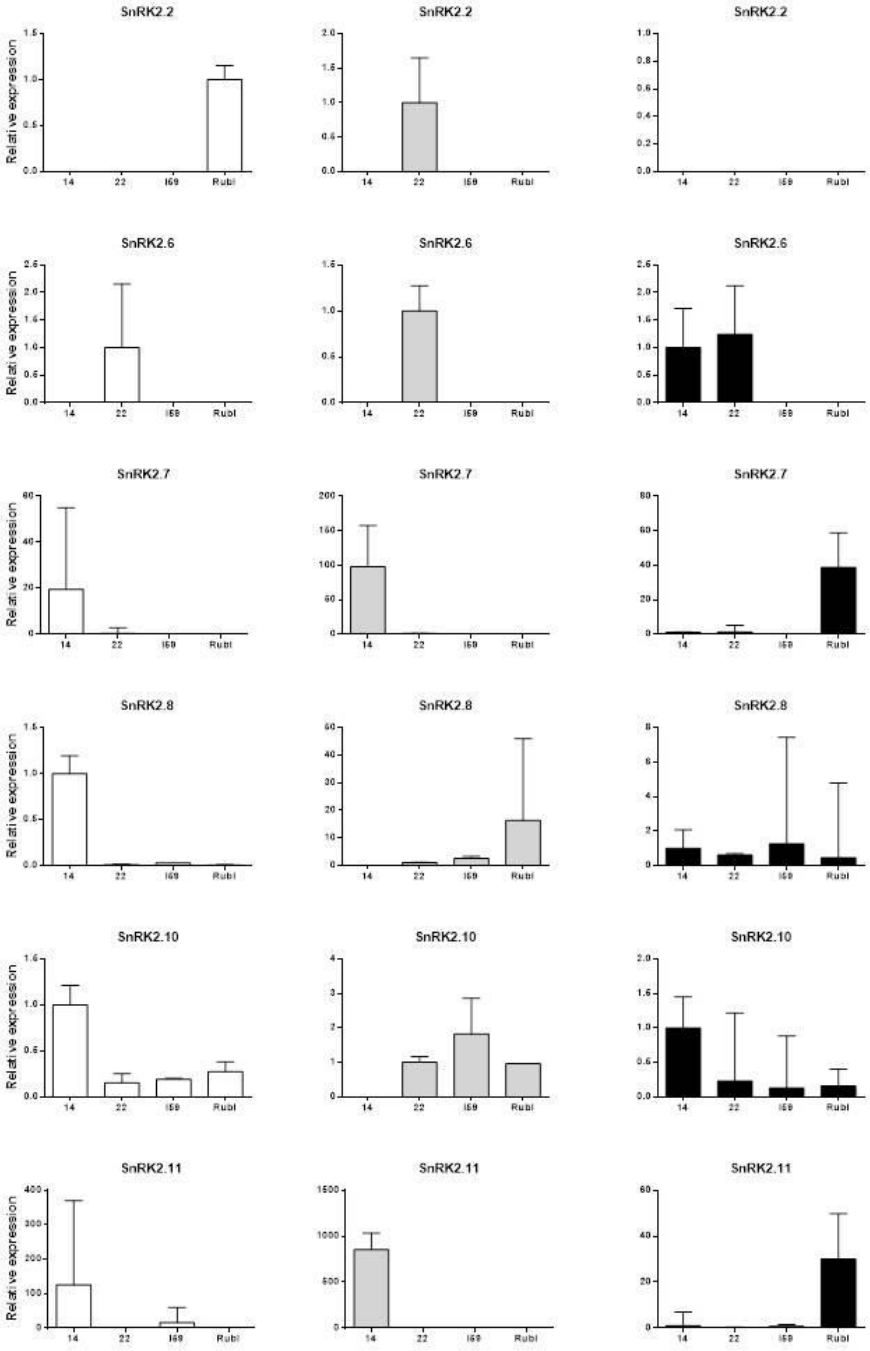


Figure 4 Expression profiles of *SnRK2* genes in leaves of *C. canephora* D^T (clone 14) and D^S (clone 22) and *C. arabica* D^T (I59) and D^S (Rubi) plants in response to exogenous ABA. RNA were extracted from leaves of coffee plantlets without exogenous ABA (control, white bars) as well as after 24 (grey bars) or 72 hours (black bars) under ABA treatment (500 µM), *SnRK2* genes studied corresponded to *SnRK2.2, SnRK2.6, SnRK2.7, SnRK2.8, SnRK2.10, SnRK2.11* genes. Values are the mean of at least three technical repetitions ± SD which are standardized independently with *CcUBQ10* (ubiquitin) as reference gene. The clone 14 was choose as preferential internal calibrator (RE=1).

Expression of SnRK2 genes

Among the *CcSnRK2* previously identified (chapter 1), six of them (*CcSnRK2.2*, *CcSnRK2.6*, *CcSnR2.7*, *CcSnRK2.8*, *CcSnRK2.10* and *SnRK2.11*) were studied by qPCR experiments (Figure 30). While *CcSnRK2.6* gene was expressed in both D^T and D^S clones of *C. canephora*, it is worth noting that expression of this gene was not detected in both cultivars of *C. arabica*. On the other hand, we can point out that *CcSnRK2.10* expression profiles detected in all coffee genotypes were not greatly affected by ABA treatments. For other *SnRK2* genes, the main differences observed between coffee species, genotypes and ABA treatments are given below.

In *C. canephora*, expression of *CcSnRK2.2* was undetectable in leaves of both clones under control condition. Under these conditions, it is worth noting higher expression level in D^T clone 14 than D^S clone 22 mainly for *CcSnR2.7* and *CcSnR2.11* genes, and to a lesser extend for *CcSnRK2.8*, and *CcSnRK2.10*. The contrary was observed for *CcSnRK2.6* that had higher expression in D^S clone 22 than in D^T clone 14. Expression of *CcSnRK2.7* and *CcSnRK2.11* genes appeared greatly up-regulated in D^T clone 14 after 24h of ABA treatment, and decreased drastically hereafter at 72h of ABA treatment. In parallel and whatever the tested conditions, expression of *CcSnRK2.7* and *CcSnR2.11* gene was always undetected in leaves of D^S clone 22.

In *C. arabica*, it is worth noting that expression profiles of all *SnRK2* genes were always low, up to undetectable in the D^S cultivar IAPAR59. In the D^S cultivar Rubi, expression of *SnRK2.2* clearly decreased after 24h of ABA treatment while the contrary was observed for *SnRK2.8* gene. In this cultivar, expression of *SnRK2.7* and *SnRK2.11* was highly up-regulated after 72h of ABA treatment, while *SnRK2.8* gene expression decreased.

Effects of ABA treatments on stomatal closure in D^T and D^S clones of C. canephora.

In leaves, the D^T and D^S plants of *C. canephora* presented differences in stomatal cell responses under ABA treatment (Figure 5).

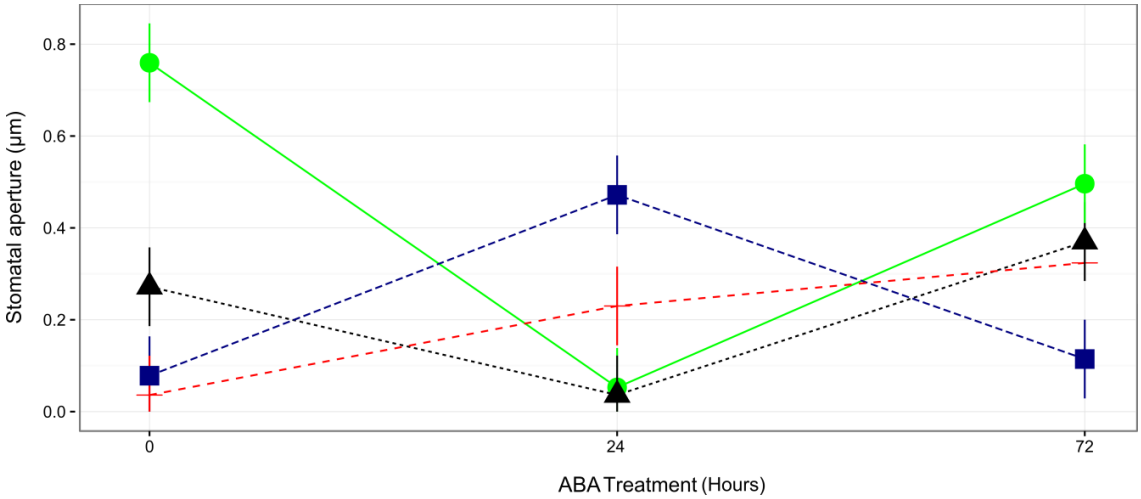


Figure 5 Evaluation of ABA effect in modulating the guard cells stomatal aperture in coffee leaves. The stomatal aperture length was measured in the guard cells of *C. canephora* D^T 14 (circles) and D^S 22 clones (triangles) and *C. arabica* D^T I59 (squares) and D^S Rubi (crosses) cultivars in 24 and 72 hours after application of exogenous ABA 500 µM solution and under control conditions (without ABA, 0h). The stomatal aperture values are given as an average of a hundred cells measurements for each clone/cultivar.

Under control condition, the *C. canephora* D^T clone 14 showed in average larger stomatal aperture than D^S clone 22 (Figure 5). After 24h and 72h of ABA treatment, no significant difference was observed between D^T and D^S clones. When analyzed separately, either clones showed significant responses to ABA treatments. After the first 24h, the clones D^T 14 and D^S 22 showed an increase of stomatal closure (decrease in guard cells stomatal aperture). After 72h, this process was fully reversed in D^S clone 22, where none significant difference was observed in stomatal aperture between 0h and 72h, and partially reversed in D^T clone 14, where despite of stomatal aperture increase significant differences were still observed between 0h and 72h. In contrast, *C. arabica* cultivars D^T I59 and D^S Rubi presented smaller stomatal aperture values than *C. canephora* clones under control conditions. Significant changes in stomatal aperture were observed between guard cells of D^T I59 and D^S Rubi after 24h and 72h of ABA treatment. The D^T I59 cultivar increases guard cells opening following 24h of ABA exogenous stimulation and further increases guard cells closing during the last 48h. None statistically significant changes in stomatal aperture of D^S Rubi guard cells were promoted with exogenous ABA stimulation. Clearly, exogenous ABA stimulation affected distinctively the stomatal control in guard cells of D^T *C. arabica* and *C. canephora* plants. Overall, guard cells of *C. canephora* clones dispose of larger stomatal aperture in natural conditions and responds to ABA stimulation by inducing stomatal closure in the first 24h but followed by stomatal reopening in the last 48h. Distinctively, in the first 24h guard cells of *C. arabica* D^T I59 induces stomatal opening upon ABA stimulation and further promotes stomatal closure in the 48h.

DISCUSSION

In this part of the work, we focused our attention to study the effects of exogenous ABA to affect the expression of *PYL-PP2C-SnRK2* genes of ABA tripartite in D^T and D^S genotypes of *C. canephora* and *C. arabica*. Among the nine *PYL* genes characterized in *C. canephora* (Chapter I), *PYR1*, *PYL4*, *PYL8a* and *PYL8b*, *PYL9* were the genes that presented the most relevant differences of expression profiles between *C. canephora* and *C. arabica* species, but also between D^T and D^S genotypes of the same coffee species and ABA treatments. Except for *PYL2*, the results presented here are in accordance with those described as expressed genes in leaves of *C. canephora* plants under I or NI conditions (Chapter I).

Regarding the first step of ABA tripartite system, we clearly highlighted that the D^S Rubi cultivar of *C. arabica* up-regulated the expression of *CcPYLs* genes latter (after 72 hours of ABA) compared to earlier responses observed for the same genes in other genotypes. Besides that, it is worth noting that *PYL9* gene expression was not detected in control or under ABA treatment only for *C. arabica* var. Rubi. It was recently suggested that *PYL9* promoted drought resistance not only by limiting transpiration water loss but also, by causing summer dormancy-like responses, such as senescence (Zhao *et al.*, 2016). In plants, leaf senescence increases the transfer of nutrients to developing and storage tissues. Moreover, transgenic tobacco showed that delayed leaf senescence increases plant resistance to drought (Rivero *et al.*, 2007). These evidences also corroborate with the physiological and molecular responses previously observed for the D^T and D^S *C. canephora* and *C. arabica* plants submitted to drought conditions (Pinheiro *et al.*, 2005; Marraccini *et al.*, 2011; Mofatto *et al.*, 2016). The *C. canephora* D^S clone 22 maintained the same expression levels of *PYL9* in control or ABA treatments while the D^T genotypes of *C. canephora* or *C. arabica* up-regulated *PYL9* expression in control conditions and ABA treatments.

It is important highlighting that the *CcPYR1* and *CcPYL8b* genes are highly up-regulated mainly in the *C. canephora* D^T clone 14 in a fast response (24 h to exogenous ABA. Previous microarray data and GUS expression studies have shown that *PYR1* and *PYL8* were expressed in guard cells (Gonzalez-Guzman *et al.*, 2012). We have previously shown (Chapter I) that *CcPYR1* was significantly down-regulated under drought in all clones of *C. canephora* except in D^T clone 14 that maintained similar expression levels in leaves under I or NI conditions. *CcPYL8b* expression levels also not presented significant difference between I or NI in clone 14. *Arabidopsis* transgenic *PYL8-OX* plants were generated and showed drought tolerance phenotype through enhanced stomatal closure in response to ABA (Lim *et al.*, 2013). HAB1 interacts with PYL8 and also with PYR1, however, the interaction with PYL8 was not ABA-dependent while with PYR1 did not occur in the absence of exogenous ABA in Y2H interaction (Santiago, *et al.* 2009; Park *et al.* 2009). Recent work showed that subcellular localization of PYL8 changes in response to ABA (Lee *et al.*, 2015). PYL8 protein moves into the

nucleus in response to ABA and the subcellular localization of PYL8 is regulated by abiotic stress signals. These result were also observed for PYL9 (Lee *et al.*, 2015).

Interestingly, under control conditions *ABI1* and *HAB* genes had higher expression levels in D^T clone 14 and Rubi cultivar of *C. canephora* and *C. arabica*, respectively. With ABA treatment, the clone 14 maintained expression levels of *CcHAB* gene at 24 and 72 h of ABA treatment. In contrast, I59 maintained *HAB* expression level during the first 24 h of ABA treatment, since its expression decreased at 72 h in this genotype. In contrast, in leaves of the *C. arabica* D^S Rubi, the *HAB* gene was upregulated after 24 h ABA treatment and the expression levels continue to increase at 72 h. The D^S clone 22 showed an uniform low expression of this gene from control to 72 hours ABA treatment.

HAB1 was originally cloned on the basis of sequence homology to *ABI1* and *ABI2*. In the case of *ABI1/ABI2*, the level of expression in response to ABA is notably higher for *ABI1* than *ABI2* (Saez *et al.*, 2003). This evidence was in accordance with our results where *CcABI1* was most expressed than *CcABI2* gene under ABA treatment *for C. arabica* plants. After 72 hours, the most expressed gene in Rubi was *CcABI1*. However, there was a peak of expression in this gene in Rubi after 24 hours under ABA treatment which suggests that this PP2C was highly expressed in this clone which could repress the transcription of kinases as *SnRK2.2* and *SnRK2.6*. In this sense, the drought-response genes could be later activated in ABA pathway.

It is known that the regulatory domain of SnRK2E/OST1/SnRK2.6 interacts with ABI1 and integrates abscisic acid and osmotic stress signals controlling stomatal closure in *Arabidopsis* (Yoshida *et al.*, 2006). It is worth noting that *CcSnRK2.6* was expressed only in leaves of the *C. canephora* clones after 72 h of ABA treatment. On the other hand, no expression was observed in *C. arabica* plants under control or ABA treatment. Regarding the *C. canephora* clones, it is worth noting that the D^S clone 22 present a basal regulation of this gene which was constant from control to ABA treatment.

These results are in accordance with the previous works in literature and also with the stomata measurements carried out during the hydroponic assay where there are significant differences among control and ABA 72 hours in the *C. canephora* D^T clone 14 but not in D^S clone 22. In the first 24h guard cells of *C. arabica* D^T I59 induced stomatal opening upon ABA stimulation and further promoted stomatal closure in the last 48h while no statistical differences were observed for Rubi. These evidences suggested that D^T clone 14 and I59 has been more efficient in the stomatal regulation under ABA exogenous treatment than D^S clone 22 or Rubi. Besides that, the absence of stomatal closure in response to ABA until 72h for Rubi is also in accordance with the delay in ABA signalling observed in gene expression analyses.

Finally, it is important to draw attention to the fact of *CcPYR1*, *CcPYL8b* and *CcSnRK2.7* and *CcSnRK2.11* were highly up-regulated in the D^T clone 14 and it suggest that they could act

synergistically in the ABA pathway as key agents in a drought-tolerant response. All those evidences could be used to select molecular markers to improve genotypes selection in field.

CONCLUSION

Altogether, the results presented herein showed the expression of genes maintained or activated preferentially in response to ABA hormone. The ABA responses from *C. canephora* plants revealed to be different to *C. arabica* genotypes. In *C. canephora*, the D^T clone 14 presented higher expression for the *AHG2*, *AHG3*, *HAI* (*PP2Cs*) and *SnRK2.7*, *SnRK2.11* (*SnRK2s*) compared to the D^S clone 22 under control conditions which suggest the existence of the tripartite system (mode off) ready to be activated in the D^T plants. With ABA (24h) it was observed a higher and faster expression of *PYR1*, *PYL8b*, *SnRK2.7*, *SnRK2.11* concomitant with a drastic decrease for *AHG2*, *ABI* and *ABI2* showing an activation of tripartite system (mode on). On the other hand, the D^S clone 22 in response to ABA (24h and 72h) could not activate the synthesis of new ABA receptors or kinases, on the contrary, it activated the synthesis of *AHG2* gene which coding a phosphatase that negatively control ABA pathway. All this evidences support the phenotype differences (e.g. stomatal control) observed for drought tolerance between the D^T clone 14 and the D^S 22 suggesting that it could be consequence of the differences observed in the expression profiles of *PYL-PP2C-SnRK2* genes.

In *C. arabica*, it was clearly that the D^T I59 had a faster response to ABA stimuli compared to the D^S Rubi. With 24h it was observed that the D^T I59 up-expressed the *PYL8a*, *ABI1* and *2*, *AHG3*, *HAI*, *SnRK2.8* and *SnRK2.11* genes. The expressed phosphatases inhibit the activity of the kinases which could explain the absence of stomatal closure responses in leaves of I59 at 24h (mode off). Besides that, in both 24h and 72h was not possible detected the expression profiles of the tripartite system in I59 which could explain the stomatal closure at 72h, suggesting that some genes could be up-regulated between 24h and 72 h in *C. arabica* I59.

Regarding *C. canephora* and *C. arabica*, it was observed that *PYL8a*, *ABI1*, *HAB*, *SnRK2.8*, *SnRK2.11* genes were up-regulated in the D^T (clone 14 and I59) compared to D^S (clone 22 and Rubi) plants under control conditions. In response to ABA treatment, *PYR1* and *HAI* were up-regulated after 24h while *PYL9* and *SnRK2.8* after 72h. Considering the differences between species, it was showed that *SnRK2.6* gene was expressed at 72h only in *C. canephora* plants (clones 14 and 22). It was also observed that *CcPYL9* was up-regulated in the *C. canephora* (clones 14 and 22) and in *C. arabica* (I59) all those presented significant stomatal closure in response to exogenous ABA.

REFERENCES

BARSALOBRES-CAVALLARI, C. F. *et al.* Identification of suitable internal control genes for expression studies in *Coffea arabica* under different experimental conditions. **BMC Molecular Biology**, London, v. 10, p.1, 2009.

DENOEUD, F. *et al.* The coffee genome provides insight into the convergent evolution of caffeine biosynthesis. **Science**, New York, v. 5, p. 1181-1184, 2014.

GONZALEZ-GUZMAN, M. *et al.* Arabidopsis PYR/PYL/RCAR receptors play a major role in quantitative regulation of stomatal aperture and transcriptional response to abscisic acid. **The Plant Cell**, Rockville, v. 24, n. 6, p. 2483-96, 2012.

GRAY, J. Guard cells: transcription factors regulate stomatal movements. **Current Biology**, London, v. 15, p. R593-R595, 2005.

HETHERINGTON, A. M.; WOODWARD, F. I. The role of stomata in sensing and driving environmental change. **Nature**, London, v. 424, p. 901-908, 2003.

HOAGLAND, D. R.; ARNON, D. I. **The water culture method for growing plants without soils**. Berkeley: The College of Agriculture University of California; California Agricultural Experimental Station, 1950. 32 p.

IMBER, D.; TAL, M. Phenotypic reversion of flacca, a wilty mutant of tomato, by abscisic acid. **Science**, New York, v. 169, p. 592-593, 1970.

JONES, R. J.; MANSFIELD, T. A. Suppression of stomatal opening in leaves treated with abscisic acid. **Journal of Experimental Botany**, Oxford, v. 21, p. 714-719, 1970.

KIM, T. H. *et al.* Guard cell signal transduction network: advances in understanding abscisic acid, CO 2 , and Ca 2+ Signaling. **Annual Review of Plant Biology**, Palo Alto, v. 61, p. 561-591, 2010.

LEE, H. Y. *et al.* The soluble ABA receptor PYL8 regulates drought resistance by controlling ABA signaling in Arabidopsis. **Plant Biotechnology Reports**, New York, v. 9, n. 5, p. 319-330, 2015.

LIM, C. W. *et al.* Arabidopsis PYL8 plays an important role for ABA signaling and drought stress responses. **The Plant Pathology Journal**, Seul, v. 29, n. 4, p. 471-476,. 2013.

MARRACCINI, P. *et al.* Differentially expressed genes and proteins upon drought acclimation in tolerant and sensitive genotypes of Coffea canephora. **Journal of Experimental Botany**, Oxford, v. 63, n. 11, p. 4191-4212, 2012.

MARRACCINI, P. *et al.* RBCS1 expression in coffee: coffea orthologs, Coffea arabica homeologs, and expression variability between genotypes and under drought stress RBCS1 expression in coffee: *Coffea* orthologs, *Coffea arabica* homeologs, and expression variability between. **BMC Plant Biology**, London, v. 11, p. 85, 2011.

MASLE, J.; GILMORE, S. R.; FARQUHAR, G. D. The ERECTA gene regulates plant transpiration efficiency in Arabidopsis. **Nature**, London, v. 436, p. 866-870, 2005.

MERILO, E. *et al.* PYR/RCAR receptors contribute to ozone-, reduced air humidity-, darkness-, and CO2-induced stomatal regulation. **Plant Physiology**, Bethesda, v. 162, n. 3, p. 1652-1668, 2013.

MOFATTO, L. S. *et al.* Identification of candidate genes for drought tolerance in coffee by high-throughput sequencing in the shoot apex of different *Coffea arabica* cultivars. **BMC Plant Biology**, London, v. 16, p. 94, 2016.

NILSON, S. E.; ASSMANN, S. M. The control of transpiration: insights from Arabidopsis. **Plant Physiology**, Bethesda, v. 143, p. 19-27, 2007.

PANTIN, F. *et al.* The dual effect of abscisic acid on stomata. **New Phytologist**, Cambridge, v. 197, n. 1, p. 65-72, 2013.

PARK, S. Y. *et al.* Abscisic acid inhibits PP2Cs via the PYR/PYL family of ABA-binding START proteins. **Science**, New York, v. 324, p. 1068-1071, 2009.

PINHEIRO, H. A. *et al.* Drought tolerance is associated with rooting depth and stomatal control of water use in clones of *Coffea canephora*. **Annals of Botany**, London, v. 96, p. 101-108, 2005.

RIVERO, R. M. *et al.* Delayed leaf senescence induces extreme drought tolerance in a flowering plant. **Proceedings of the National Academy of Sciences**, Washington, v. 104, p. 19631-19636, 2007.

SAEZ, A. *et al.* Gain-of-function and loss-of-function phenotypes of the protein phosphatase 2C HAB1 reveal its role as a negative regulator of abscisic acid signalling. **Plant Journal**, Shannon, v. 37, n. 1, p. 354-369, 2003.

SANTIAGO, J. *et al.* The abscisic acid receptor PYR1 in complex with abscisic acid. **Nature**, London, v. 462, p. 665-668, 2009a.

SANTIAGO, J. *et al.* Modulation of drought resistance by the abscisic acid receptor PYL5 through inhibition of clade IIA PP2Cs. **The Plant Journal**, Oxford, v. 60, p. 575-588, 2009b.

SCHROEDER, J. I. *et al.* Guard cell signal transduction. **Annual Review of Plant Physiology and Plant Molecular Biology**, Palo Alto, v. 52, p. 627-658, 2001.

SHATIL-COHEN, A.; ATTIA, Z.; MOSHELION, M. Bundle-sheath cell regulation of xylem-mesophyll water transport via aquaporins under drought stress: a target of xylem-borne ABA? **The Plant Journal**, Shannon, v. 67, n. 1, p. 72-80, 2011.

SIRICHANDRA, C. *et al.* The guard cell as a single-cell model towards understanding drought tolerance and abscisic acid action. **Journal of Experimental Botany**, Oxford, v. 60, p. 1439-1463, 2009.

TAL, M.; IMBER, D.; ITAI, C. Abnormal stomatal behavior and hormonal imbalance in flacca, a wilty mutant of tomato: I., root effect and kinetin-like activity. **Plant Physiology**, Bethesda, v. 46, p. 367-72, 1970.

TARDIEU, F.; SIMONNEAU, T. Variability among species of stomatal control under fluctuating soil water status and evaporative demand: modelling isohydric and anisohydric behaviours. **Journal of Experimental Botany**, Oxford, v. 49, p. 419-432, 1998.

YOSHIDA, R. *et al.* The regulatory domain of SRK2E/OST1/SnRK2.6 interacts with ABI1 and integrates abscisic acid (ABA) and osmotic stress signals controlling stomatal closure in Arabidopsis. **Journal of Biological Chemistry**, Bethesda, v. 281, p. 5310-5318, 2006.

ZHAO, Y. *et al.* ABA receptor PYL9 promotes drought resistance and leaf senescence. **Proceedings of the National Academy of Sciences**, Washington, v. 113, n. 7, p. 1949-1954, 2016.

GENERAL CONCLUSION AND PERSPECTIVES

The results presented in this work are one of the first that use the data generated by *C. canephora* sequencing, recently published, to analyze gene families such as those that codify proteins belonging to the tripartite system of ABA perception and signal transduction pathway. Comparing to studies developed in other species, our results showed for the first time the existence of duplication event in the *PYL* gene-family, notable for *CcPYL8*.

The results of expression analyses allowed us to confirm that the majority of the selected genes are functional in leaves and roots tissues. Similarly, several works has evidenced the importance of ABA tripartite system genes to fruit maturation highlighting the importance of further studies characterizing the *PYL* gene-family expression during coffee seeds development.

Despite the relevant information assessed with exogenous ABA experiment where genes expressed in response to this phytohormone presented distinguished regulation profile (mode on) in the D^T clone 14 and (mode off) in the D^S clone 22, a similar hydroponic test with different ABA concentration (lower) and number of plants (higher) in a different timepoint could be of interest. The evaluation of homoeologous gene expression in *C. arabica* subgenomes could provide useful information on this species plasticity to regulate ABA signaling and response pathways.

Even if our results did not present significative differences among clones regarding the amount of ABA in leaves and roots, it could be interesting to quantify this phytohormone under water deficit during a timecourse. Indeed, ABA quantification in plants with -3,0 MPA Ψ_{pd} value in stress condition (after 6 days watering withheld for the D^S clone 22, and between 12 and 15 days for the D^T clones, Marraccini et., 2011), did not allow to know if ABA content could variate in leaves and roots early after stress application. To verify that ABA metabolism is not altered in the different *C. canephora* clones, it could be also interesting to test the gene expression of *CcNCED3* and *CcCYP707A1*, which are respectively involved in synthesis and catabolism of ABA. This work is also underway in the laboratory (Costa *et al.*, manuscript in preparation).

The results presented in this study confirm those previously obtained (Vieira *et al.*, 2013) which showed that drought tolerance response in *C. canephora* is a result of several correlated mechanisms rather than a single one. In addition, it would be interesting to search for single nucleotide polymorphisms (SNPs: *single-nucleotide polymorphisms* and indels: *INsertion*/*DELetion*) in the genes identified in this work, for example, in the genomes of D^T clones (14, 73 and 120) and D^S (22) of *C. canephora* since these are sequenced (AC Andrade, personal communication).

This research could be conducted both in the coding sequence, to search for proteins modifications in the tripartite system genes of D^T and D^S clones used in this work, and within their regulatory sequences (promoters) to verify the occurrence of sequence variations in *cis*-regulatory

elements that could explain the different expression profiles observed for some genes in D^T and D^S clones, as has recently been observed for *CcDREB1D* gene of *C. canephora* (Alves *et al.*, submitted).

Finally, those genes with higher correlated drought-induced expression identified during this work (e.g *CcAHG2* and *CcSnRK2.2*) could be tested in other *C. canephora* drought tolerant and sensitive clones (Carneiro *et al.*, 2015) to find out if their expression profiles are kept. If that is the case, then one might consider using them as molecular markers in the coffee breeding programs for the generation of new drought tolerant varieties.

ANNEX: ARTICLES PUBLISHED DURING THE PHD

Plant Mol Biol (2014) 85:11–31
DOI 10.1007/s11103-013-0166-5

Lipid transfer proteins in coffee: isolation of *Coffea* orthologs, *Coffea arabica* homeologs, expression during coffee fruit development and promoter analysis in transgenic tobacco plants

Michelle G. Cotta · Leila M. G. Barros · Juliana D. de Almeida · Fréderic de Lamotte · Eder A. Barbosa · Natalia G. Vieira · Gabriel S. C. Alves · Felipe Vinecky · Alan C. Andrade · Pierre Marraccini

Received: 10 September 2013 / Accepted: 6 December 2013 / Published online: 28 January 2014

Abstract The aim of the present study was to perform a genomic analysis of non-specific lipid-transfer proteins (nsLTPs) in coffee. Several nsLTPs-encoding cDNA and gene sequences were cloned from *Coffea arabica* and *Coffea canephora* species. In this work, their analyses revealed that coffee nsLTPs belong to Type II LTP characterized under their mature forms by a molecular weight of around 7.3 kDa, a basic isoelectric points of 8.5 and the presence of typical CXC pattern, with X being an hydrophobic residue facing towards the hydrophobic cavity. Even if several single nucleotide polymorphisms were identified in these nsLTP-coding sequences, 3D predictions showed that they do not have a significant impact on protein functions. Northern blot and RT-qPCR experiments revealed specific expression of Type II nsLTPs-encoding genes in coffee fruits, mainly during the early development of endosperm of both *C. arabica* and *C. canephora*. As part of our search for tissue-specific promoters in coffee, an *nsLTP* promoter region of around 1.2 kb was isolated. It contained several DNA repeats including boxes identified as essential for grain specific expression in other plants. The whole fragment, and a series of 5′ deletions, were fused to the reporter gene *β*-glucuronidase (*uidA*) and analyzed in transgenic *Nicotiana tabacum* plants. Histochemical and fluorimetric GUS assays showed that the shorter (345 bp) and medium (827 bp) fragments of *nsLTP* promoter function as grain-specific promoters in transgenic tobacco plants.

Keywords Bean development · *Coffea* · Endosperm-specific promoter · Gene expression · Lipid transfer proteins

M. G. Cotta · L. M. G. Barros · J. D. de Almeida · E. A. Barbosa · N. G. Vieira · G. S. C. Alves · F. Vinecky · A. C. Andrade · P. Marraccini
Embrapa Genetic Resources and Biotechnology, Parque Estação Biológica, CP 02372, Brasília, DF 70770-917, Brazil

M. G. Cotta · N. G. Vieira · G. S. C. Alves
University of Lavras, Campus Universitário, CP 3037, Lavras, MG 37200-000, Brazil

F. de Lamotte
INRA UMR AGAP, 34060 Montpellier, France

E. A. Barbosa · F. Vinecky
University of Brasilia, Campus Universitário Darcy Ribeiro, Brasília, DF 70910-900, Brazil

P. Marraccini (✉)
CIRAD UMR AGAP, 34398 Montpellier, France
e-mail: marraccini@cirad.fr

Introduction

Lipid-transfer proteins (LTPs) are characterized by their ability to bind fatty acids and to transfer in vitro lipids (e.g. phospholipids, cholesterol) between membranes (Kader 1996). As LTPs can associate with various phospholipids with broad specificity, these proteins are more referred to nsLTPs for non-specific lipid transfer proteins (Ostergaard et al. 1993). Plant nsLTPs have been purified from various sources of plant tissue (e.g. leaves and seedlings) and are characterized by small molecular weights (usually ≈6.5–10 kDa) and basic isoelectric points (pI) ranging between 8.8 and 12 (Kader 1997). In *Brassica oleracea*, nsLTPs were found associated with the waxy surface of the leaves and expressed at high levels accounting for 50 % of proteins in young leaves (Pyee et al. 1994). In addition to this role in mediating phospholipid transfer, nsLTPs may also be involved in other biological functions such as plant defense mechanisms against fungal and

bacterial pathogens (Molina et al. 1993; Kristensen et al. 2000), and may participate in the assembly of hydrophobic protective layers of surface polymers such as cutin or waxes (Cameron et al. 2006; Yeats and Rose 2008). Several studies have highlighted that the expression of nsLTP genes is also induced in response to environmental stresses such as drought (Trevino and O'Connell 1998; Rizhsky et al. 2004; Tapia et al. 2013) and abscisic acid (Garcia-Garrido et al. 1998).

The first cDNA encoding a plant LTP was isolated from maize seedlings (Tchang et al. 1988). In the same plant, it is common to observe several LTP-encoding cDNAs suggesting the existence of a complex gene family (Kader 1997). A mechanism of alternative splicing might exist in the RNA coding for these proteins (Arondel et al. 1991). In terms of structure, nsLTPs consist of 91–95 amino acid residues differing in sequence but containing eight strictly conserved cysteine residues (Kader 1997). It is probably due to the four disulfide bridges formed between these residues (Takishima et al. 1986) that nsLTPs are remarkably stable after purification (Lindorff-Larsen and Winther 2001).

Regarding their classification, nsLTPs were divided into two main groups according to their molecular weight: Type I (9 kDa) and Type II (7 kDa) (Marion et al. 2000). In 2008, a new classification of nsLTPs was proposed using the putative mature form of 267 proteins from rice, wheat, and *Arabidopsis thaliana* and showed that the major functional diversification within the nsLTP family predated the monocot/dicot divergence (Boutrot et al. 2008). Genome wide analysis carried out by these authors allowed to identify eight types of LTPs, as well as 33 subfamilies. Type I LTPs were the more abundant and displayed a specific disulfide bond pattern different from Type II constituting the second more abundant type. Molecular studies revealed complex expression patterns for the various types (Boutrot et al. 2005) and work is still in progress to decipher the specificity of these different nsLTP Types. More recently, Wang et al. (2012) classified nsLTPs into five different types based on the sequence similarity matrix and the properties of their 8-cysteine motifs and showed that Types I and II are shared by all the plant species possessing nsLTPs. Regarding the 517 plant nsLTPs analyzed by these authors, 391 (from 88 species) and 102 (from 23 species) were classified in Types I and II, respectively. However, no relationship between proteins of the same Type and the function was clearly established and the precise physiological roles of plant nsLTPs are still a matter of debate. For example, Type I nsLTP from maize, spinach, arabidopsis, radish, onion, and broccoli exhibit antimicrobial activity (García-Olmedo et al. 1995). This was also shown for the Cc-LTP1 purified from *Coffea canephora* seeds that exhibited strong antifungal activity (Zottich et al. 2011). On the other hand, probable role of Type II nsLTPs as plant defensins were also reported in rice (Samuel et al. 2002) and tobacco leaves (Harada et al. 2010). Type I nsLTPs found primarily in aerial organs, nsLTP1 is proposed to transport cutin monomers whereas Type 2 mainly found in roots, is involved in the transport of suberin monomers (Samuel et al. 2002).

Expression patterns of plant *nsLTP* genes are usually complex and controlled temporally and spatially. Depending on the gene considered, expression is often detected in the aerial portions of plants (leaves, stems, shoot meristems) as well as in inflorescences, but also early in development, such as in embryo cotyledons and leaf primordia of *A. thaliana* (Thoma et al. 1994). In addition, Fleming et al. (1992) showed that *LTP* gene expression was higher in young tobacco leaves than in old ones, but also higher in the upper parts of the plant compared to the basal parts, suggesting that nsLTP-encoding genes were expressed according to a developmental gradient. However, no *nsLTP* transcripts were detected in the roots of various plants. The expression of nsLTP-encoding genes has also been analyzed in seeds, such as those of barley, for example, where expression was well detected in the aleurone layer, which is rich in lipid bodies (Jakobsen et al. 1989).

The tissue-specificity of the *nsLTP* promoter region was also investigated by transgenic plant assays involving promoter fusions to the *GUS* reporter gene. For example, the promoters of *ltp1* and *ltp2* genes from barley were able to direct aleurone-specific expression in barley and rice seeds, respectively (Kalla et al. 1994; Skriver et al. 1992). The sequence analyses of these promoters detected the presence of MYB and MYC protein binding sites (Linnestad et al. 1991), like those also found within the promoter region of the strawberry *Fxaltp* gene (Yubero-Serrano et al. 2003).

Despite the fact that coffee is one of the most important agricultural commodities in the world, basic knowledge is missing regarding many aspects of this crop, particularly lipid metabolism during bean development, especially considering the importance of those compounds in organoleptic features (Leroy et al. 2006). In the genus *Coffea*, two species account for almost all coffee bean production: *Coffea canephora* and *C. arabica*. *C. canephora* is diploid ($2n = 2x = 22$) and allogamous while *C. arabica* is amphidiploid (allotetraploid, $2n = 4x = 44$) and autogamous. As *C. arabica* arose from natural hybridization between *C. canephora* and *C. eugenioides*, its transcriptome is a mixture of homeologous genes expressed from these two sub-genomes (Vidal et al. 2010). In these two coffee species, the lipid content of green coffee beans are 15 and 10 %, respectively, and mainly consists of triacylglycerols, sterols, tocopherols and diterpens (Speer and Kölling-Speer 2006). Most of these lipids are located in

the endosperm of green coffee beans but a small amount is also present on the outer layer of the bean (Wilson et al. 1997). It is known that the lipid fraction of the beans is little changed during roasting and protects the aroma compounds from degradation (Folstar 1985). Even though the effect of fat contents on the sensory quality of coffees remains to be determined, several studies have highlighted that the increase in fat content with shading and altitude is positively correlated with cup quality (Decazy et al. 2003; Vaast et al. 2006).

In terms of lipid synthesis, the expression burst of genes involved in lipid assembly and storage has been observed at mid-stages of bean development (Joët et al. 2009). This supports the observations by electronic microscopy of oil body accumulation in endosperm cells at 110–150 days after flowering (DAF) (Dentan 1985).

Using the recent advances in coffee genomics (De Kochko et al. 2010), our study set out to (1) identify the different coffee *nsLTP* homeologs corresponding to the *C. canephora* and *C. eugenioides* ancestor sub-genomes of *C. arabica* and (2) evaluate the expression of these alleles during bean development. We also report on the cloning of an *nsLTP* promoter that was tested in transgenic tobacco for analyzing its ability to control the expression of the *uidA* reporter gene in seeds.

Methods

Plant materials

Three cultivars of *Coffea arabica* (IAPAR59, Catuaí Amarelo and Mundo Novo) and one clone (L6P35) of *C. canephora* conilon were used in this study. Eight-year-old plants of IAPAR59 and the 3-years-old clone L6P35 of *C. canephora* conilon cultivated under field conditions at the experimental station of Embrapa Cerrado research center (Planaltina-DF, Brazil 15°35′43″S–47°43′52″O) in full sunlight were used for fruit expression studies. Fruits were collected between 1 pm and 4 pm regularly (every 4 weeks) after flowering (mid September) up to complete maturation (May ≈210 DAF) over 2 years (2006/2007 and 2008/2009 harvests) for *C. arabica* cv. IAPAR59 and from the flowering up to complete maturation (July ≈300 DAF, harvest 2011/2012) for the clone L6P35 of *C. canephora*. The cultivars of Catuaí Amarelo and Mundo Novo were in the coffee collection of Embrapa Genetic Resources and Biotechnology research center (Brasilia-DF, Brazil). For expression studies, leaves and roots were from *C. arabica* cv. Catuaí Amarelo. Tobacco (*Nicotiana tabacum* L. cv. Xanthi XHFD8) was grown in vitro (25 °C, 16-h photoperiod) on solid MS (Murashige and Skoog 1962) or in the greenhouse under environmental conditions.

RNA isolation

Total RNA was extracted from roots and leaves and from whole fruits at 120 DAF (2007/2008 harvest) of *C. arabica* cv. Catuaí Amarelo. Total RNA was also extracted from whole fruits (2008/2009 harvest), separated perisperm and endosperm (2006/2007 harvest) of *C. arabica* cv. IAPAR59 and separated endosperm of clone L6P35 of *C. canephora* conilon (2011/2012 harvest). After collected, all samples were immediately frozen in liquid nitrogen and stored at −80 °C before being ground and treated as described previously (Marraccini et al. 2011). RNA quantification was performed using a NanoDrop™ 1000 Spectrophotometer (Waltham, MA, USA).

DNA extraction

To isolate the *nsLTP* genes from *C. arabica* cv. Mundo Novo and from the clone 120 of *C. canephora* (Marraccini et al. 2012) and promoter from *C. arabica* cv. Catuaí Amarelo, genomic DNA was extracted from leaves according to the Doyle and Doyle (1987) method modified as follows. Once ground in liquid nitrogen, 2.5 g of young leaves was transferred and mixed with 20 mL of extraction buffer (1.4 M NaCl, 100 mM Tris–HCl, 2 % CTAB, 0.05 M EDTA, 1 % PVP, 1 % β-mercaptoethanol) pre-heated to 65 °C and incubated for 1 h. One volume of phenolic acid (pH 8):chloroform:isoamyl alcohol (25:24:1) was then added before removal of cell debris by centrifugation (10 min, 15,000*g*). The aqueous phase was further mixed with one volume of chloroform:isoamyl alcohol (24:1) and centrifuged. The upper phase was then treated by an equal volume of isopropanol in order to precipitate nucleic acids by centrifugation (30 min, 15,000*g*) which were resuspended in sterile water and incubated (30 min, 37 °C) with 200 µg of RNaseA. DNA concentration and quality were determined using a NanoDrop™ 1000 Spectrophotometer (Waltham, MA, USA).

Northern-blot experiment

Twenty micrograms of total RNA was fractionated on a 1.2 % (w/v) agarose gel containing 2.2 M formaldehyde in MOPS buffer. Equal amounts of loaded RNA samples were controlled by the abundance of 26S and 18S rRNA on gels stained with ethidium bromide. The *nsLTP* internal cDNA probe was amplified by conventional PCR reaction using the LTP-F3 and LTP-F4 primers (Table 1) and the coffee EST GT669102 as a template. This probe of 117 bp in length (from nucleotides 62–178 of contig22413, Fig. 1) was further labeled by random-priming with α-^{32}P-dCTP as described by the supplier (GE Healthcare). Total RNAs were transferred to Hybond N+ membranes which

Table 1 List of primers

Primers	Sequences
GSP1	5′-CAGATCCACCAGCAACAGTACAACC-3′
GSP2	5′-CAGTGCAACCCCAGATGATTTCTTC-3′
LTP-F1	5′-GCGCTTTTTGCTTTTCATAAAGAT-3′
LTP-F2	5′-GCACTTTTTGCTTTTCATAATGATG-3′
LTP-F3	5′-GAAATCATCTGGGGTTGCAC-3′
LTP-F4	5′-AAGCATGGACTCAATGCTTG-3′
LTP-R1	5′-ATTCAACACCATTACTAGTTTTCGAGC-3′
LTP-R2	5′-CACCATTACATGGGAACGTTGC-3′
LTP-FT	5′-CTGTGGTCTGAAATGGCCAACT-3′
BUBI-F	5′-AAGACAGCTTCAACAGAGTACAGCAT-3′
BUBI-R	5′-GGCAGGACCTTGGCTGACTATA-3′
F1-pBI	5′-***CGCAAGCTT***CATCCTAAAATACATTCG-3′
F2-pBI	5′-***CGCAAGCTT***TCCATGAAAAATGCAATCC-3′
F3-pBI	5′-***CGCAAGCTT***CCAAGACATTATTAATGATC-3′
F4-pBI	5′-***CGCAAGCTT***CTCCCACTTCTCAAAACTTGG-3′
R1-pBI	5′-***CGCGGATCC***GAAAAGCAAAAAGTGCAGAAGAG-3′
FORmax	5′-GCCAGGGTTTTCCCAGTCACGACGTTGTAA-3′
REVmax	5′-CACACAGGAAACAGCTATGACCATGATTA-3′

GSP1 and GSP2 primers were used during the genome walking experiment to amplify *nsLTP* promoter fragments. LTP-primers were used to clone nsLTP-encoding nucleic sequences and for qPCR experiments. The BUBI-F/R primer pair was used for the ubiquitine (*UBI*) reference gene in RT-qPCR experiments. Primers (pBI) were used for the construction of transformation vectors. The sequences (bold and italics) added to the 5′ end of F-pBI primers corresponded to the *HindIII* restriction sites and the one added to the 5′ end of the R1-pBI primer corresponded to the *BamHI* restriction site. The FORmax and REVmax primers used for DNA sequencing were also indicated

were further hybridized at 65 °C in modified Church and Gilbert buffer (7 % SDS, 1 % BSA, 10 mM EDTA, 0.5 M sodium phosphate monobasic pH 7.2) and washed at 65 °C in 2 × standard saline citrate (SSC; 1× = 150 mM sodium chloride and 15 mM sodium citrate, pH 7.0)—0.1 % SDS (2 × 15 min) with a final stringent wash in 0.1 × SSC–0.1 % SDS (2 × 15 min). Membranes were exposed with BAS-MS 2340 IP support and data were acquired using an FLA-3000 Fluorescent Image Analyzer (Fujifilm Life Science).

Real-time RT-PCR assays

To eliminate contaminant genomic DNA, RNA samples were treated with RQ1 RNase-free DNase according to the manufacturer's instructions (Promega, Madison, WI, USA) and RNA quality was verified by agarose gel electrophoresis for visual inspection of the ribosomal RNA bands upon ethidium bromide staining. First strand cDNA was synthesized by treating 1 µg of total RNA with the ImProm-II™ Reverse Transcription System with oligos (dT_{15}) according to the manufacturer's recommendations (Promega). The absence of contaminating genomic DNA was checked as previously described (Marraccini et al. 2011). Quantitative PCR was carried out with synthesized single-strand cDNA described above using the protocol recommended for the use of 7500 Fast Real-Time PCR Systems (Applied Biosystems, Foster City, CA, USA). cDNA preparations were diluted (1/25–1/100) and tested by qPCR. Primers (Table 1) were designed using Primer Express software (Applied Biosystems) and preliminarily tested for their specificity and efficiency against a mix of cDNA extracted from fruits of *C. arabica* cv. IAPAR59 (data not shown). qPCR was performed with 1 µL of diluted ss-cDNA and 0.2 µM (final concentration) of each primer in a final volume of 10 µL with SYBR green fluorochrome (SYBRGreen qPCR Mix-UDG/ROX, Invitrogen). The reaction mixture was incubated for 2 min at 50 °C (Uracil DNA-Glycosilase treatment), then 5 min at 95 °C (inactivation of UDGase), followed by 40 amplification cycles of 3 s at 95 °C and 30 s at 60 °C (annealing and elongation). Data were analyzed using SDS 2.1 software (Applied Biosystems) to determine cycle threshold (Ct) values. The specificity of the PCR products generated for each set of primers was verified by analyzing the Tm (dissociation) of amplified products. PCR efficiency (E) was estimated using absolute fluorescence data captured during the exponential phase of amplification of each reaction with the equation $(1 + E) = 10(-1/^{slope})$ (Ramakers et al. 2003). Efficiency values were taken into account in all subsequent calculations. Expression levels were calculated by applying the formula $(1 + E)^{-\Delta\Delta Ct}$ where $\Delta Ct_{\text{ target}} = Ct_{\text{ target gene}} - Ct_{\text{ CcUBQ10}}$ and $\Delta\Delta Ct = \Delta Ct_{\text{ target}} - \Delta Ct_{\text{ reference sample}}$, the perisperm-60 DAF and endosperm 150 DAF being used as reference samples in *C. arabica* and *C. canephora* experiments, respectively. Gene expression levels were normalized (SDS 2.1 software) with the expression ubiquitin gene as endogenous control (GW488515; Cruz et al. 2009; Marraccini et al. 2012).

Isolation of nsLTP-encoding cDNA and gene sequences

The *CaLTP1a*, *CaLTP2* and *CaLTP3a* cDNA sequences were obtained by PCR. The template used was 10 ng of a cDNA mixture from perisperm, endosperm and pericarp tissues separated from fruits of *C. arabica* cv. IAPAR59 harvested at different maturation times (from 30 to 210 DAF). The PCR reaction (PTC-100 Thermocycler, MJ Research) was performed using primer combinations LTP-F2/LTP-R2, LTP-F1/LTP-R2 and LTP-F1/LTP-R1 and *Taq* Platinum DNA polymerase according to the supplier's instructions (Invitrogen) under the following conditions: initial denaturation (94 °C-2 min) followed by 40 cycles (94 °C-30 s, 60 °C-30 s, 72 °C-3 min) and a final

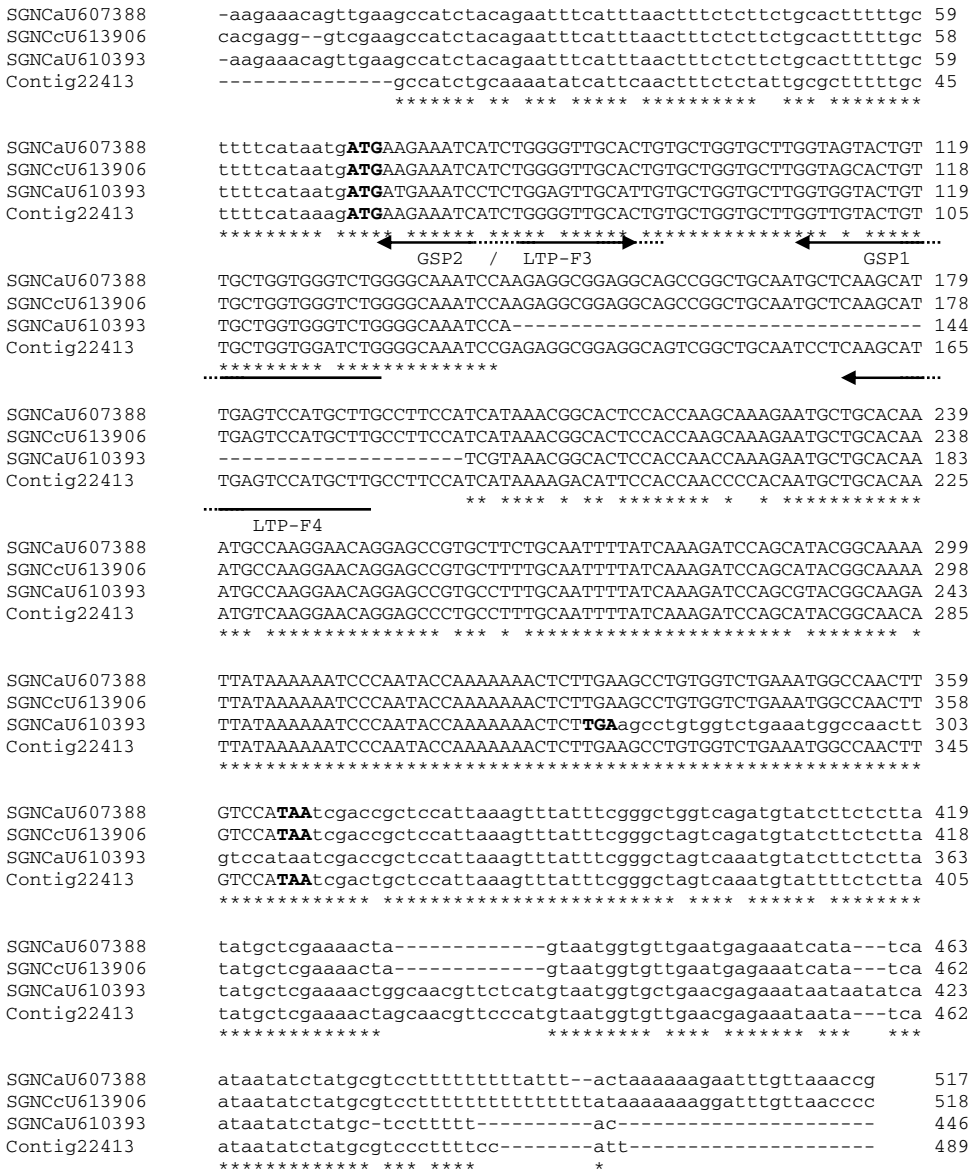


Fig. 1 Alignment of nsLTP-encoding nucleic sequences found in public databases. Contig22413 was found in the database of the Coffee Genome Project (http://bioinfo04.ibi.unicamp.br). Contigs SGNCaU607388, SGNCcU613906 and SGNCaU610393 were found in the SOL genomic database (http://solgenomics.net). The nsLTP-encoding sequences are in upper case, the 5′ and 3′ UTR regions in *lower case* and the start and stop codons in *bold*. The *nsLTP* cDNA probe was amplified using the LTP-F3 and LTP-F4 primers (Table 1) which are indicated as well as the GSP1 and GSP2 primers used during the genome walking experiment to amplify *nsLTP* promoter fragments. The *stars* below the alignments indicate identical bases and the nucleotides are numbered (*right*) on each *lane*

extension step (72 °C-7 min). Amplicon quality was verified by electrophoresis. PCR fragments were cloned in the pCR2.1TOPO (Invitrogen) vector and amplified in *Escherichia coli* TOP10 cells (Invitrogen). For each PCR reaction, two recombinant plasmids were extracted independently and double-strand sequenced. The *CaLTP1a*, *CaLTP1b*, *CaLTP2* and *CaLTP3a* genes were amplified by PCR from genomic DNA (10 ng) of *C. arabica* cv. Mundo Novo and the *CcLTP3* gene from clone 120 of *C. canephora* conilon. The primer combinations and the PCR conditions corresponded to those described before to isolate the *nsLTP* cDNAs. The fragments obtained were cloned and, for each primer combination, five recombinant plasmids were double-strand sequenced and further analyzed.

Isolation of *nsLTP* promoter and plasmid constructions for tobacco transformation

The *nsLTP* promoter was cloned from *C. arabica* cv. Catuaí Amarelo using the 5′ RACE strategy, combined with

a nested PCR approach according to the recommendation of the supplier (Genome Walker Universal Kit, Clontech). The GSP1 and GSP2 primers (Table 1) used were designed using Primer3 software (http://frodo.wi.mit.edu). This led to the amplification of three fragments that were sequenced and used to design the pBI-primers (Table 1) that enabled the amplification of *nsLTP* promoter fragments by PCR reactions carried out using 5 ng of genomic DNA, the forward (F-pBI) and R1-pBI primers (0.2 μM final concentration) and *Pfu* DNA polymerase under the following conditions: initial denaturation (94 °C-1 min) followed by 30 cycles (94 °C-30 s, 51 °C-30 s, 72 °C-2 min) and a final extension step (72 °C-7 min). The *HindIII* and *BamHI* restriction sites were included in the 5′-end of the F-pBI and R1-pBI primers, respectively. Amplified DNA fragments were purified from agarose gel by the Wizard®SV Gel and Clean Up System (Promega), double-digested with *HindIII* and *BamHI* and further ligated into the pBI121 (Clontech) vector previously cut by the same enzymes. Following ligation and *E. coli* transformation, the vectors here called pCaLTP-S (F4-pBI/R1-pBI, 345 bp), pCaLTP-M1 (F3-pBI/R1-pBI, 827 bp), pCaLTP-M2 (F2-pBI/R1-pBI, 1,047 bp) and pCaLTP-L (F1-pBI/R1-pBI, 1,252 bp) were obtained. For all these constructs, *nsLTP* promoter fragments were sequenced to certify that they were identical to the original promoter.

Genetic transformation and analysis of transgenic tobacco plants

The pCaLTP-S, pCaLTP-M1, pCaLTP-M2 and pCaLTP-L vectors, as well as the pCaMV35S positive control vector, were introduced independently into the disarmed strain *Agrobacterium tumefaciens* C58pMP90 as previously described by An et al. (1993). The genetic transformation of *N. tabacum* was accomplished according to Horsch et al. (1993). After transformation, around 20 independent transformants were regenerated for each construct and self-fertilized. The seeds were aseptically sown in MS medium containing 100 mg L^{-1} of kanamycin sulfate to identify the tobacco containing a unique locus of T-DNA insertion by measuring the frequency of kanamycin-resistant plants among the T1 progenies (data not shown).

Histochemical GUS assays

The histochemical GUS assay was performed with slices of leaves and roots, floral organs and seeds of transformed plants. The samples were incubated overnight at 37 °C in an X-Gluc (5-bromo-4-choloro-3-indolyl-β-D-glucuronide) solution (100 mM NaH_2PO_4, 500 mM EDTA, 0.1 % Triton X-100, 0.5 mM $K_3Fe(CN)_6$, 1 mM X-Gluc solubilized in DMSO) for blue color development. After staining, the sections were kept in 70 % ethanol until chlorophyll removal, and then photographed under a stereo microscope (Zeiss).

Fluorometric GUS assays

Quantitative determination of GUS activity was accomplished by fluorometric GUS assay (Jefferson et al. 1987) from transgenic plants containing a unique locus of T-DNA insertion. Leaves (~100 mg) and mature seeds (~50 mg) of self-fertilized T0 tobacco plants were ground in 500 μL of extraction buffer (100 mM sodium phosphate pH 7, 10 mM Na_2EDTA, 0.1 % sarkosyl, 0.1 %, Triton X-100, 1 mM DTT). Protein concentrations were determined as described by Bradford (1976) using a Bio-Rad kit and BSA as the standard. Protein extracts (50–100 μL) were incubated in extraction buffer containing 1 mM MUG (4-methylumbelliferyl-β-D-glucuronide) in a 200 μL final reaction mixture. Fluorescence was measured at 15 min intervals for 60 min using a VersaFluor fluorometer (Bio-Rad). A standard curve for 4-methylumbelliferone (MU) in extraction buffer was used to convert levels of fluorescence into mmol MU g^{-1} protein min^{-1}.

DNA sequencing and analysis

All DNAs isolated during this work were cloned in the pCR2.1TOPO (Invitrogen) vector and double-strand sequenced using FORmax and REVmax primers (Table 1) related to M13 For/Rev universal primers and BigDye Terminator Sequencing Kit v3.1 chemistry on an ABI 3130xl Genetic Analyzer (Applied Biosystems). Multiple alignments were made using the CLUSTALW program (Thompson et al. 1994) and nucleic and protein sequences found in public databases such as the SOL Genomics Network (SGN, http://solgenomics.net/content/coffee.pl; Mueller et al. 2005), Harvest (http://harvest.ucr.edu/; Lin et al. 2005) and the Coffee Genome Project (http://www.lge.ibi.unicamp.br/cafe/; Mondego et al. 2011). The TargetP program (Emanuelsson et al. 2007) was used to search for putative signal peptides. DNA motifs were sought using the PlantPAN (http://PlantPAN.mbc.nctu.edu.tw, Chang et al. 2008) and the TSSP/Prediction of Plant Promoters (SoftBerry: http://www.softberry.com, Shahmuradov et al. 2003) web interfaces.

Phylogenetic analysis and 3D modeling for coffee nsLTPs

All analyses were carried out on the South Green Bioinformatics Platform (SGBP: http://southgreen.cirad.fr) using Galaxy (Giardine et al. 2005). The Coffee nsLTPs protein sequences identified in this work together with others 216 nsLTPs plant protein sequences from *A. thaliana,* wheat

Fig. 2 Alignment of coffee nsLTP-encoding genes and cDNA sequences from *C. arabica* and *C. canephora*. The genes *CaLTP1a* (HG323818) and *CaLTP1b* (HG323819) amplified with primer pair LTP-F2/LTP-R2, *CaLTP2* (HG323820) with LTP-F1/LTP-R2 and *CaLTP3b* (HG323821) with LTP-F1/LTP-R1 were from *C. arabica* cv. Mundo Novo. The *CaLTP1a* (HG008739), *CaLTP2* (HG008740) and *CaLTP3a* (HG008741) cDNA sequences were from cloned fruits of *C. arabica* cv. IAPAR59. The *CcLTP3* (HG323822) gene was from clone 120 of *C. canephora* conilon. The nsLTP-encoding sequence is in upper case, the 5′ and 3′ UTR regions in *lower case* and the start and stop codons in *bold*. Identical genes and cDNA sequences are identified by (*equal to*). The *stars* below the alignments indicate identical bases and the nucleotides are numbered (*right*) on each *lane*. Nucleotides diverging between the sequences are *boxed* in *gray*. *Horizontal arrows* indicate primers (Table 1) used to amplify *LTP* cDNAs and genes and to perform qPCR experiments

and rice (Boutrot et al. 2008) were aligned using MAFFT program (Katoh and Toh 2008). In order to keep only their reliable parts, these sequences were filtered with Gblocks (Castresana 2000). These 223 sequences were used to construct a phylogenetic tree using PhyML algorithm (Guindon et al. 2010). Trees were reconciliated with RAP-Green algorithm (Dufayard et al. 2005) for the comparison of the gene tree with the species tree. The reference species tree used is the one provided by the NCBI taxonomic database: http://www.ncbi.nlm.nih.gov/sites/entrez?db=taxonomy. Dendroscope software (Huson et al. 2007) was used for visualization of nsLTPs tree. For modelisation 3D, @tome-2 platform (Pons and Labesse 2009) (http://atome.cbs.cnrs.fr) was utilized considering different homology modeling: PSI-Blast (Altschul et al. 1997), HHSearch (Söding 2005), Fugue (Shi et al. 2001), and SP3 (Zhou and Zhou 2005). PyMOL program (http://www.pymol.org/) was used for visualization and manipulation of 3D structures. SNAP method (Bromberg and Rost 2007) was used for prediction of the functional effects of non-synonymous SNPs.

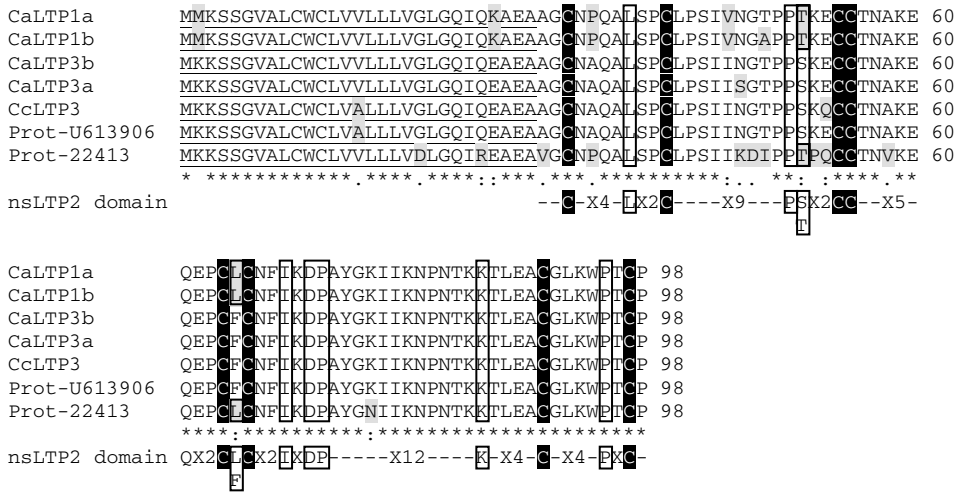
Results

Characterization of nsLTP-encoding cDNA and gene sequences

Electronic Northern and Fisher's statistical tests based on the Coffee Genome Project data (http://bioinfo04.ibi.unicamp.br) pointed to a unigene called contig22413 (Fig. 1) that is highly and specifically expressed in coffee fruits. This 489 bp sequence was formed by the alignment of 32 ESTs, all from the FV1, FV2 and FR4 (fruits at all stages of development) cDNA of *Coffea racemosa* libraries (Vieira et al. 2006). It contains a 5′ untranslated region (UTR) of 56 bases, a 3′ UTR of 136 bases and an open reading frame of 294 bases encoding for a putative nsLTP protein of 98 amino acids. In the GenBank database (release 191, 08/2012), more than 250 ESTs (E-values ranging from 0 to 1e-100) highly identical to contig22413 were found mainly from fruit cDNA libraries for *C. racemosa*, *C. arabica* (Moncada et al. unpublished) and *C. canephora* (Lin et al. 2005). Other searches in the SOL database (http://solgenomics.net) also identified (1) the *C. arabica* unigenes SGN-U607388 and SGN-U610393 formed by ESTs from fruits harvested at 15, 26 and 28 weeks after flowering (WAF) and (2) *C. canephora* SGN-U613906 formed by the clusterization of 34 ESTs from *C. canephora* fruits (perisperm and endosperm tissues) harvested at 18 and 30 WAF. Once aligned, despite a gap of 56 bp found in SGN-U610393, these sequences showed high identity to the putative nsLTP-coding regions. However, several divergences were observed in the UTRs, such as an insertion of 13 bp in the 3′ UTR of SGN-U610393 and contig22413 sequences that was not present in SGN-U607388 and SGN-U613906 sequences (Fig. 1).

The differences observed in silico between these nsLTP-encoding unigenes enabled the design of specific primers that led to the isolation of *CaLTP1a*, *CaLTP2* and *CaLTP3a* cDNAs using the primer combinations LTP-F2/LTP-R2, LTP-F1/LTP-R2 and LTP-F1/LTP-R1, respectively (Table 1), all expressed in fruits of *C. arabica* cv. IAPAR59 (Fig. 2). The same primer combinations were also used to amplify corresponding nsLTP-encoding genes from *C. arabica* cv. Mundo Novo and *C. canephora*. From *C. arabica*, four sequences were isolated: *CaLTP1a*

(A)



(B)

	Pre-protein (98 aa)		Mature protein (69 aa)	
	MW	pI	MW	pI
CaLTP1a	10434.44	8.56	7390.64	8.48
CaLTP1b	10404.41	8.56	7360.62	8.48
CaLTP3b	10440.34	8.34	7398.62	8.48
CaLTP3a	10413.31	8.34	7371.60	8.48
CcLTP3	10411.30	8.56	7397.64	8.69
Prot-U613906	10412.29	8.34	7398.62 (2)	8.48
Prot-22413	10626.64	8.04	7498.83	8.17

Fig. 3 Sequence alignment and characteristics of the coffee nsLTP proteins. **A** The amino acids corresponding to the putative signal peptide (1–29) are *underlined*. CaLTP1a (CDF66370), CaLTP1b (CDG03097), CaLTP3a (CDF66372), CaLTP3b (CDG03099) and CcLTP3 (CDG03100) were deduced from the corresponding nucleic sequences presented in Fig. 2. The proteins Prot-U613906 and Prot-22413 were deduced from the corresponding contigs presented in Fig. 1. Below the alignment, identical amino acids are indicated by stars, conservative substitutions are indicated by two vertically *stacked dots* and semi-conservative substitutions are indicated by *single dots*. Divergent amino acids between nsLTP proteins are also underlined in gray. The nsLTP2 domain (cd01959) is also indicated and amino acids of nsLTP matching with this domain are boxed. **B** Characteristics of coffee nsLTPs: molecular weights (MW in Daltons), amino acids (aa) and isoelectric points (pI) are indicated for pre-proteins and mature proteins (without the signal peptide). The CaLTP2 (CDF66371) protein identical to the CaLTP1a was not represented

and *CaLTP1b* amplified with the LTP-F2/LTP-R2 primers, *CaLTP2* amplified with the LTP-F1/LTP-R2 primers and *CaLTP3b* with the LTP-F1/LTP-R1 primers. From *C. canephora*, the primer pair LTP-F1/LTP-R1 was the only one able to function and led to the isolation of *CcLTP3* gene. Nucleic alignments of these sequences demonstrated that the *CaLTP1a* cDNA and gene were strictly identical to *CaLTP2* and that the *CaLTP1a* and *CaLTP1b* genes differed by only one base (in position 164) (Fig. 2). In addition, *CaLTP2* cDNA and gene sequences were also strictly identical but diverged from *CaLTP1a* by only 2 bases in the 5′ region corresponding to the annealing of primers LTP-F1 and -F2 (positions 3 and 21, respectively) (Fig. 2). The LTP3-encoding sequences were clearly grouped together and characterized by 13-bp changes, and were very well conserved in their corresponding nsLTP-encoding regions which distinguished them from the LTP1 and LTP2 encoding regions. In addition, *LTP3* and *LTP1–LTP2* sequences also diverged by an insertion/deletion of 13 bp in their 3′ UTR regions as observed when aligning *nsLTP* unigenes.

Characterization of coffee nsLTP proteins

The proteins deduced from nsLTP cDNAs and genes were aligned to be compared (Fig. 3A). Apart from the protein deduced from contig SGNCaU610393, which was shorter in its C-ter region than the other nsLTP, CaLTP2 was identical to CaLTP1a, and the protein deduced from contig SGNCaU607388 was identical to CaLTP3b. All these proteins had the same length (98 amino acids) with a similar

theoretical molecular mass (≈10.4 kDa) and estimated isoelectric point (pI ≈ 8.5) (Fig. 3B). They also contained a putative signal peptide with a cleavage site between the amino acid residues A_{29} and V_{30}. As pre-protein, the CaLTP1a (deduced from *CaLTP1a* cDNA and gene) and CaLTP2 (deduced from *CaLTP2* cDNA and gene) were strictly identical (not shown), The CaLTP1a and CaLTP1b proteins diverged by only one amino acid residue in position 48 of the unprocessed proteins. Strict identity was observed between the proteins CaLTP3b and P-U607388. Only one amino acid residue (in position 53) differed between the CcLTP3 and P-U613906 proteins. Within the coffee nsLTPs studied, the P-22413 protein was the only one mostly diverging from the others.

In their processed forms (without the putative signal peptide), all these nsLTPs should contain 69 amino acids and have a similar theoretical molecular mass (≈7.4 kDa) and an estimated pI between 8.17 and 8.69 (Fig. 3B). The processed forms of CaLTP3b and P-U613906 proteins also appeared identical. Searches for sequence similarities with the Blastp program against the GenBank data base were also performed. For example, CaLTP1a had the highest similarity (e-value: 4e-19, identity 42 %, similarity 65 %) with a probable nsLTP AKCS9 (XP_003528957) from soybean (Schmutz et al. 2010). On the other hand, LTP3 proteins had the highest homology (e-value: 2e-18) with the Type 2 nsLTP precursor (CAH69201) from *Triticum aestivum* (Boutrot et al. 2007). Finally, the P-U613906 protein showed the highest scores (e value: 7e-18) with the nsLTP (Q43681) from *Vigna unguiculata* (Krause et al. 1994). All the coffee nsLTPs contained a conserved nsLTP2 domain (cd01959) in the conserved protein domain CDD database (Marchler-Bauer et al. 2013) formed by 8 cysteine residues (C_1 to C_8), the C_5XC_6 motif (characterized by the presence of hydrophobic residues such as leucine or phenylalanine at position 65 of the full-length proteins) and followed the protein pattern CX4LX2CX9-11P[S,T] X2CCX5QX2-4C[L,F]CX2[A,L,I]X[D,N]PX10-12[K,R] X4-5CX3-4PX0-2C.

Phylogenetic analysis reveals that all *coffea* proteins identified in this work belong to Type II (Fig. 4A). There is an evidence of evolutionary similarity between these coffee nsLTPs sequences and *Arabidopsis* nsLTPs: AtLtpII.1 (At1g43665), AtLtpII.12 (At5g38160.1) and AtLtpII.13 (At5g38170) from fruit (Fig. 4B). At the sequence level, coffee proteins exhibit CXC pattern, where X is either a leucine or a phenylalanine residue, both hydrophobic. In addition, 3D modeling indicates this residue is facing the hydrophobic cavity in support to our classification analysis.

The most common type of genetic variations in organisms is single nucleotide polymorphism (SNPs). Several coding-SNPs were identified in this work as analysis of amino acid substitution revealed 24 different amino acids on 11 positions (Fig. 5A). Non-synonymous SNPs (promotes amino acid change) are 'neutral' if the resulting point-mutated protein is not functionally discernible from the wild type. Although non-synonymous SNPs generally have the most obvious functional/biochemical effects, they do not necessarily associate with functional or structural consequences (Bromberg and Rost 2007).

Thus, to predict effect of non-synonymous polymorphisms in nsLTPs protein functions within coffee species (Cc or Ca), we analyze the amino acids proprieties as well as the localization in protein structure. The replaced amino acids conserved the physic-chemical proprieties in most cases (Fig. 5C). For instance, the amino acids located at positions 1, 17, 18, 19, 23, 24 and 47 are hydrophilic and replaced by hydrophilic amino acid. Likewise, the amino acids at positions 5, 16, 29 and 36 are hydrophobic and replaced by another with the same property. To display the position of these amino acids substitution at protein structure, 3D modeling was done (Fig. 5B).

Expression of Type II *nsLTP* genes

Using the *nsLTP* cDNA fragment as a probe, a Northern-blot experiment detected transcripts with an expected length of approximately 500 bases in fruits (at 120 DAF) but not in roots and leaves of *C. arabica* cv. Catuaí Amarelo (Fig. 6). The expression of nsLTP-encoding genes was also analyzed in developing fruits of *C. arabica* cv. IAPAR59 collected regularly between 30 and 210 DAF (Fig. 6). This confirmed the high expression in fruits at 120 DAF, while gene expression was undetectable in the earlier and later stages of fruit development.

Primer pairs LTP-FT/LTP-R1 and LTP-FT/LTP-R2 specific to the *LTP3* and *LTP1–LTP2*, gene sequences, respectively, were used in quantitative PCR experiments to analyze the expression of nsLTP-encoding genes individually in pericarp, perisperm and endosperm tissues separated from fruits of *C. arabica* cv. IAPAR59 as well as in separated endosperm from fruits of *C. canephora* harvested at regular stages of maturation. In *C. arabica*, apart from the low expression of LTP1–LTP2 genes in the pericarp at 30 DAF, these genes were not expressed in the pericarp (Fig. 7A). On the other hand, *LTP3* gene expression was very low at 30 and 60 DAF, increased afterwards to reach a peak at 120 DAF and decreased to be barely detectable in the latest stages of pericarp development (150–210 DAF). No *nsLTP* gene expression was observed in perisperm at 60 and 90 DAF (Fig. 7B). In the endosperm, expression of both the *LTP3* and *LTP1–LTP2* genes was high at 90 and 120 DAF and undetectable in the latest stages of maturation. When detected, *LTP3* expression was always higher (three to fourfold) than expression of *LTP1–LTP2* genes (Fig. 7B). In *C. canephora*, LTP3 gene expression was

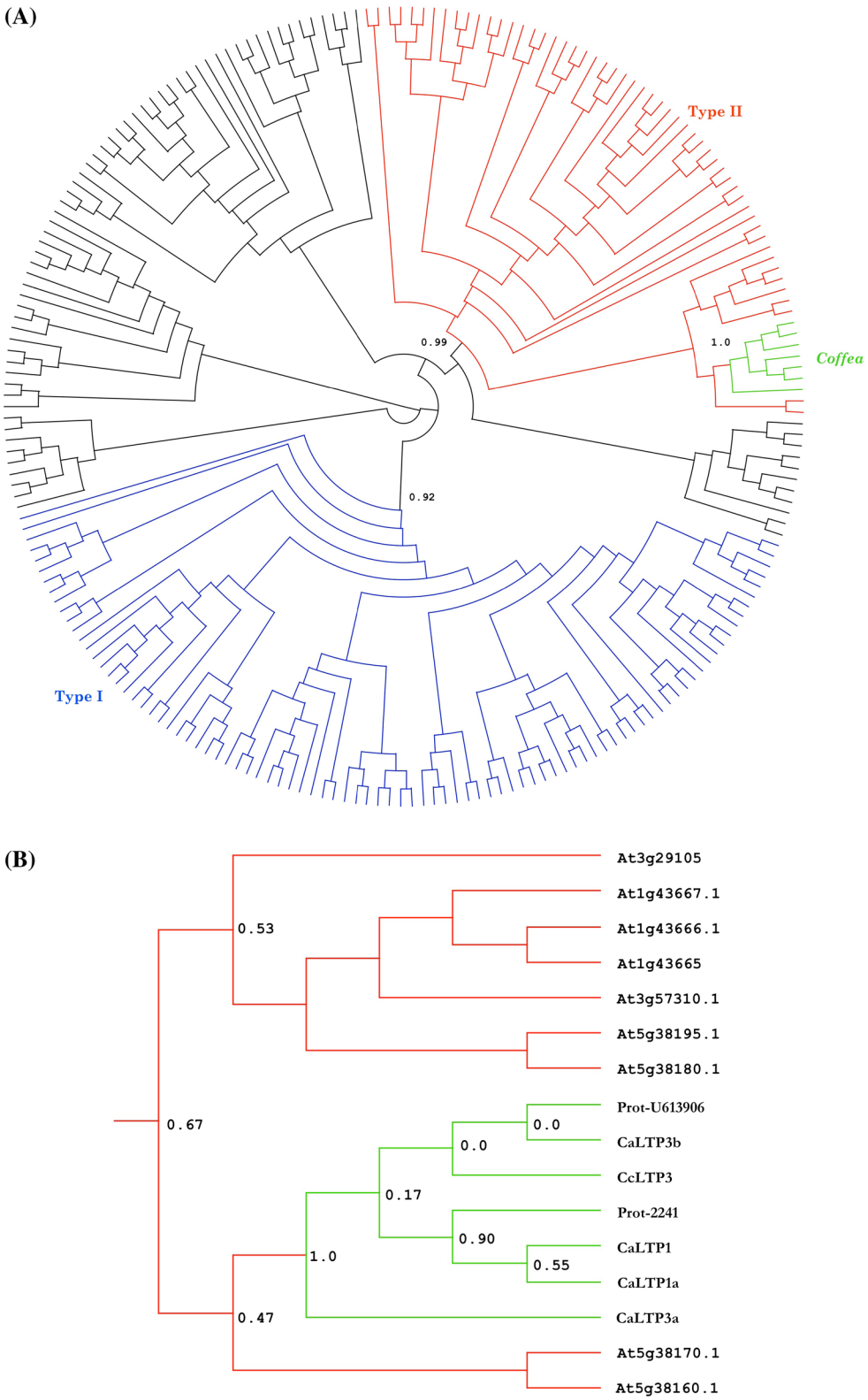
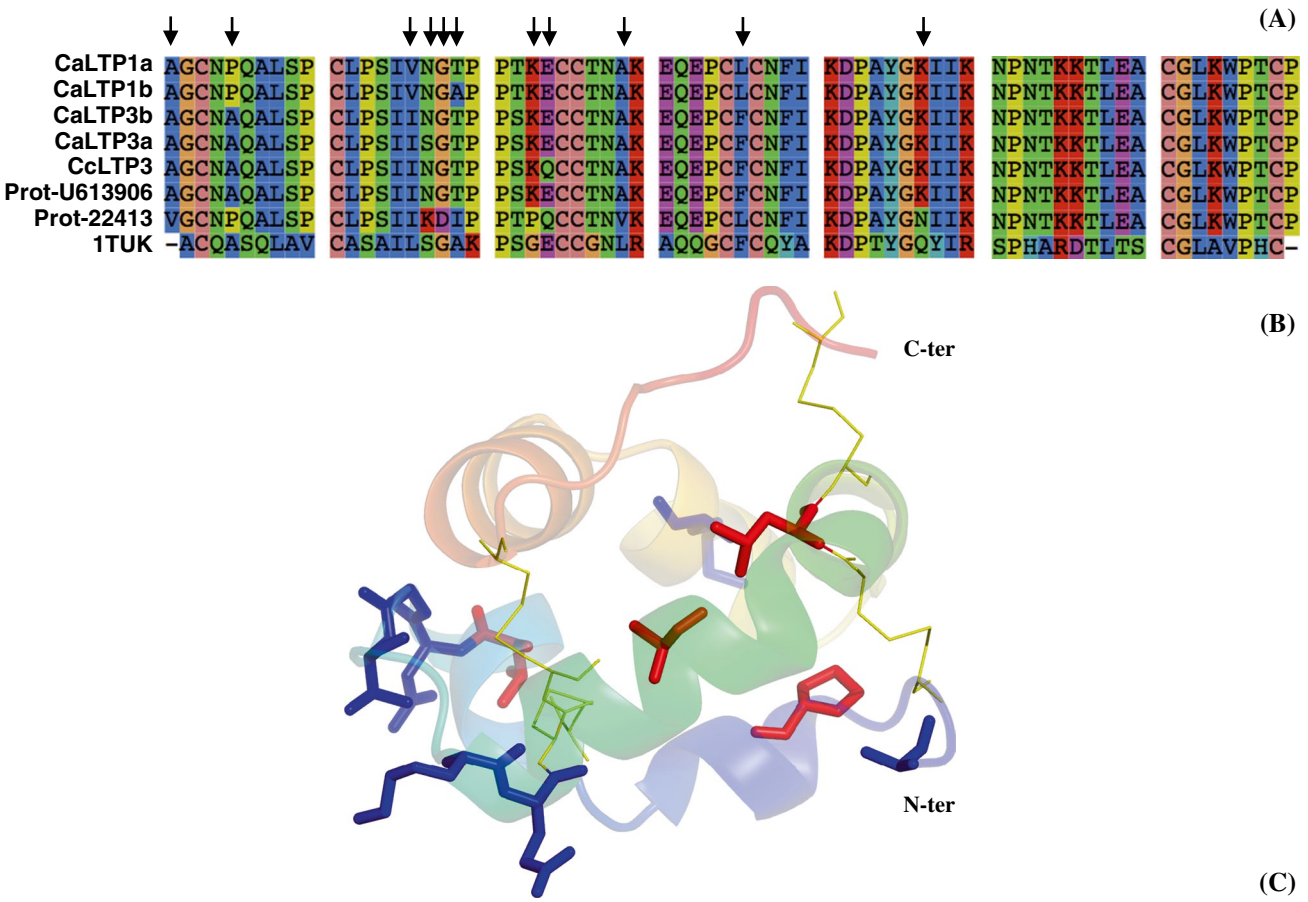


Fig. 4 Phylogenic analyses. **A** *Coffea* nsLTPs sequences (in *green*) are members of Type II plant nsLTP (in *red*). Sequences in *black* belong to other nsLTP Types (see Boutrot et al. 2008). **B** Close up showing the close relationship of coffee nsLTPs (*green lines*) with those of *Arabidopsis thaliana* (At) (*red lines*)

highly detected in earliest stages of endosperm development (120–150 DAF) and no more after (Fig. 7C). In this species, expression of no *LTP1–LTP2* genes was undetectable in the endosperm (data not shown).

Isolation and characterization of the *nsLTP* promoter region

A primer-assisted genome walking experiment led to the isolation of three fragments of 1.9, 1.3, and 0.85 kb



(C)

Sequence position	1	5	16	17	18	19	23	24	29	36	47
Residue type	A	P	V	N	G	T	K	E	A	L	K
	V	A	I	S	D	A	P	Q	V	F	N
			L	K		I					
Location	S	C	S	S	S	S	S	S	C	C	S

Fig. 5 nsLTP Alignments and 3D prediction. **A** Alignment of the 7 coffee nsLTP together with 1TUK (the wheat nsLTP used for modelisation). The SNPs (amongst the *Coffea* sequences) are indicated with *arrows*. **B** 3D model for the first CaLTP1a nsLTP. The backbone is displayed as cartoon with rainbow colors from the N-ter to C-ter. The side chains of the residues affected by the SNPs are displayed as stick (*red* for hydrophobic and *blue* for hydrophylic) and disulfide bonds are displayed as *yellow sticks*. **C** Table with the sequence position of the SNPs, residue type induced by SNPs and their location (exposed to solvent (S) or facing the cavity (C)—*red* for hydrophobic and *blue* for hydrophylic). The amino acid coloring scheme is as follow: *red*, *blue*, *green*, *cyan*, *pink*, *fuchsia*, *yellow* and *orange* for [KR], [AFILMVW], [NQST], [HY], [C], [DE], [P] and [G]

(respectively GW4, GW1, and GW2 in Fig. 8) that were sequenced and aligned to obtain an *nsLTP* promoter consensus sequence that was used to design four F-pBI primers and the R1-pBI primer (Table 1), containing the *HindIII* and the *BamHI* restriction sites, respectively. With this new amplification round, four *nsLTP* promoter fragments (1.2, 1.0, 0.82 and 0.345 kb) were obtained from the genomic DNA of *C. arabica* cv. Catuaí Amarelo, then sequenced and aligned, giving the *nsLTP* promoter sequence (Fig. 9) that was identical to the *nsLTP* promoter consensus sequence previously amplified by the genome walking strategy. This promoter contained a putative TATA box (TATAAAT) located 96 bp upstream of the start codon. However, no obvious CCAAT sequence could be identified. Despite the fact that the transcriptional start site of this gene was not determined, it was assumed to be localized 27 bp downstream of the TATA box by the TSSP program for the prediction of plant promoters. Several *cis*-regulatory elements known to be responsible for the spatial and temporal specificity of gene expression in other plants were identified, such as the TGCAC motif (1162/1166), a prolamine box CAAAGT (235/240), the CAAGTG box (1071/1076)

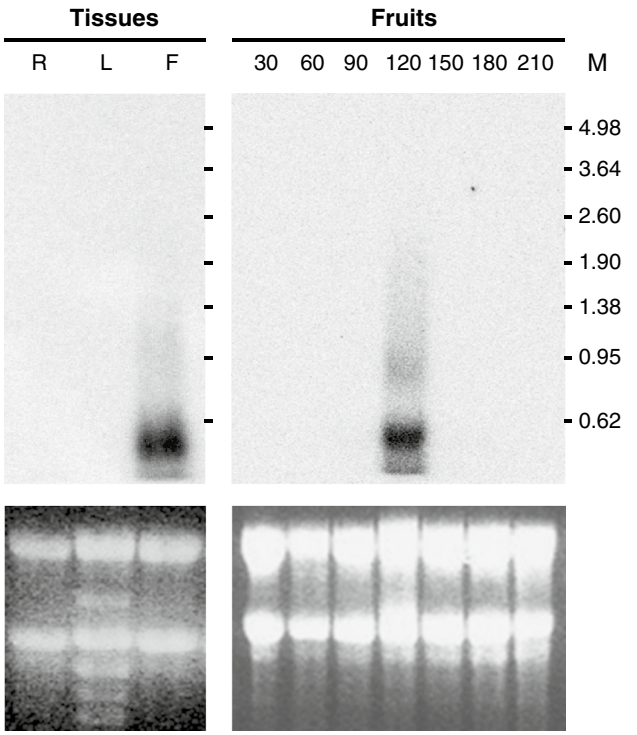


Fig. 6 Expression profiles of Type II nsLTP-encoding genes. Expression was tested in different tissues of *C. arabica* such as roots (R), leaves (L) and from whole fruits (F) at 120 days after flowering (DAF) of *C. arabica* cv. Catuaí Amarelo (2007/2008 harvest) and in whole fruits of *C. arabica* cv. IAPAR59 (2008/2009 harvest) harvested at regular DAF (indicated for each *lane*). Total RNAs (20 μg) were separated by formaldehyde-agarose gel, transferred onto a nylon membrane and hybridized with the *nsLTP* cDNA internal probe (*upper part*). Sizes of RNA molecular weight markers (Promega) are noted on the right (M). rRNAs stained by ethidium bromide were used to monitor the equal loading of RNA samples (*lower part*)

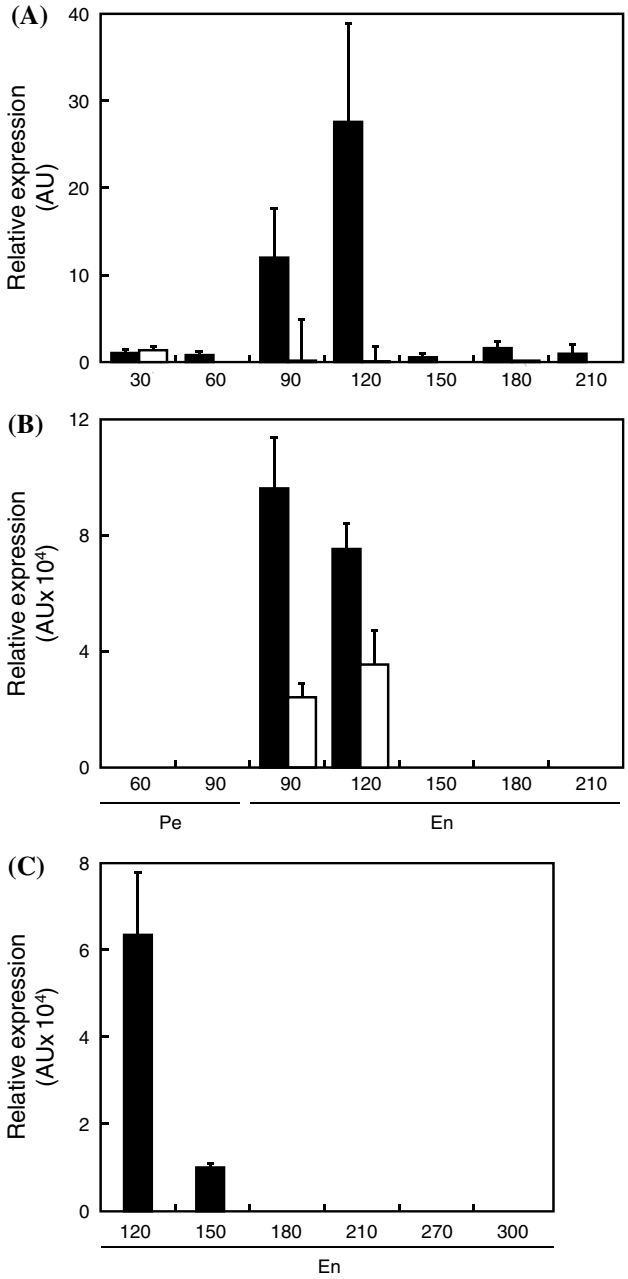


Fig. 7 Expression of nsLTP-encoding genes during coffee fruit development. The expression of *LTP1–LTP2* (*CaCe*, *white isobars*) and *LTP3* (*CaCc*, *black isobars*) genes was analyzed by q-PCR using the LTP-FT/LTP-R2 and LTP-FT/LTP-R1 primer pairs, respectively. Tissues corresponded to **A** pericarp, **B** perisperm (Pe) and endosperm (En) separated from fruits of *C. arabica* cv. IAPAR59 (2006/2007 harvest) and **C** to endosperm from fruits of *C. canephora* clone L6P35 (2011/2012 harvest) collected at regular days after flowering (DAF). Expression levels are expressed in arbitrary units (AU) of nsLTP-encoding genes using the expression of the *UBI* gene as endogenous control. Values are the mean of three biological replications ±SD

and five TGATTCA motifs (564/570; 665/671, 837/843, 943/949 and 1113/1119). This sequence also contained two boxes (480/486 and 1058/1064) matching with the RTTTTTR element, six ACGT boxes and four MYB-binding boxes (CNGTTR). The *nsLTP* promoter also contained boxes known as essential elements for many light-regulated genes such as several GT-1 binding sites (GRWAAW) as well as an *rbcS* general consensus sequence (AATCCAA), all mainly located between nucleotides 210 and 420. From a structural viewpoint, the *LTP* promoter also presented a well conserved and repeated DNA like the DNA-1 motifs (51 bp: 534/584, 635/685, 713/763, 807/857 and 913/963), the DNA-2 motifs (16/17 bp: 507/522, 585/601, 686/702 and 886/901), the DNA-3 motifs (10 bp: 523/532, 602/611, 703/715, 796/805 and 902/911) and the DNA-4 motifs (29 bp: 764/790, 858/886 and 964/989). We also noted particular arrangements of these motifs, the DNA-2 and -3 motifs being associated four times and flanking the first DNA-1 sequence, for example. Separately, the DNA-3 motif was also present alone before the fourth DNA-1 motif. On the other hand, the three DNA-4 motifs followed the last three DNA-1 motifs.

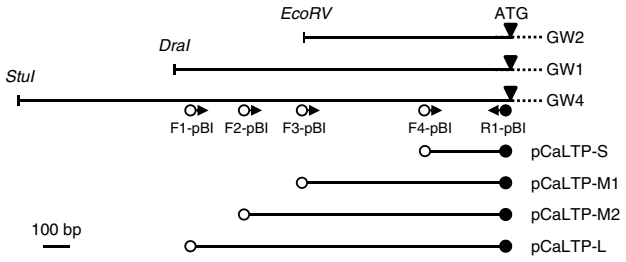


Fig. 8 Diagrammatic representation of the *nsLTP* promoter. The fragments amplified from the genomic DNA of *C. arabica* cv. Catuaí Amarelo by the genome walking experiment are indicated (GW) as well as the restriction enzymes used. The primers used to construct the pCaLTP-S (345 bp), pCaLTP-M1 (827 bp), pCaLTP-M2 (1047 bp) and pCaLTP-L (1252 bp) vectors are shown (*arrows*) with the *HindIII* (*white circles*) and *BamHI* (*black circles*) restriction sites

Analysis of *LTP* promoter in transgenic tobacco plants

A deletion analysis was carried out to define precisely the *LTP* promoter regions essential for its expression. Four constructions were made by fusing 345 bp (pCaLTP-S), 827 bp (pCaLTP-M1), 1,047 bp (pCaLTP-M2) and 1,252 bp (pCaLTP-L) respectively of this promoter to the *uidA* reporter gene (Fig. 8), and further introduced separately into *N. tabacum* by *A. tumefaciens*-mediated transformation. Several T0 transformants were regenerated and used to perform histochemical assays by checking β-glucuronidase (GUS) activities in roots, leaves, fruits, seeds, petals, stamens and anthers (Fig. 10). The plants transformed by pCaLTP-S showed histochemical

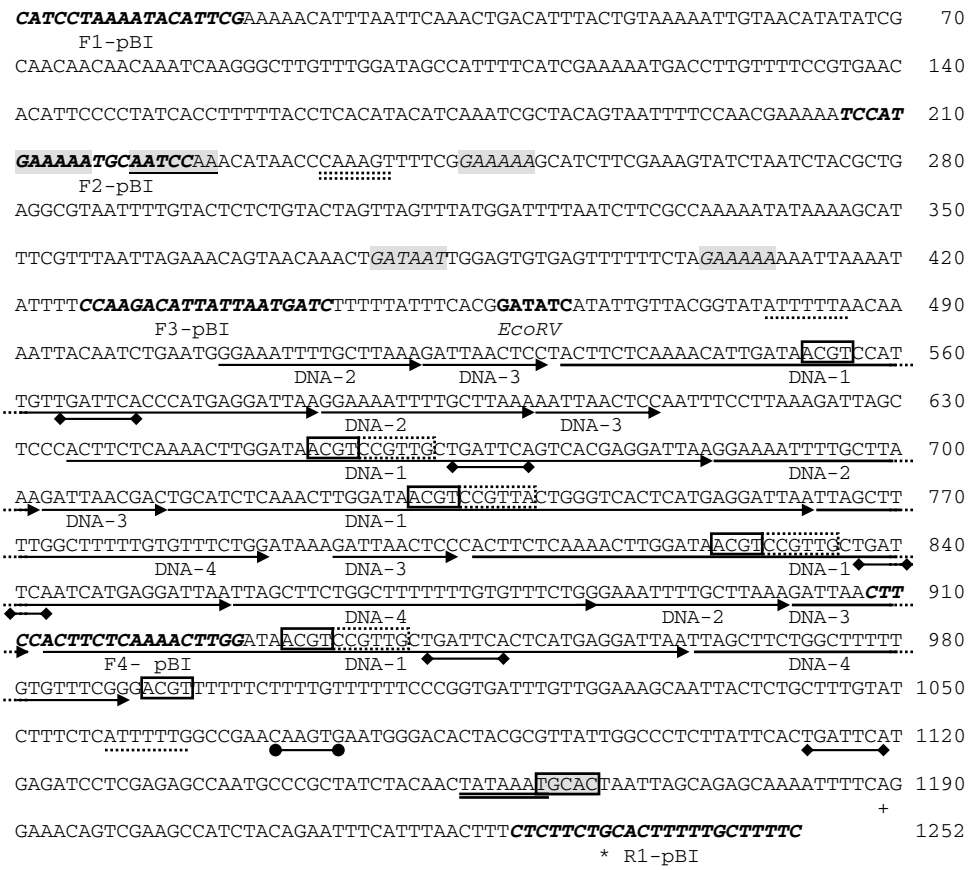


Fig. 9 Complete nucleotide sequence of the *nsLTP* promoter region. Nucleotide numbers are indicated to the right of each *lane*. The putative transcriptional start site (*plus*) is indicated as well as the 5′-end of the *nsLTP*-cDNA and gene sequences (*asterisk*). The TATA box (*double underlined*), the prolamine box (CAAAGT, double-dotted line), TGATTCA motifs (*filled diamond*), RTTTTTR motifs (*dotted line*), RY-like TGCAC motif (*gray box*), ACGT motif (*white box*), CNGTTR MYB-binding motifs (*dotted box*), CAAGTG E-box (*filled circle*), GRWAAW GT-1 binding sites (in *gray and italics*) and the AATCCAA *rbcS* consensus sequence (in *gray and underlined*) are also presented. The primers F1-pBI, F2-pBI, F3-pBI, F4-pBI and R1-pBI used to construct the pCaLTP vectors are in *bold* and *italics*. The *horizontal arrows* indicate the position of repeated DNA motifs (DNA-1, -2, -3 and -4) found in the *nsLTP* promoter. The *EcoRV* restriction site (GATATC) of the GW2 fragment is in *bold*. This sequence was deposited in the EMBL/GenBank database under number HG323817

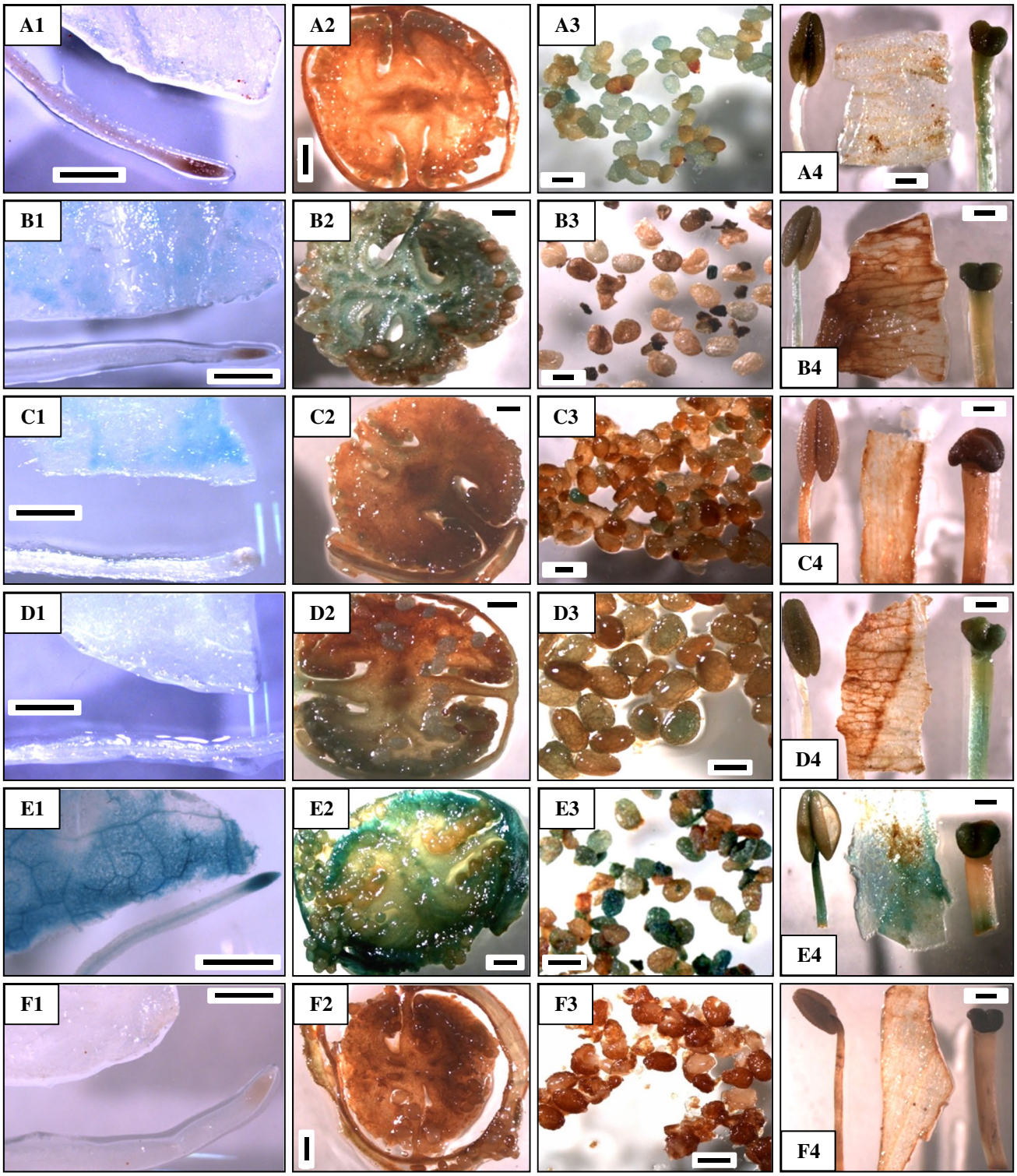


Fig. 10 Histochemical localization of GUS activity in transgenic tobacco plants transformed with the pCaLTP vectors. GUS activities were tested in **1** leaves (*top*) and roots (*bottom*), **2** unripe capsules and immature seeds, **3** isolated mature seeds and **4** stamens (filament and anther, *left*), petal (*middle*) and pistil (style and stigma, *right*) of plants transformed with **A** pCaLTP-S, **B** pCaLTP-M1, **C** pCaLTP-M2, **D** pCaLTP-L, **E** pCaMV35S (35S::*uidA* cassette, positive control) and **F** untransformed tobacco plants (negative control). For images **1**, **2** and **4**, the *black bars* represent 1 mm. For images **3**, *black bars* represent 0.5 mm

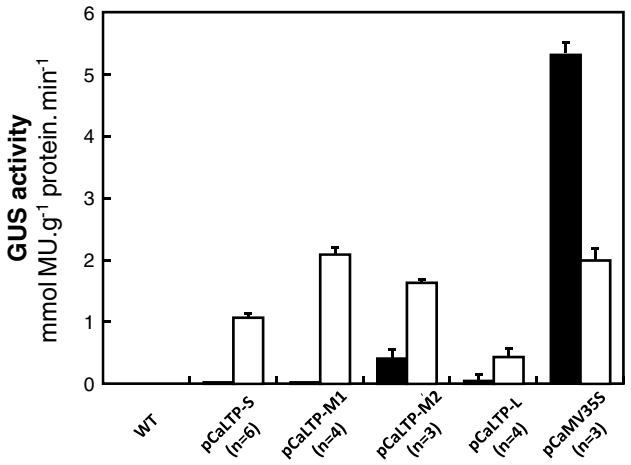


Fig. 11 GUS activities in transgenic tobacco plants. Activities were measured in mature seeds (*white isobars*) and leaves (*black isobars*) from tobacco plants transformed with pCaLTP-S (345 bp), pCaLTP-M1 (827 bp), pCaLTP-M2 (1,047 bp), pCaLTP-L (1,252 bp), pCaMV35S (positive control) and untransformed tobacco plants (WT, negative control). GUS activities (in mmol MU g^{-1} protein min^{-1}) are the mean of three different experiments. The *bar* indicates the standard error of the mean. The number of independent transformed lines tested (*in brackets*) is indicated

staining with X-gluc in isolated mature seeds, but not in unripe capsules, immature seeds, leaves, roots, petals and other flower tissues (Fig. 10A). In tobacco plants transformed by pCaLTP-M1 (Fig. 10B), GUS activity was detected in placental (inner) tissue of fruits but also weakly in immature seeds and slight staining also occurred in leaves but not in roots and flower organs. The tobacco transformed by pCaLTP-M2 showed GUS staining in leaves and also in isolated seeds but not in root and in unripe capsules and flower organs (Fig. 10C). For the tobacco transformed by pCaLTP-L (Fig. 10D), GUS activity was observed in immature seeds and placental tissue of the capsules (low expression), in mature seeds, as well as in styles, and weakly in leaves but not in roots, petals and stamens. As a positive control (Fig. 10E), GUS activity was well detected in all the tissues of T0 plants transformed by pCaMV35S vector carrying the CaMV35S::*uidA* cassette. One the other hand, all the tissues of untransformed tobacco plants remained unstained (negative control).

For each construction, quantitative fluorometric GUS assays were performed using total proteins extracted from leaves and mature seeds (Fig. 11). As expected, GUS activity was not detected in either seeds or leaves of untransformed (WT) tobacco plants. On the other hand, GUS activity was well detected in both seeds and leaves of transgenic tobacco plants transformed by the pCaMV35S vector, confirming the constitutive expression of the CaMV 35S promoter in higher plants.

No GUS activity was observed in leaves of plants transformed with pCaLTP-S and pCaLTP-M1 vectors and very low activity was observed in leaves of plants transformed with pCaLTP-L vector. The pCaLTP-M2 vector was the only one leading to significant GUS activity in leaves of transformed tobacco that corresponded to approximately 7.7 % of that observed in the leaves of pCaMV35S plants. For seeds, GUS activities were detected in all of pCaLTP constructions tested. GUS levels were high in seeds of the plants transformed with pCaLTP-M1 and similar to those found in seeds of the plants transformed by pCaMV35S. In seeds of the plants transformed with pCaLTP-M2, pCaLTP-S and pCaLTP-L vectors, GUS activities were 78, 50 and 21 % respectively than those of pCaLTP-M1 seeds.

Discussion

The main purpose of this work was to characterize the nsLTP-encoding genes specifically expressed in coffee fruits. The search for coffee *nsLTP*-EST in public databases revealed several contigs used to define primer pairs that enabled the identification of nsLTP-encoding cDNA and gene sequences from the *C. arabica* and *C. canephora* species. For both *CaLTP1a* and *CaLTP2* sequences of *C. arabica*, cDNA cloned from fruits of IAPAR59 and genes from the genomic DNA of Mundo Novo were strictly identical. This suggested the existence of two closely related nsLTP-encoding genes in this species. However, as the *CaLTP1a* and *CaLTP2* nucleic sequences diverged by only 2 bases located in their 5′ region used for primer designs, it is possible that these were introduced during the amplification cycles by primer mismatches or mistakes. If this occurred, both sequences should be considered as equal and coming from the same *nsLTP* gene (e.g. *CaLTP1a*). The *CaLTP1b* gene was also amplified from *C. arabica* and diverged from *CaLTP1a* by only one base in the nsLTP-coding sequence suggesting that it was an allele of this gene. Two additional sequences were also isolated in *C. arabica*: *CaLTP3a* corresponding to a cDNA isolated from fruits of IAPAR59 and *CaLTP3b*, corresponding to the *nsLTP* gene of Mundo Novo. Both sequences were also highly identical and diverged by only one base in their nsLTP-encoding sequence. In *C. canephora*, the *CcLTP3* gene was the only sequence obtained that appeared to be highly identical to *CaLTP3a* and *CaLTP3b* of *C. arabica*. Nucleic sequence alignments revealed that *LTP3* diverged from *LTP1* to *LTP2* by few bases in the nsLTP-encoding sequence but also by the insertion/deletion of a 13 bp sequence in their common 3′ UTR region. The fact that no *LTP1–LTP2* sequences were amplified from *C. arabica* and that *LTP3* sequences were amplified from both the *C. arabica* and *C. canephora* species, suggested that the

CaLTP1–CaLTP2 genes corresponded to *nsLTP* sequences carried by the *C. eugenioides* sub-genome of *C. arabica* (hereafter called *CaCe*) and that *CaLTP3* gene was carried by the *C. canephora* sub-genome of *C. arabica* (hereafter called *CaCc*). This is also supported by the fact that (1) SGNCaU607388 and SGNCcU613906 contigs, respectively formed by the assembly of coffee EST from both *C. arabica* and *C. canephora*, were identical to *LTP3* and that (2) expression of *CaCc*, but not of *CaCe,* was detected in *C. canephora* endosperm.

Recently, the nsLTPs from rice, wheat, and *A. thaliana* were classified in nine different types on the basis of sequence similarity (Boutrot et al. 2008). These proteins can differentiate into two major groups. Type I (9 kDa nsLTPs), which represents 50 % of nsLTP, shows a characteristic fold with four helices and the residue X of the sequence CXC located on the third helix is a hydrophilic residue exposed to the solvent, towards the outer part of the protein. The others nsLTPs are classified in 8 other groups which share similarities in term of size (7 kDa), folding (five helices) and the X in CXC pattern is an hydrophobic residue facing towards hydrophobic cavity.

In this study, phylogenetic analyses revealed that all nsLTPs deduced from cloned cDNA and genes corresponded to Type II LTP with the same length and a conserved peptide signal of 29 residues. They also displayed similar MW of roughly 10.4 kDa in their pre-protein form and 7.4 kDa in their processed form characterizing these proteins. Whether processed or not, they also had the same basic pI. CaLTP1a and CaLTP2 diverged from CaLTP1b by only one amino acid (in position 48). On the other hand, CaLTP3a, CaLTP3b and CcLTP3 appeared very similar, diverging from each other by only one amino acid in their processed form. Together, LTP1–LTP2 and LTP3 proteins showed 94 % identity and 97 % homology and also contained the conserved nsLTP2 domain.

Several studies noted the localization of residue within 3D structure is relevant for the effect of a particular substitution on function (Chasman and Adams 2001; Saunders and Baker 2002; Kharabian 2010). In fact, changes of a hydrophobic into a non-hydrophobic amino acid may be non-neutral in the protein core while it may not matter on the surface (Bromberg and Rost 2007). Studies in rice (*Oryza sativa*) (Larkin and Park 2003) have already reported codon-SNPs at exons 9 and 10 of GBSSI (Granule Bound Starch Synthase) gene, with non-functional and functional effects, respectively. They also have verified that one SNP in each of the, exon/intron1 boundary site, exon 6 and exon 10, are inherited as haplotypes and expressed as combination together to regulate the GBSSI function. Chen et al. (2008a, b) have also showed that these SNPs can alter the amylose content and pasting properties of rice. For predicting the effect of SNP [C/A] at exon 6, the simulation of native protein structure (Y) and mutant (S) was done (Kharabian 2010). The results showed a distinctive deformed loop at the mutation position, located at the outer layer (surface) of the GBSSI molecule which possibilities to affect the efficiency of the protein binding site. For the coffee Type II nsLTPs, most of the residues affected by the SNPs were located in surface and mainly in one side of molecule. These results corroborate with literature data which evidence internal residues are more constrained by evolution and surfaces tend to be less conserved (Bromberg and Rost 2007). Then, main variations identified in the present work are evolutionarily common and the few modifications within the molecule do not seem to affect the functional site. Based on these results, we may assume that these SNPs do not have a significant impact on function of the nsLTPs within *Coffea* species.

The expression of Type II nsLTP-encoding genes was tested in different tissues of *C. arabica*. This was done by Northern blot experiments with a probe able to recognize all *nsLTP* transcripts. It demonstrated the absence of *LTP* gene expression in roots and leaves of *C. arabica* but high expression of that gene in fruits at 120 DAF. The detected expression in fruits was refined in isolated pericarp, perisperm and endosperm by RT-qPCR experiments using primer pairs localized in the 3′ UTR region of the *nsLTP* sequences and specific to *CaCe* (*LTP1–LTP2*) and *CaCc* (*LTP3*) homeologous genes. In *C. arabica*, *CaCc* expression was clearly observed in the pericarp at 90 and 120 DAF while *CaCe* expression was negligible in that tissue. These homeologous genes were not expressed in the perisperm but were concomitantly expressed early (90 and 120 DAF) during the endosperm development. Afterwards, expression of *CaCe* and *CaCc* was negligible up to the end of bean maturation. The comparison of expression levels in the pericarp and endosperm tissues revealed higher expression (10^3 fold) of *nsLTP* genes in the endosperm than in the pericarp. These results also highlighted the predominant expression of the homeologous *CaCc* genes over the *CaCe* genes in the perisperm and endosperm tissues. Like in *C. arabica*, *CaCc* expression was also highly during the earliest stages of endosperm development in *C. canephora*. However, *CaCe* expression was not detected in this species, therefore confirming that the *CaLTP1* and *CaLTP2* genes from *C. arabica* were cloned from its *C. eugenioides* sub-genome. Few publications have investigated the expression of homeologous genes in *C. arabica* (Petitot et al. 2008; Marraccini et al. 2011). Vidal et al. (2010) reported that, in this species, the *C. eugenioides* sub-genome may express genes coding for proteins that assume basal biological processes while the *C. canephora* sub-genome contributes to adjusting Arabica gene expression by expressing genes coding for regulatory proteins. To the authors' knowledge, the results presented here with an undetectable expression

of *CaCe* in the pericarp and co-expression of *CaCc* and *CaCe* homeologs in the grain tissues (perisperm and endosperm), are the first describing differential expression of homeologous genes within different tissues of the same organ (e.g. fruit).

Expression studies also highlighted that maximum expression of Type II nsLTP-encoding genes was observed at 120 DAF and 90–120 DAF by Northern blot and qPCR experiments, respectively, carried out with the cultivar (IAPAR59) of *C. arabica*. This discrepancy could be explained by the fact that in whole fruits at 90 DAF the perisperm forms the main tissue while the endosperm is a small developing tissue (Geromel et al. 2006), as the Northern blot was done with whole fruit the nsLTP mRNA was diluted in 90 DAF, therefore not detectable by this technique. Otherwise, the qPCR was done using separate tissues, not having this dilution effect. Another explanation is that the beans collected from plants grown under field conditions were subjected to different meteorological conditions that could affected fruit development. For example, plants used for harvesting fruits in 2008/2009 (Northern blot experiment) suffered from drought after blooming, which delayed their fruit development by around a month compared to those harvested in 2006/2007 and analyzed by qPCR experiments (data not shown). Whatever the situation, the expression peaks of Type II nsLTP-encoding genes coincided with the decline of the perisperm and the expansion of the endosperm (De Castro and Marraccini 2006; Geromel et al. 2006). In seeds of *C. arabica* cv. Laurina, Joët et al. (2009) showed that lipids began to be synthesized in the perisperm and then loaded into the developing endosperm where their synthesis and mobilization continued. The same process was also suggested to occur for kahweol and cafestol diterpens (Dias et al. 2010). These data also tallied with the peaks for the transcriptional activity of genes encoding proteins involved in fatty acid synthesis (e.g. acetyl-CoA carboxylase, diacylglycerol acyltransferase, enoyl-ACP reductase, hydroxyacyl-ACP dehydrase, ketoacyl-ACP reductase, ketoacyl-ACP synthase) or participating in oil body formation (e.g. oleosin, caleosin and steroleosin) (Salmona et al. 2008; Joët et al. 2009). A tobacco nsLTP1 (TobLTP2) has been shown to be involved in cell wall loosening suggesting that the association of LTP with hydrophobic wall compounds promotes non-hydrolytic modifications in the cell wall which facilitate cell extension (Nieuwland et al. 2005). This is also in accordance with the fact that nsLTPS were found in cell wall compartment (Thoma et al. 1993). In that sense, it is possible that Type II nsLTPs play an important function during the tissue rearrangements observed during coffee bean development and characterized by the rapid expansion of "liquid" endosperm (De Castro and Marraccini 2006). Even if Type II nsLTPs reported here are not related to Type I LTPs exhibiting α-amylase inhibitor properties (Zottich et al. 2011), it cannot be completely ruled out that high accumulation of Type II nsLTPs supposed to occur concomitantly to high expression of *nsLTP* genes, could represent a defense mechanism against fungal and bacterial pathogens during coffee bean development. This also does not preclude the participation of coffee Type II nsLTP in other biological processes like in response to drought, as suggested by high expression of Type II nsLTP-encoding genes in leaf primordial and plagiotropic meristems of drought-tolerant cultivar of *C. arabica* grown without irrigation (Vidal et al. 2013).

It is of particular interest to develop a repertoire of seed-specific promoters for future studies on transcriptional control in coffee, particularly to direct the expression of recombinant genes in the grain. Several coffee endosperm-specific promoters have already been described in the literature (Lashermes et al. 2008). As the expression of *nsLTP* genes was strong and seed-specific, the promoter was isolated and studied. Its analysis revealed the presence of several DNA boxes known to be important mainly in the regulation of genes expressed in seeds. This was the case of TGCAC motifs known to constitute the core region of the legumin DNA box that controls the expression of many storage protein-encoding genes and shares significant homology with the RY repeat (CATGCATG) involved in the regulation of genes coding for legumin storage proteins (Shirsat et al. 1989), a prolamine box known to be involved in quantitative regulation of the rice glutelin gene *GluB-1* (Wu et al. 2000) and the CAAGTG boxes closely related to the E-box CANNTG involved in the seed-specific expression of phaseolin (Kawagoe and Murai 1992). This promoter also contained five TGATTCA motifs closely related to the TGAGTCATCA (TGAC-like) motif essential for seed-specific expression of pea lectin (de Pater et al. 1993), two RTTTTTR elements corresponding to the binding site of the SEF 4 transcription factor reported to activate expression of the β-conglycinin 7S storage protein in soybean (Lessard et al. 1991), six ACGT boxes required for seed-specific expression of a 2S storage protein (Vincentz et al. 1997) and the *erd1* (early response to dehydration) gene in *Arabidopsis* responsible for etiolation-induced increase (Simpson et al. 2003) and four MYB-binding boxes (CNGTTR) involved in water stress responsive regulation of gene expression (Lüscher and Eisenman 1990). It is worth noting that the CNGTTR boxes were always linked to the ACGT boxes. This promoter was also characterized by the presence of several DNA repeats which are known to play an important role in regulating gene expression. For instance, a tandem-repeat of the *rsus3* endosperm specific promoter from rice (*Oryza sativa*) fused to the *uidA* reporter gene displayed an activity three times greater than the single copy construct

(Rasmussen and Donaldson 2006). In *Arabidopsis*, gene promoters enriched in GGCCCAWW and AAACCCTA repeat sequences appeared up-regulated while those enriched with repeated TTATCC motifs were down-regulated (Tatematsu et al. 2005). To our knowledge, the structure of the *nsLTP* promoter reported here, with very well conserved and long DNA repeats organized in tandem and concentrated in a region of less than 500 bp, is quite original. It is also remarkable that most of the ACGT, CNGTTR and TGATTCA boxes were included in these repeats and always arranged in the same order, suggesting they play an important role in regulating the expression of the *LTP* promoter.

The deletion analysis of the coffee *nsLTP* promoter performed in transgenic tobacco plants revealed that it functioned as a seed-specific promoter. This was particularly the case of the shorter (pCaLTP-S) and medium (pCaLTP-M1) fragments of the *nsLTP* promoter, suggesting that the long DNA repeats organized in tandem might play an essential role in the strength and the tissue-specificity of this promoter. In a previous study, Marraccini et al. (1999) also reported that a 245 bp fragment of the 11S coffee promoter was also able to confer seed-specificity of this sequence in transgenic tobacco. It is possible that the putative TGATTCA, ACGT and CNGTTR binding sites included in the DNA-1 tandem-repeat carried out this function. By comparison with the pCaLTP-M1 plants, the detection of GUS activity in leaves of pCaLTP-M2 plants could be explained by the presence of GT-1 binding sites (Terzaghi and Cashmore 1995) and an *rbcS* general consensus sequence (Donald and Cashmore 1990) in the sequence flanked by primers F2-pBI and F3-pBI. Even though no particular motifs were found within the sequence flanked by primers F1-pBI and F2-pBI, this region seemed to function as a silencer since the GUS activities of pCaLTP-L plants were greatly reduced in both seeds and leaves compared to those of pCaLTP-M2 plants. The results presented here, together with those demonstrating that the coffee *RBCS* promoter was highly expressed and light-regulated in transgenic tobacco (Marraccini et al. 2003), support the idea that the mechanisms implicated in the transcriptional control of *nsLTP* gene expression were highly conserved between these two species.

To better understand the function of the *nsLTP* promoter, it will be of interest to dissect this sequence more deeply, for example by performing site-directed mutagenesis of the DNA motifs and repeat sequences reported here. The effects of such changes on the capability of the *nsLTP* promoter to drive the expression of genes of interest in coffee seeds could also be evaluated, since genetic transformations of coffee by *Agrobacterium* mediated systems are now available (Lashermes et al. 2008).

Accession numbers

The *CaLTP1a*, *CaLTP2* and *CaLTP3a* cDNA sequences were deposited in the GenBank database under the accession numbers HG008739, HG008740 and HG008741, respectively. The *CaLTP1a*, *CaLTP1b, CaLTP2, CaLTP3b, CcLTP3* gene and promoter sequences were deposited in the GenBank database under the accession numbers HG323818, HG323819, HG323820, HG323821, HG323822 and HG323817, respectively.

Acknowledgments We should like to thank Gustavo Costa Rodrigues (Embrapa CNPTIA) for providing the *C. arabica* fruits used in this study and the members of the Apomix laboratory of Embrapa Genetic Resources and Biotechnology for assistance with microscopy. We are especially indebted to Manuel Ruiz and Jean-François Dufayard (CIRAD-UMR AGAP) for fruitful discussions and bioinformatics assistance. We are also grateful to Luciano V. Paiva and Antônio Chalfun Jr. (Lavras Federal University-UFLA) for supporting this project, to Renata H. Santana for their technical assistance and to Peter Biggins (CIRAD) for its critical reading of the manuscript. This work was funded by the Embrapa macroprogram project. The authors acknowledge the scholarships from the Brazilian agencies CNPq (Michelle G. Cotta, Eder A. Barbosa), Brazilian Consortium of Coffee Research (Felipe Vinecky, Natalia G. Vieira) and CAPES (Michelle G. Cotta, CAPES-COFECUB Project Sv738-12). The 3′ end fragment (411 bp) of the *LTP* promoter was patented (BR1020120081628) by INPI-Brazil.

References

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402

An G, Ebert PR, Mitra A, Ha SB (1993) Binary vectors. In: Gelvin SB, Schilperoort RA, Verma DPS (eds) Plant molecular biology manual. Kluwer Academic Publishers, Dordrecht, pp A3/1–A3/19

Arondel V, Tchang F, Baillet B, Vignols F, Grellet F, Delseny M, Kader JC, Puigdomenech P (1991) Multiple mRNA coding for phospholipid-transfer protein from *Zea mays* arise from alternative splicing. Gene 99:133–136

Boutrot F, Guirao A, Alary R, Joudrier P, Gautier M-F (2005) Wheat non-specific lipid transfer protein genes display a complex pattern of expression in developing seeds. Biochim Biophys Acta 1730:114–125

Boutrot F, Meynard D, Guiderdoni E, Joudrier P, Gautier MF (2007) The *Triticum aestivum* non-specific lipid transfer protein (*TaLtp*) gene family: comparative promoter activity of six TaLtp genes in transgenic rice. Planta 225:843–862

Boutrot F, Chantret N, Gautier MF (2008) Genome-wide analysis of the rice and Arabidopsis non-specific lipid transfer protein (nsLtp) gene families and identification of wheat *nsLtp* genes by EST data mining. BMC Genomics 9:86

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72:248–254

Bromberg Y, Rost B (2007) SNAP: predict effect of non-synonymous polymorphisms on function. Nucl Acids Res 35:3823–3835

Cameron KD, Teece MA, Smart LB (2006) Increased accumulation of cuticular wax and expression of lipid transfer protein in response to periodic drying events in leaves of tree tobacco. Plant Physiol 140:176–183

Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol Biol Evol 17:540–552

Chang WC, Lee TY, Huang HD, Huang HY, Pan RL (2008) PlantPAN: plant promoter analysis navigator, for identifying combinatorial *cis*-regulatory elements with distance constraint in plant gene groups. BMC Genomics 9:561

Chasman D, Adams RM (2001) Predicting the functional consequences of non-synonymous single nucleotide polymorphisms: structure based assessment of amino acid variation. J Mol Biol 307:683–706

Chen MH, Bergman C, Pinson S, Fjellstrom R (2008a) Waxy gene haplotypes: associations with apparent amylose content and the effect by the environment in an international rice germplasm collection. J Cereal Sci 47:536–545

Chen MH, Bergman CJ, Pinson S, Fjellstrom R (2008b) Waxy gene haplotypes: associations with pasting properties in an international rice germplasm collection. J Cereal Sci 48:781–788

Cruz F, Kalaoun S, Nobile P, Colombo C, Almeida J, Barros LMG, Romano E, Grossi-de-Sa MF, Vaslin M, Alves-Ferreira M (2009) Evaluation of coffee reference genes for relative expression studies by quantitative real-time RT-PCR. Mol Breed 23:607–616

De Castro RD, Marraccini P (2006) Cytology, biochemistry and molecular changes during coffee fruit development. Braz J Plant Physiol 18:175–199

De Kochko A, Akaffou S, Andrade AC, Campa C, Crouzillat D, Guyot R, Hamon P, Ming R, Mueller LA, Poncet V, Tranchant-Dubreuil C, Hamon S (2010) Advances in *Coffea* genomics. In: Kader JC, Delseny M (eds) Advances in botanical research, vol 53. Academic Press, Oxford, pp 23–53

De Pater S, Pham K, Chua NH, Memelink J, Kijne J (1993) A 22-bp fragment of the pea lectin promoter containing essential TGAC-like motifs confers seed-specific gene expression. Plant Cell 5:877–886

Decazy F, Avelino J, Guyot B, Perriot JJ, Pineda C, Cilas C (2003) Quality of different Honduran coffees in relation to several environments. J Food Sci 68:2356–2361

Dentan E (1985) Etude microscopique du développement et de la maturation du grain de café. In: Proceedings of the international congress of ASIC 11, pp 381–398

Dias RCE, Campanha FG, Vieira LGE, Ferreira LP, Pot D, Marraccini P, Benassi MT (2010) Evaluation of kahweol and cafestol in coffee tissues and roasted coffee by a new high-performance liquid chromatography methodology. J Agric Food Chem 58:88–93

Donald RGK, Cashmore AR (1990) Mutation of either G box or l box sequences profoundly affects expression from the *Arabidopsis* rbcS-1A promoter. EMBO J 9:1717–1726

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19:11–15

Dufayard JF, Duret L, Penel S, Gouy M, Rechenmann F, Perrière G (2005) Tree pattern matching in phylogenetic trees: automatic search for orthologs or paralogs in homologous gene sequence databases. Bioinformatics 21:2596–2603

Emanuelsson O, Brunak S, von Heijne G, Nielsen H (2007) TargetP: locating proteins in the cell using TargetP, SignalP, and related tools. Nat Protoc 2:953–971

Fleming AJ, Mandel T, Hofmann S, Sterk P, de Vries SC, Kuhlemeier C (1992) Expression pattern of a tobacco lipid transfer protein gene within the shoot apex. Plant J 2:855–862

Folstar P (1985) Lipids. In: Clarke RJ, Macrae R (eds) Coffee-chemistry, vol 1. Elsevier Applied Science Publishers, London, pp 203–220

Garcia-Garrido JM, Menossi M, Puigdoménech P, Martinez-Izquierdo JA, Delseny M (1998) Characterization of a gene encoding an abscisic acid-inducible type-2 lipid transfer protein from rice. FEBS Lett 428:193–199

García-Olmedo F, Molina A, Segura A, Moreno M (1995) The defensive role of nonspecific lipid-transfer proteins in plants. Trends Microbiol 3:72–74

Geromel C, Ferreira LP, Guerreiro SM, Cavalari AA, Pot D, Pereira LF, Leroy T, Vieira LG, Mazzafera P, Marraccini P (2006) Biochemical and genomic analysis of sucrose metabolism during coffee (*Coffea arabica*) fruit development. J Exp Bot 57:3243–3258

Giardine B, Riemer C, Hardison RC et al (2005) Galaxy: a platform for interactive large-scale genome analysis. Genome Res 15:1451–1455

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol 59:307–321

Harada E, Kim JA, Meyer AJ, Hell R, Clemens S, Choi YE (2010) Expression profiling of tobacco leaf trichomes identifies genes for biotic and abiotic stresses. Plant Cell Physiol 51:1627–1637

Horsch RB, Fry J, Hoffmann N, Neidermeyer J, Rogers SG, Fraley RT (1993) Leaf disk transformation. In: Gelvin SB, Schilperoort RA, Verma DPS (eds) Plant molecular biology manual. Kluwer Academic Publishers, Dordrecht, pp A5/1–A5/9

Huson DH, Richter DC, Rausch C, Dezulian T, Franz M, Rupp R (2007) Dendroscope: an interactive viewer for large phylogenetic trees. BMC Bioinform 8:460

Jakobsen K, Klemsdal SS, Aalen RB, Bosnes M, Alexander D, Olsen OA (1989) Barley aleurone development: molecular cloning of aleurone-specific cDNAs from immature grains. Plant Mol Biol 12:285–293

Jefferson RA, Kavanagh TA, Bevan MW (1987) GUS fusions: β-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J 6:3901–3907

Joët T, Laffargue A, Salmona J, Doulbeau S, Descroix F, Bertrand B, De Kochko A, Dussert S (2009) Metabolic pathways in tropical dicotyledonous albuminous seeds: *Coffea arabica* as a case study. New Phytol 182:146–162

Kader JC (1996) Lipid transfer proteins in plants. Annu Rev Plant Physiol Plant Mol Biol 47:627–654

Kader JC (1997) Lipid-transfer proteins: a puzzling family of plant proteins. Trends Plant Sci 2:66–70

Kalla R, Shimamoto K, Potter R, Nielsen PS, Linnestad C, Olsen OA (1994) The promoter of the barley aleurone specific gene encoding a putative 7-kDa lipid transfer protein confers aleurone cell-specific expression in transgenic rice. Plant J 6:849–860

Katoh K, Toh H (2008) Recent developments in the MAFFT multiple sequence alignment program. Brief Bioinform 9:286–298

Kawagoe Y, Murai N (1992) Four distinct nuclear proteins recognize in vitro the proximal promoter of the bean seed storage protein β-phaseolin gene conferring spatial and temporal control. Plant J 2:927–936

Kharabian A (2010) An efficient computational method for screening functional SNPs in plants. J Theor Biol 265:55–62

Krause A, Sigrist CJ, Dehning I, Sommer H, Broughton WJ (1994) Accumulation of transcripts encoding a lipid transfer-like protein during deformation of nodulation-competent *Vigna unguiculata* root hairs. Mol Plant Microbe Interact 7:411–418

Kristensen AK, Brunstedt J, Nielsen KK, Roepstorff P, Mikkelsen JD (2000) Characterization of a new antifungal non-specific

lipid transfer protein (nsLTP) from sugar beet leaves. Plant Sci 155:31–40

Larkin PD, Park WD (2003) Association of waxy gene single nucleotide polymorphisms with starch characteristics in rice (*Oryza sativa* L.). Mol Breed 12:335–339

Lashermes P, Andrade AC, Etienne H (2008) Genomics of coffee, one of the world's largest traded commodities. In: Moore H, Ming R (eds) Genomics of tropical crop plants. Springer, Berlin, pp 203–226

Leroy T, Ribeyre F, Bertrand B, Charmetant P, Dufour M, Montagnon C, Marraccini P, Pot D (2006) Genetics of coffee quality. Braz J Plant Physiol 18:229–242

Lessard PA, Allen RD, Bernier F, Crispino JD, Fujiwara T, Beachy RN (1991) Multiple nuclear factors interact with upstream sequences of differentially regulated β-conglycinin genes. Plant Mol Biol 16:397–413

Lin CW, Mueller LA, Mc Carthy J, Crouzillat D, Pétiard V, Tanksley SD (2005) Coffee and tomato share common gene repertoires as revealed by deep sequencing of seed and cherry transcripts. Theor Appl Genet 112:114–130

Lindorff-Larsen K, Winther JR (2001) Surprisingly high stability of barley lipid transfer protein, LTP1, towards denaturant, heat and proteases. FEBS Lett 488:145–148

Linnestad C, Lönneborg A, Kalla R, Olsen OA (1991) Promoter of a lipid transfer protein gene expressed in barley aleurone cells contains similar *myb* and *myc* recognition sites as the maize *Bz-McC* allele. Plant Physiol 97:841–843

Lüscher B, Eisenman RN (1990) New light on Myc and Myb. Part II. Myb. Genes Dev 4:2235–2241

Marchler-Bauer A, Zheng C, Chitsaz F et al (2013) CDD: conserved domains and protein three-dimensional structure. Nucl Acids Res 41(D1):D348–D352

Marion D, Douliez JP, Michon T, Elmorjani K (2000) Structure, biological and technological functions of lipid transfer proteins and indolines, the major lipid binding proteins from cereal kernels. J Cereal Sci 32:1–20

Marraccini P, Deshayes A, Pétiard V, Rogers WJ (1999) Molecular cloning of the complete 11S seed storage protein gene of *Coffea arabica* and promoter analysis in transgenic tobacco plants. Plant Physiol Biochem 37:273–282

Marraccini P, Courjault C, Caillet V, Lausanne F, Lepage B, Rogers WJ, Tessereau S, Deshayes A (2003) Rubisco small subunit of *Coffea arabica*: cDNA sequence, gene cloning and promoter analysis in transgenic tobacco plants. Plant Physiol Biochem 41:17–25

Marraccini P, Freire LP, Alves GSC et al (2011) *RBCS1* expression in coffee: *Coffea* orthologs, *Coffea arabica* homeologs, and expression variability between genotypes and under drought stress. BMC Plant Biol 11:85

Marraccini P, Vinecky F, Alves GSC et al (2012) Differentially expressed genes and proteins upon drought acclimation in tolerant and sensitive genotypes of *Coffea canephora*. J Exp Bot 63:4191–4212

Molina A, Segura A, García-Olmedo F (1993) Lipid transfer proteins (nsLTPs) from barley and maize leaves are potent inhibitors of bacterial and fungal plant pathogens. FEBS Lett 316:119–122

Mondego JMC, Vidal RO, Carazzolle MF et al (2011) An EST-based analysis identifies new genes and reveals distinctive gene expression features of *Coffea arabica* and *Coffea canephora*. BMC Plant Biol 11:30

Mueller LA, Solow TH, Taylor N et al (2005) The SOL genomics network. A comparative resource for Solanaceae biology and beyond. Plant Physiol 138:1310–1317

Murashige T, Skoog F (1962) A revised medium for rapid growth and bio assays with tobacco tissue cultures. Physiol Plant 15:473–497

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

Ostergaard J, Vergnolle C, Schoentgen F, Kader JC (1993) Acyl-binding lipid-transfer proteins from rape seedlings, a novel category of proteins interacting with lipids. Biochim Biophys Acta 1170:109–117

Petitot AS, Lecouls AC, Fernandez D (2008) Sub-genomic origin and regulation patterns of a duplicated WRKY gene in the allotetraploid species *Coffea arabica*. Tree Genet Genomes 3:379–390

Pons JL, Labesse G (2009) @TOME-2: a new pipeline for comparative modeling of protein-ligand complexes. Nucleic Acids Res 37(Suppl 2):W485–W491

Pyee J, Yu H, Kolattukudy PE (1994) Identification of a lipid transfer protein as the major protein in the surface wax of broccoli (*Brassica oleracea*) leaves. Arch Biochem Biophys 311:460–468

Ramakers C, Ruijter JM, Deprez RH, Moorman AF (2003) Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. Neurosci Lett 339:62–66

Rasmussen TB, Donaldson IA (2006) Investigation of the endosperm specific sucrose synthase promoter from rice using transient expression of reporter genes in guar seed tissue. Plant Cell Rep 25:1035–1042

Rizhsky L, Liang H, Shuman J, Shulaev V, Davletova S, Mittler R (2004) When defense pathways collide: the response of Arabidopsis to a combination of drought and heat stress. Plant Physiol 134:1683–1696

Salmona J, Dussert S, Descroix F, de Kochko A, Bertrand B, Joët T (2008) Deciphering transcriptional networks that govern *Coffea arabica* seed development using combined cDNA array and real-time RT-PCR approaches. Plant Mol Biol 66:105–124

Samuel D, Liu YJ, Cheng CS, Lyu PC (2002) Solution structure of plant nonspecific lipid transfer protein-2 from rice (*Oryza sativa*). J Biol Chem 277:35267–35273

Saunders CT, Baker D (2002) Evaluation of structural and evolutionary contributions to deleterious mutations prediction. J Mol Biol 332:891–901

Schmutz J, Cannon SB, Schlueter J et al (2010) Genome sequence of the palaeopolyploid soybean. Nature 14:178–183

Shahmuradov IA, Gammerman AJ, Hancock JM, Bramley PM, Solovyev VV (2003) PlantProm: a database of plant promoter sequences. Nucl Acids Res 31:114–117

Shi J, Blundell TL, Mizuguchi K (2001) FUGUE: sequence structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. J Mol Biol 310:243–257

Shirsat AH, Wilford N, Croy R, Boulter D (1989) Sequences responsible for tissue specific promoter activity of a pea legumin gene in tobacco. Mol Gen Genet 215:326–331

Simpson SD, Nakashima K, Narusaka Y, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (2003) Two different novel *cis*-acting elements of *erd1*, a *clpA* homologous *Arabidopsis* gene function in induction by dehydration stress and dark-induced senescence. Plant J 33:259–270

Skriver K, Leah R, Müller-Uri F, Olsen FL, Mundy J (1992) Structure and expression of the barley lipid transfer protein gene *Ltp1*. Plant Mol Biol 18:585–589

Söding J (2005) Protein homology detection by HMM–HMM comparison. Bioinformatics 21:951–960

Speer K, Kölling-Speer I (2006) The lipid fraction of the coffee bean. Braz J Plant Physiol 18:201–216

Takishima K, Watanabe S, Yamada M, Mamiya G (1986) The amino-acid sequence of the nonspecific lipid transfer protein from germinated castor bean endosperms. Biochim Biophys Acta 870:248–255

Tapia G, Morales-Quintana L, Parra C, Berbel A, Alcorta M (2013) Study of nsLTPs in *Lotus japonicus* genome reveal a specific epidermal cell member (LjLTP10) regulated by drought stress in aerial organs with a putative role in cutin formation. Plant Mol Biol 82:485–501

Tatematsu K, Ward S, Leyser O, Kamiya Y, Nambara E (2005) Identification of *cis*-elements that regulate gene expression during initiation of axillary bud outgrowth in *Arabidopsis*. Plant Physiol 138:757–766

Tchang F, This P, Stiefel V et al (1988) Phospholipid transfer protein: full-length cDNA and amino-acid sequence in maize. Aminoacid sequence homologies between plant phospholipid transfer proteins. J Biol Chem 263:16849–16855

Terzaghi WB, Cashmore AR (1995) Light-regulated transcription. Annu Rev Plant Physiol Plant Mol Biol 46:445–474

Thoma SL, Kaneko Y, Somerville C (1993) An Arabidopsis lipid transfer protein is a cell wall protein. Plant J 3:427–437

Thoma S, Hecht U, Kippers A, Botella J, De Vries S, Somerville CR (1994) Tissue-specific expression of a gene encoding a cell wall-localized lipid transfer protein from *Arabidopsis*. Plant Physiol 105:35–45

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. Nucl Acids Res 22:4673–4680

Trevino MB, O'Connell MA (1998) Three drought responsive members of the nonspecific lipid-transfer protein gene family in *Lycopersicon pennellii* show different developmental patterns of expression. Plant Physiol 116:1461–1468

Vaast P, Bertrand B, Perriot JJ, Guyot B, Génard M (2006) Fruit thinning and shade improve bean characteristics and beverage quality of coffee (*Coffea arabica* L.) under optimal conditions. J Sci Food Agric 86:197–204

Vidal RO, Mondego JMC, Pot D, Ambrosio AB, Andrade AC, Pereira LFP, Colombo CA, Vieira LGE, Carazzolle MF, Pereira GAG (2010) A high-throughput data mining of single nucleotide polymorphisms in *Coffea* species expressed sequence tags suggests differential homeologous gene expression in the allotetraploid *Coffea arabica*. Plant Physiol 154:1053–1066

Vidal R, Alekcevetch JC, Leroy T, de Bellis F, Pot D, Rodrigues GC, Carazzolle M, Pereira GAG, Andrade AC, Marraccini P (2013) High-throughput sequencing of cDNA shows that *cv.* Rubi and IAPAR59 of *Coffea arabica* have different molecular response to water privation. In: 24th International colloquium on the chemistry of Coffee, ASIC, Paris

Vieira LGE, Andrade AC, Colombo CA et al (2006) Brazilian coffee genome project: an EST-based genomic resource. Braz J Plant Physiol 18:95–108

Vincentz M, Leite A, Neshich G (1997) ACGT and vicilin core sequences in a promoter domain required for seed-specific expression of a 2S storage protein gene are recognized by the opaque-2 regulatory protein. Plant Mol Biol 34:879–889

Wang NJ, Lee CC, Cheng CS, Lo WC, Yang YF, Chen MN, Lyu PC (2012) Construction and analysis of a plant non-specific lipid transfer protein database (nsLTPDB). BMC Genom 13(Suppl 1):S9

Wilson AJ, Petracco M, Illy E (1997) Some preliminary investigations of oil biosynthesis in the coffee fruit and its subsequent re-distribution within green and roasted beans. In: 17th International colloquium on the chemistry of Coffee, ASIC, Paris

Wu CY, Washida H, Onodera Y, Harada K, Takaiwa F (2000) Quantitative nature of the Prolamin-box, ACGT and AACA motifs in a rice glutelin gene promoter: minimal *cis*-element requirements for endosperm-specific gene expression. Plant J 23:415–421

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

Yubero-Serrano EM, Moyano E, Medina-Escobar N, Muñoz-Blanco J, Caballero JL (2003) Identification of a strawberry gene encoding anon-specific lipid transfer protein that responds to ABA, wounding and cold stress. J Exp Bot 54:1865–1877

Zhou H, Zhou Y (2005) SPARKS 2 and SP3 servers in CASP6. Proteins 61:152–156

Zottich U, Da Cunha M, Carvalho AO (2011) Purification, biochemical characterization and antifungal activity of a new lipid transfer protein (LTP) from *Coffea canephora* seeds with α-amylase inhibitor properties. Biochim Biophys Acta 1810:375–383

Mofatto et al. BMC Plant Biology (2016) 16:94
DOI 10.1186/s12870-016-0777-5

BMC Plant Biology

RESEARCH ARTICLE Open Access

Identification of candidate genes for drought tolerance in coffee by high-throughput sequencing in the shoot apex of different *Coffea arabica* cultivars

Luciana Souto Mofatto[1†], Fernanda de Araújo Carneiro[2†], Natalia Gomes Vieira[2†], Karoline Estefani Duarte[2], Ramon Oliveira Vidal[1], Jean Carlos Alekcevetch[2], Michelle Guitton Cotta[2], Jean-Luc Verdeil[3], Fabienne Lapeyre-Montes[3], Marc Lartaud[3], Thierry Leroy[3], Fabien De Bellis[3], David Pot[3], Gustavo Costa Rodrigues[4], Marcelo Falsarella Carazzolle[1], Gonçalo Amarante Guimarães Pereira[1], Alan Carvalho Andrade[2,5] and Pierre Marraccini[2,3*]

Abstract

Background: Drought is a widespread limiting factor in coffee plants. It affects plant development, fruit production, bean development and consequently beverage quality. Genetic diversity for drought tolerance exists within the coffee genus. However, the molecular mechanisms underlying the adaptation of coffee plants to drought are largely unknown. In this study, we compared the molecular responses to drought in two commercial cultivars (IAPAR59, drought-tolerant and Rubi, drought-susceptible) of *Coffea arabica* grown in the field under control (irrigation) and drought conditions using the pyrosequencing of RNA extracted from shoot apices and analysing the expression of 38 candidate genes.

Results: Pyrosequencing from shoot apices generated a total of 34.7 Mbp and 535,544 reads enabling the identification of 43,087 clusters (41,512 contigs and 1,575 singletons). These data included 17,719 clusters (16,238 contigs and 1,575 singletons) exclusively from 454 sequencing reads, along with 25,368 hybrid clusters assembled with 454 sequences. The comparison of DNA libraries identified new candidate genes ($n = 20$) presenting differential expression between IAPAR59 and Rubi and/or drought conditions. Their expression was monitored in plagiotropic buds, together with those of other ($n = 18$) candidates genes. Under drought conditions, up-regulated expression was observed in IAPAR59 but not in Rubi for *CaSTK1* (protein kinase), *CaSAMT1* (SAM-dependent methyltransferase), *CaSLP1* (plant development) and *CaMAS1* (ABA biosynthesis). Interestingly, the expression of lipid-transfer protein (nsLTP) genes was also highly up-regulated under drought conditions in IAPAR59. This may have been related to the thicker cuticle observed on the abaxial leaf surface in IAPAR59 compared to Rubi.

(Continued on next page)

* Correspondence: marraccini@cirad.fr
†Equal contributors
[2]Embrapa Recursos Genéticos e Biotecnologia (LGM-NTBio), Parque Estação Biológica, CP 02372, 70770-917, Brasilia, DF, Brazil
[3]CIRAD UMR AGAP, F-34398 Montpellier, France
Full list of author information is available at the end of the article

(Continued from previous page)

Conclusions: The full transcriptome assembly of *C. arabica*, followed by functional annotation, enabled us to identify differentially expressed genes related to drought conditions. Using these data, candidate genes were selected and their differential expression profiles were confirmed by qPCR experiments in plagiotropic buds of IAPAR59 and Rubi under drought conditions. As regards the genes up-regulated under drought conditions, specifically in the drought-tolerant IAPAR59, several corresponded to orphan genes but also to genes coding proteins involved in signal transduction pathways, as well as ABA and lipid metabolism, for example. The identification of these genes should help advance our understanding of the genetic determinism of drought tolerance in coffee.

Keywords: Candidate gene, Coffee, Drought, Differential gene expression, RNA-Seq, Real-time PCR (RT-qPCR)

Background

Coffee is the single most important tropical commodity traded worldwide and is a source of income for many developing countries in Tropics [1]. In the coffee genus, *Coffea arabica* accounts for approximately 70 % of total production worldwide, estimated at 8.5 million tons in 2015 [2]. Coffee production is subject to regular fluctuations mainly due to the natural biennial cycle but also caused by adverse climatic effects. Among them, drought is a widespread limiting factor and affects flowering and bean development, hence coffee yield [3]. Marked variations in rainfall also increase bean defects and modify the biochemical composition of beans, hence the final quality of the beverage [4]. Periods of drought may become more pronounced as a consequence of global climate change and geographical coffee growing regions may shift considerably, leading to environmental, economic and social problems [5]. In such a context, the creation of drought-tolerant coffee varieties has now become a priority for coffee research.

Genetic variability for drought tolerance exits in the coffee genus, particularly in *Coffea canephora* [6, 7] but also in *C. arabica* [8]. Although molecular mechanisms of drought tolerance have been widely studied in model plants [9], they are less well understood in *Coffee sp*. In a previous study analysing the effects of drought on gene expression, we recently identified a set of 30 genes differentially expressed in the leaves of drought-tolerant and drought-susceptible clones of *C. canephora* grown in the greenhouse under control (unstressed) and drought conditions [10, 11]. In that case, the expression of genes encoding glycine-rich proteins, heat shock proteins, dehydrins, ascorbate peroxidase, as well as trans-acting factors (such as DREB1D), for example, increased under drought conditions.

In *Coffea sp.*, EST resources have been developed for various species and tissues including roots, leaves, and fruits [12–16]. However, no genomic resources are available for shoot apices, which are considered as key organs for plant development by integrating several signals, such as environmental stimuli as well as hormones (abscisic acid [ABA], auxins, cytokinins) and transcription [17]. On the other hand, next-generation sequencing (NGS) provides new opportunities to study transcriptomic responses and to combine high-throughput sequencing with the functional annotation capacity of generated ESTs [18].

In order to identify candidate genes involved in drought tolerance in coffee plants, we collected the shoot apices from drought-tolerant IAPAR59 and drought-susceptible Rubi cultivars of *C. arabica* under control and drought conditions to generate libraries that were sequenced using the GS-FLX Titanium strategy. A reference full transcriptome was annotated and compared to pre-identify genes differentially expressed between cultivars and drought conditions. The transcription profiles of these genes were further analysed by qPCR in the plagiotropic buds of these plants.

Methods

Plant material

We compared two cultivars of *Coffea arabica*, the drought-susceptible (D^S) Rubi MG1192 (also referred to hereafter as RUB) and the drought-tolerant (D^T) IAPAR59 (also referred to hereafter as I59). Rubi did not undergo recent introgression with *C. canephora* genomic DNA, while IAPAR59 is the result of a cross between the Timor hybrid HT832/2 and the Villa Sarchi cultivar [19].

Field experiment

Seeds of these two commercial cultivars came from fruits harvested in May 2007 in the coffee experimental fields of the Institute for Research and Rural Assistance (Incaper, Vitoria, Espirito Santo, Brazil) and germinated (September 2007) in greenhouse of this institute. Five-month-old plantlets of the Rubi and IAPAR59 were then planted (January 2008) in a field experiment (0.7 m spacing between plants and 3 m spacing between rows) at the Cerrado Agricultural Research Center (Planaltina-DF, Brazil 15°35'44"S - 47°43'52"W) under full-sunlight conditions in two blocks of 30 plants for each cultivar. Under the conditions of the Cerrado climate [20], the rainfall pattern is divided into a dry season (from May to September) followed by a wet season (from October to April) that concentrates more than 80 % of annual

precipitations. For each cultivar, one control (C) block was irrigated while the drought (D) block was not irrigated during the dry seasons. For the control condition, irrigation was supplied by sprinklers (1.5 m in height) set up in the field in such a way that irrigation was uniform. Soil water content was monitored using PR2 profile probes (Delta-T Devices Ltd), and irrigation was applied regularly so as to maintain a moisture content above 0.27 cm^3 $H_2O.cm^{-1}$.

Sampling

For both cultivars and experiments, leaf predawn water potentials (Ψ_{pd}) were measured once a week during the 2009 dry season (from May to October) of (23-month-old plants) and only once in 2011 (at the end of the dry season) (47-month-old plants) using a Scholander-type pressure chamber (Plant Water Status Console, Model 3000 F01, Soil Moisture Equipment Corp, Santa Barbara, CA USA) in fully expanded leaves (8–15 cm long) from the third pair from the apex of plagiotropic branches located in the upper third of the plant canopy. For 454 sequencing, between 30 and 50 shoot apices were collected (between 10:00 and 11:00 am) from three different plants at the end of the dry season from Rubi and IAPAR59 under the control and drought conditions, and further dissected to isolate the shoot apex (Fig. 1b). For microscopic analyses, leaves identical to those used for Ψ_{pd} measurements were also collected from the same plants. At the end of the 2011 dry season, Ψ_{pd} were measured once for Rubi and IAPAR59 plants under control and drought treatments, and shoot apices were collected (Fig. 1a) for gene expression analyses (qPCR).

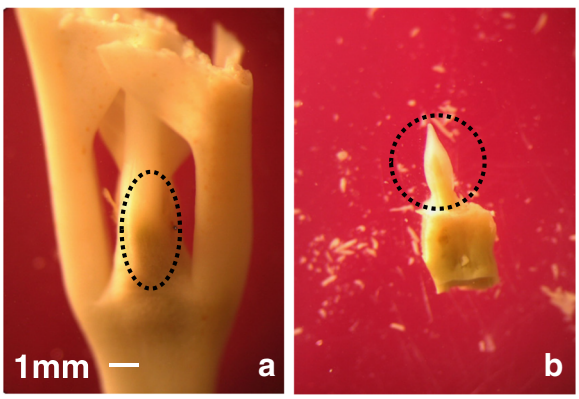


Fig. 1 Tissue dissection of plagiotropic buds. **a** The plagiotropic buds (including small leaves) were collected from plants during the 2011 dry season and used to extract RNA for qPCR expression analysis. **b** Meristem and leaf primordium dissected from plagiotropic buds harvested during the 2009 dry season and used to extract RNA for pyrosequencing. The dotted circles show the position of meristem and leaf primordium. The same scale (white bar = 1 mm) is used for both documents

RNA isolation, DNA synthesis and 454-sequencing

The plagiotropic buds were incubated for 5 min in the washing buffer (66 % chloroform, 33 % methanol, 1 % HCl) [21] and further incubated twice for 30 min under a vacuum in the fixation buffer (25 % acetic acid, 75 % ethanol RNAse-free) then cooled to 4 °C. Samples were stored in 75 % RNAse-free ethanol. For the control and drought conditions, shoot apices (meristems and primordium leaves) of three different plants were separated from plagiotropic buds under a binocular microscope by dissection and then ground to powder in liquid nitrogen using a pestle and mortar. Total RNA was extracted using the Nucleospin RNA Plant kit (Macherey-Nagel), including a DNAse-I treatment. The quality and quantity of RNA were checked with a Bioanalyzer (2100, RNA Nano 6000 Agilent). The 1st strand cDNA synthesis was performed using 1 μg total RNA and the SMARTer™ PCR cDNA Synthesis Kit (Clontech). Double-stranded DNA was then produced for each library (I59-C, I59-D, RUB-C and RUB-D). For each sample, DNA (around 5 μg) was nebulized to a mean fragment size of 650 bp, ligated to an adapter using standard procedures [22] and then sequenced by performing two runs (1 library per DNA sample x 2) using GS-FLX Titanium (Beckman Coulter Genomics SA, Grenoble, France) which generated one million reads corresponding to more than 255 Mb.

Transcriptome assembly and automatic annotation

All 454-sequencing reads were inspected for low quality reads and 454 adapters that were identified by SSAHA2 software [23]. A reference full transcriptome was then built using *C. arabica* reads originating from the present project and from the Brazilian Coffee Genome Project (BCGP) available in the GenBank public database [14, 24]. The Sanger and 454 reads were submitted for a trimming pipeline using *bdtrimmer* software [25] that was used to exclude ribosomal, vector, low quality (regions with a PHRED score less than 20) and short sequences (less than 100 bp). All sequences (454 and Sanger reads) were assembled using MIRA software [26]. The contigs formed by only Sanger reads were discarded from the full transcriptome assembly. The reference full transcriptome was annotated by Blast2GO software version 2.8 [27] using Non-Redundant protein (NCBI/NR), InterPro and Gene Ontology (GO) databases. The same program was also used to group datasets in GO according to the biological process. Further details on the automatic annotation of all contigs are provided in Additional file 1: Table S1. The complete bioinformatic pipeline used for this work is described in Additional file 2: Figure S1.

Digital gene expression analysis

The reference full transcriptome was also used to count all 454 reads/libraries individually by parsing the ACE

file generated by MIRA software. The number of sequences anchored in each contig (read counts) was subjected to differential expression analysis between the libraries using DEseq [28] and EdgeR [29] software in the R/Bioconductor package. A unigene was considered as differentially expressed when it was identified in at least one software considering fold-change ≥ 2 (or fold-change ≤ -2) and p-value ≤ 0.05. The libraries were compared based on (1) differentially expressed genes in IAPAR59 between C (control) and D (drought) conditions (with the calculation of fold-change based on the I59-D/I59-C ratio), (2) differentially expressed genes in Rubi between C and D conditions (RUB-D/RUB-C), (3) differentially expressed genes in the control library between Rubi and IAPAR59 (RUB-C/I59-C) and (4) differentially expressed genes in the drought library between Rubi and IAPAR59 (RUB-D/I59-D). Further information about differentially expressed genes in all the libraries is given in Additional file 3: Table S2.

Functional annotation of differentially expressed genes

The lists of differentially expressed genes in each analysis were separated into UP and DOWN regulated and subjected to GO enrichment analysis to identify significantly enriched GO slim terms (Plant GO slim) using Blast2GO software and a p-value ≤ 0.05.

Selection of candidate genes

The comparison of DNA libraries led to the identification of 80 (20 for each library) candidate genes (CGs) that were up- and down-regulated (see Additional file 3: Table S2). For each CG, primer pairs were designed using Primer Express software (Applied Biosystems) and tested of their specificity and efficiency against a mix of ss-DNAs of plagiotropic buds (data not shown). The best primer pairs ($n = 20$) were used to monitor the expression of corresponding CGs in plagiotropic buds of Rubi and IAPAR59 under control and drought conditions. These genes corresponded to *CaAEP1*, *CaCAB2*, *CaCHI1*, *CaCHI2*, *CaCHI3*, *CaDLP1*, *CaELIP3*, *CaGAS2*, *CaGRP2*, *CaH2A*, *CaHSP3*, *CaIPS1*, *CaJAMT1*, *CaMAS1*, *CaPP2*, *CaPSBB*, *CaSAMT1*, *CaSDC1*, *CaSLP1* and *CaSTK1* (Table 1). This list of CGs was increased by adding other genes such as 14 orphan genes (*CaUNK2-CaUNK7*, *CaUNK9* and *CaUNK11-CaUNK17* already described to present differential gene expression profiles in different organs of *C. canephora* [30]. This list was finally completed by including the *CaUNK1*, *CaUNK8* and *CaUNK10* orphan genes, and *LTP* genes that were already studied in *C. canephora* [10, 11, 31] and *C. arabica* [32], respectively.

Real-time quantitative PCR assays

For qPCR experiments, plagiotropic buds containing shoot apices and small leaves (Fig. 1a) were immediately frozen in liquid nitrogen after collection, and stored at -80 °C before being extracted and converted into single-strand cDNA as previously described [33]. Real-time qPCR assays were carried out using the protocol recommended for the use of 7500 Fast Real-Time PCR Systems (Applied Biosystems, Foster City, CA, USA). DNA preparations were diluted (1/50) and tested by qPCR using CG primer pairs (Table 1). RT-qPCR was performed with 1 μl of diluted ss-DNA and 0.2 μM (final concentration) of each primer in a final volume of 10 μl with SYBR green fluorochrome (SYBRGreen qPCR Mix-UDG/ROX, Invitrogen). The reaction mixture was incubated for 2 min at 50 °C (Uracil DNA-Glycosilase treatment), then for 5 min at 95 °C (inactivation of UDGase), followed by 40 amplification cycles of 3 sec at 95 °C and finally for 30 sec at 60 °C. Data were analysed using SDS 2.1 software (Applied Biosystems) to determine cycle threshold (Ct) values. The specificity of the PCR products generated for each set of primers was verified by analysing the Tm (dissociation) of amplified products. PCR efficiency (E) was estimated using absolute fluorescence data captured during the exponential phase of amplification of each reaction with the equation E (in %) = $(10^{(-1/slope)} -1) \times 100$ [34]. Efficiency values were taken into account in all subsequent calculations. Gene expression levels were normalized to expression levels of *CaUBQ10* as a constitutive reference. Relative expression was quantified by applying the formula $(1 + E)^{-\Delta\Delta Ct}$ where $\Delta Ct_{\text{target}} = Ct_{\text{target gene}} - Ct_{\text{reference gene}}$ and $\Delta\Delta Ct = \Delta Ct_{\text{target}} - \Delta Ct_{\text{internal calibrator}}$, with the internal reference always being the Rubi-control (RUB-C) sample with relative expression equal to 1.

Leaf histological analysis of cuticle

Mature leaves of the IAPAR59 and Rubi genotypes were fixed for 48 h in 100 mM phosphate buffer at pH 7.2, supplemented with 1 % (v/v) glutaraldehyde, 2 % (v/v) paraformaldehyde, and 1 % (w/v) caffeine, at room temperature [35]. The samples were dehydrated and embedded in Technovit 7100 resin (Heraeus Kulzer) according to the manufacturer's recommendations. Three-micrometer semi-thin sections were cut with glass knives on a Leica RM2065 Microtome. The resulting sections were double stained according to Buffard-Morel et al. [36]. Briefly, polysaccharides were stained dark pink with periodic acid Schiff (PAS) and soluble proteins were stained blue with naphthol blue-black (NBB) [37]. Sections were then mounted in Mowiol. The slides were observed with a Leica DM6000 microscope (Leica, Germany) under bright field or epifluorescent light (A4 filter). Pictures were taken with a Retiga 2000R camera (QImaging Co.) and the images were processed with Volocity 4.0.1 (Improvision, Lexington, MA, USA). Cuticle thickness was measured with the freeware Image J software (http://imagej.nih.gov/ij/). Experiments were conducted on

Table 1 Candidate genes and corresponding primers used for qPCR experiments

Gene	Protein name	*C. canephora*	GB	ATP	SGN	Primer	Primer sequences	bp
CaUBQ10	Ubiquitin	Cc02_g31600	GW488515	32782	U637098	BUBI-FBUBI-R	5′ AAGACAGCTTCAACAGAGTACAGCAT 3′ 5′ GGCAGGACCTTGGCTGACTATA 3′	104
CaAEP1[a]	Putative aldose 1-epimerase	Cc07_g03170	GT005185	716	U637659	716-1 F716-1R	5′ CGGTGATGTCCTCTCTGATGAG 3′ 5′ GTTGGGATGAGCTGGTTGTTC 3′	75
CaCAB2[a]	Chlorophyll a/b-binding protein	Cc09_g09030	GT003492	33540	U629601	48565-F48565-R	5′ GTTCAAGGCTGGATCCCAAA 3′ 5′ GCAAGCCCAGATAGCCAAGA 3′	100
CaCHI1[a]	Class III chitinase	Cc11_g00410	GT012279	32745	U637166	50103-F50103-R	5′ AATCAAGCGACCGTCCATTC 3′ 5′ GTGTTTCCGCTGTGGATGTG 3′	70
CaCHI2[a]	Putative chitinase	Cc00_g14300	GT011845	32737	U638035	53058-F53058-R	5′ CCTGCTCGCGGTTTCTACAC 3′ 5′ TTGTTCCAAAAGCCCCATTG 3′	70
CaCHI3[a]	Chitinase-like protein	Cc03_g13720	GW491433	32875	U645893	23638-F23638-R	5′ AAACGGCCCGTCCAGAA 3′ 5′ GCTTTGTCCTGCTGGTCCAT 3′	130
CaDLP1[a]	Dirigent-like protein	Cc00_g27410	GW477731	35149	nf	39577-F39577-R	5′ TTGGTAGTCCGGCGAGAGAA 3′ 5′ GCATATCCCCGAGCAAACCT 3′	70
CaELIP3[a]	Early light-induced protein (ELIP)	Cc03_g04300	GR985685	32771	U631550	32771-F32771-R	5′ TCGGTTGCCATGCAATCTT 3′ 5′ GCAGATGAAGCCCACAGCTT 3′	100
CaGAS2[a]	Glucosyltransferase arbutin synthase	Cc02_g39100	GT697284	3945	U632419	632419-F632419-R	5′ GCTGACGACGTTAGGATTGAGA 3′ 5′ AACTTGGCGGTGTCAACCAA 3′	101
CaGRP2[a]	Glycin-rich protein	Cc00_g16260	GW430980	32799	U635030	53139-1 F53139-1R	5′ CACATATGCTGGTGAGCCAAA 3′ 5′ AGGCATTTAAGCGCCATGAT 3′	100
CaH2A[a]	Putative histone H2A	Cc01_g12440	GT723387	33557	U630412	53417-F53417-R	5′ GCACTGGAGCTCCGGTCTAC 3′ 5′ AGCAGCATTTCCAGCCAATT 3′	80
CaHSP3[a]	Heat schock protein (HSP) 70 kDa	Cc02_g08040	GR982512	33197	U636531	33197-1 F33197-1R	5′ GGCGTCTGGCAACACGAT 3′ 5′ CGATGAGACGCTCGGTGTCT 3′	100
CaIPS1[a]	Myo-inositol 1-phosphate synthase	Cc07_g15530	GT003538	10496	U632517	10496-1 F10496-1R	5′ AAGCAACCTGAATTTGGCTGAT 3′ 5′ GAGAGGGACCATGGATTCCA 3′	100
CaJAMT1[a]	Jasmonate O-methyltransferase	Cc03_g07330	GR989151	33008	U631389	47327-F47327-R	5′ CTGTGGCTGAACCCTTGCTT 3′ 5′ TCTTTGGACATGCGATCAGAAA 3′	100
CaMAS1[a]	Momilactone-A synthase	Cc00_g13640	GW479615	33413	nf	33413-F33413-R	5′ GGGCAGAGGCACGAAAAA 3′ 5′ GGTACCCTGCCGCAACTATG 3′	60
CaPP2[a]	Putative phloem protein 2 (PP2)	Cc03_g13000	GR995691	33207	U633544	33207-F33207-R	5′ GGTGTTGGCGATGTCGAGAT 3′ 5′ TTCCTTGGGTCGAAGCTCAA 3′	90
CaPSBB[a]	Photosystem II CP47 (psbB)-like protein	nf	GW447378	22102	U630312	55586-F55586-R	5′ ATCGGAAATAATCCGGCAAA 3′ 5′ AACCATCCAATCGCTATTCCA 3′	80
CaSAMT1[a]	S-adenosyl-methionine-methyltransferase	Cc03_g05630	DV672716	754	U629783	34318-F34318-R	5′ AACGTTTGGGTGATGAATGTTG 3′ 5′ GTGCCAATAAGCCCTCTATCGT 3′	80
CaSDC1[a]	S-adenosyl-L-methionine decarboxylase	Cc11_g11130	GT002431	8508	U629687	8508-1 F8508-1R	5′ CTCGATTCCTCCCATCCTGAA 3′ 5′ TGACTGTGCCCCAGGGAATA 3′	100
CaSLP1[a]	Subtilisin-like protein	Cc00_g19100	GW430663	1620	nf	7961-F7961-R	5′ CCATCGTTCTCGGTGGTCTT 3′ 5′ GCATTGCTCCCCACATTCTT 3′	80
CaSTK1[a]	Hypothetical S/T protein kinase	Cc00_g18670	GT687049	6301	U631794	6301-1 F6301-1R	5′ CCACCCACAAGCTGTATTCTCA 3′ 5′ GACCCAATGGGATGTCATCAC 3′	80
CaUNK1[c]	Unknown protein 1	Cc03_g08880	DV689820	33062	U614843	182052-F182052-R	5′ TATAGTGTTTATGGTGTGGCTTTCAGT 3′ 5′ GTACCACCGTAGGGAGACGTATG 3′	79

Table 1 Candidate genes and corresponding primers used for qPCR experiments *(Continued)*

CaUNK2[b]	Unknown protein 2	Cc07_g01940	DV708962	31492	U637447	33353-F33353-R	5′ GAACTTACAAACGCGCGTAACC 3′ 5′ CATGGTCGAATCCAGATTTCATT 3′	80
CaUNK3[b]	Unknown protein 3	nf	nf	22823	nf	22823-F22823-R	5′ GGAAGCATGCACACAGAAAATAGA 3′ 5′ TTCCTGTTTACGTCTTTTTCAATTGA 3′	80
CaUNK4[b]	Unknown protein 4	Cc06_g11210	GW465088	39984	nf	55677-F55677-R	5′ GCTGTGGTTTTAAAGTTTTGATGGA 3′ 5′ TGCAAAATTAAGGTCCCAACAGT 3′	81
CaUNK5[b]	Unknown protein 5	Cc08_g09510	GW474926	4578	nf	4578-F4578-R	5′ GGAGTTCCTGTCCGAAGTTGTT 3′ 5′ GGCATGCTGTCACCTGAAAA 3′	80
CaUNK6[b]	Unknown protein 6	Cc03_g06850	GT002178	34993	U632634	34993-F34993-R	5′ AAGCCAATGCCGATCGATT 3′ 5′ CGCCGCCGAAGATCTCTAG 3′	100
CaUNK7[b]	Unknown protein 7	Cc03_g00560	GW444736	33613	U631416	25639-F25639-R	5′ CGAGGAAGCTGAAGGAAAGGA 3′ 5′ TCCGACTGGCCTAACAAGGT 3′	61
CaUNK8[b]	Unknown protein 8	Cc00_g04970	DV695331	33190	U640780	LP18101-FLP18100-R	5′ CTCGCGTGGCCGAGATC 3′ 5′ CCCTCACATTTCCACGTGAAT 3′	100
CaUNK9[b]	Unknown protein 9	Cc03_g08920	GT649500	32762	U636808	30926-F30926-R	5′ CGGAGGAGGCCATGGAGGT 3′ 5′ CCGTGTCCATAACCACCATGT 3′	123
CaUNK10[c]	Unknown protein 10	nf	GT648004	14813	U645073	D18240-FD18240-R	5′ TAGCCTTGTTCTTTTAGGGAGTCTTATC 3′ 5′ AGAGCTTCGTCCAGGAAGAAGA 3′	134
CaUNK11[b]	Unknown protein 11	Cc03_g14330	GR991912	8598	U637116	32792-F32792-R	5′ GCTGGGAAAGCTACAGAAACCA 3′ 5′ GAACTCCAACGCCAAGCATT 3′	100
CaUNK12[b]	Unknown protein 12	Cc10_g12840	nf	53029	nf	53029-F53029-R	5′ CTTCACACCATTCAGACAATCGA 3′ 5′ GACCGTAATTGGGCGTCAAT 3′	100
CaUNK13[b]	Unknown protein 13	Cc00_g17760	GT673421	14198	U639484	33980-F33980-R	5′ ATTGCCCTGTTTGCATGCAT 3′ 5′ CTGCATGGTGATTGTCCTCAGT 3′	100
CaUNK14[b]	Unknown protein 14	Cc00_g16260	GT672564	48325	U635030	11524-F11524-R	5′ GGCGGTTGTCATGGATACG 3′ 5′ TTTGGCTCACCAGCATATGTG 3′	119
CaUNK15[b]	Unknown protein 15	Cc00_g04970	GR983286	33190	U636790	05517-F05517-R	5′ AAAATTTCACCACGGCAAGCT 3′ 5′ TTGCCTCCCTCACATTTCCA 3′	72
CaUNK16[b]	Unknown protein 16	nf	GW464209	9761	U639049	18112-F18112-R	5′ TGTGAACTGCCATCCCAAGA 3′ 5′ AAGACTACCATGTCCAACAACTTCAG 3′	88
CaUNK17[b]	Unknown protein 17	Cc03_g08920	GT685623	32762	U636800	42747-F42747-R	5′ AGGTGGCTGCCAAGTCAGTT 3′ 5′ ATGGTACTTGGCTTCTCCTTCCTC 3′	71
CaLTP1[d]*CaLTP2*[d]	Non-specific lipid transfer protein (nsLTP)	Cc11_g09700	HG008739HG008740	46897	U632702	LTP-R2LTP-FT	5′ CACCATTACATGGGAACGTTGC 3′ 5′ CTGTGGTCTGAAATGGCCAACT 3′	120
CaLTP3[d]	Non-specific lipid transfer protein (nsLTP)	Cc04_g06890	HG008741	33368	U632702	LTP-R1LTP-FT	5′ ATTCAACACCATTACTAGTTTTCGAGC 3′ 5′ CTGTGGTCTGAAATGGCCAACT 3′	113
LTP[d]		nf		-	U632702	LTP-F100LTP-R100	5′ TGCAATTTTATCAAAGATCCAGC 3′ 5′ AGTTGGCCATTTCAGACCACA 3′	93

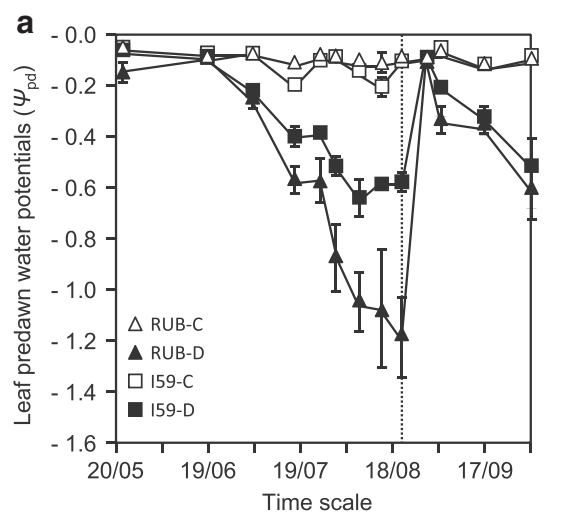
Gene names were assigned based on the best BLAST hit obtained by comparing the coffee ESTs with public databases. *C. canephora* means coffee sequences that aligned with the candidate genes using BLASTx searches against NR/NCBI and filtration (http://coffee-genome.org [59]). GenBank (GB: http://blast.ncbi.nlm.nih.gov/Blast.cgi), ATP (http://www.lge.ibi.unicamp.br/cirad/) and SGN (Sol Genomics Network, http://solgenomics.net/) accession numbers of coffee ESTs are also given, as well as the length of base pairs (bp) of amplicons. nf: no-hits found (SGN: tools/blast/SGN Clusters [current version] / Coffee species Clusters, GB: BLASTn/Nucleotide collection [nr/nt]). The size of amplicons is based on the unigene. ([a]): candidate genes ($n = 20$) identified during this study. ([b]): orphan genes ($n = 14$) previously described [35] and analysed in this study. ([c]): orphan genes ($n = 3$) with expression already been studied in leaves of D^T and D^S clones of *C. canephora* conilon [10, 11, 36]. ([d]): LTP-encoding genes were previously described [37]

the "Plate-Forme d'Histocytologie et Imagerie Cellulaire Végétale (PHIV platform)" (http://phiv.cirad.fr/) using microscopes belonging to the Montpellier Rio Imaging platform (www.mri.cnrs.fr). The results are expressed as means (μm) of 11 measured values. The data were statistically processed using (1) an analysis of variance computer program (Statistica, StatSoft, Inc.), and (2) the Student-Newman-Keuls (SNK) mean comparison test [38] when the effect of the factor tested was found to be statistically significant. A probability level of $P \leq 0.05$ was considered significant for all the statistical analyses.

Results

Monitoring drought under field conditions

In 2009, leaf predawn water potential (Ψ_{pd}) values were similar in the leaves of irrigated Rubi and IAPAR59 plants, ranging from -0.06 to -0.16 MPa (Fig. 2a). This confirmed the unstressed status of these plants which were considered as the control in our experiment. At the same time, the Ψ_{pd} values decreased gradually during the dry season in the leaves of Rubi and IAPAR59 under drought conditions reaching the lowest values at the end of the dry season (Fig. 2a). At that time, the less negative Ψ_{pd} values in IAPAR59 indicated that it had better access to soil water. The first rains then occurred and the Ψ_{pd} values of drought-stressed plants increased almost to those measured in irrigated plants, illustrating the complete recovery of stressed plants. In 2011, Ψ_{pd} was measured at the peak of the drought (end of dry season). Under drought conditions, both Rubi and IAPAR59 had similar Ψ_{pd} values that were more negative than those measured in 2009, indicating more severe drought stress in 2011 (Fig. 2b).



b

	Control	Drought
Rubi	-0.1 to -0.2	<-4.0
IAPAR59	-0.1 to -0.2	<-4.0

Fig. 2 Predawn leaf water potentials (Ψ_{pd}) measured in plants of *C. arabica*. Rubi (RUB, triangle) and IAPAR59 (I59, square) cultivars were grown under control (C, open symbols) and drought (D, black symbols) conditions. Ψ_{pd} values (expressed in mega-Pascal, MPa) were measured once a week during the 2009 dry season (23-month-old plants) (**a**). The time scale is in days and months (dd/mm, from 20/05 to 02/10). Vertical bars are standard deviations ($n = 9$ leaves) and the dashed vertical line (20/08) represents the harvest point of plagiotropic buds for RNA extraction for 454 sequencing and leaves for microscopic analyses. **b** Ψ_{pd} of Rubi and IAPAR59 plants (47-month-old plants) measured at the end of the 2011dry season. In this case, Ψ_{pd} values ranged from -0.1 to -0.2 MPa for the control conditions, but were below (< -4.0 MPa = severe drought) the range of use of a Scholander-type pressure chamber for drought conditions

Sequencing, assembly and annotation of the *Coffee* shoot apex transcriptome

The final reference assembly generated a total of 34,743,872 bp (34.7 Mbp) with coverage of 6.5x and 43,087 clusters, corresponding to 41,512 contigs and 1,575 singletons. These data are composed of: (1) 17,719 clusters (16,238 contigs and 1,575 singletons) from 454 sequences, exclusively; and (2) 25,368 hybrid clusters that contain 454 reads, and at least one contig from Sanger sequencing (public database). The contigs formed by only Sanger reads were discarded from the full transcriptome assembly. On average, 22.4 % and 55.6 % of the total raw data were discarded from Sanger and 454, respectively, due to low quality. After removing the adapters, these reads had a size of 379.2 bp (on average). The statistical data for the Sanger and 454 reads are listed in Table 2.

Transcriptome annotation by Blast2GO using Non-Redundant protein (NCBI/NR) and InterPro databases resulted in 36,965 transcriptome clusters (85.8 %) with a known protein function, 1,824 conserved proteins of unknown function (4.2 %), 1,515 proteins identified by InterPro only (3.5 %) and 2,783 unidentified proteins (6.5 % no-hits found).

Table 2 Characteristics of reads used in this work

Libraries	Total reads	Trimmed reads	Average length of reads
Public Sanger database	195,110	151,403	518
I59-C	135,304	66,641	325
I59-D	282,213	112,518	351
RUB-C	230,064	101,394	360
RUB-D	345,751	153,572	342
Total	1,188,442	585,528	379.2

Statistics of all reads used in this work: public Sanger reads and 454 sequenced reads from two cultivars under two conditions. Cultivars (RUB: Rubi and I59: IAPAR59) of *C. arabica* and treatments (*C* control and *D* drought) are indicated. The number of total reads, trimmed reads and average read length (in bp) are indicated

The results of the digital gene expression analysis (Table 3) showed more differentially expressed genes (DEG) in the cultivars Rubi (RUB) and IAPAR59 (I59) cultivars under drought (D) conditions (RUB-D/I59-D), totalling 490 clusters (1.14 % of the total), with 320 clusters classified as up-regulated. Under the control (C) conditions, a few DEG were found (RUB-C/I59-C), corresponding to 184 clusters (0.43 % of total clusters). The comparison between control and drought conditions showed a prevalence of up-regulated genes (165 clusters) and a total of 226 DEG in IAPAR59 (I59-D/I59-C) with 0.52 % of total clusters, and 343 clusters in Rubi (RUB-D/RUB-C) with 0.80 % of total clusters.

The results of the gene ontology (GO) enrichment analysis are shown in Fig. 3 and all GO enrichment data are listed in Additional file 1: Tables S1 and Additional file 3: Table S2. For IAPAR59, the comparison of drought and control conditions (I59-D/I59-C) identified over-represented GO terms characterized by up-regulated genes involved in expression (gALL_c3501) and translation (gALL_c2033, gALL_c4461, gALL_c6492) processes and in the generation of precursor metabolites and energy (gALL_c921, gALL_c4013, gALL_c4540). For Rubi, a comparison of the RUB-D/RUB-C libraries revealed an over-representation of the following GO terms which were up-regulated: protein metabolic process (gALL_c2021, gALL_c3355), response to stress (gALL_rep_c33197/*CaHSP3*) and response to abiotic stimulus (gALL_rep_c32771/*CaELIP3*, gALL_c2829, gALL_rep_c32766). When comparing both cultivars under drought conditions (RUB-D/I59-D), GO terms were identified related to increased enrichment of tropism for up-regulated genes (gALL_c1270, gALL_c1524, gALL_c1864) and photosynthesis for down-regulated genes (gALL_c27215, gALL_rep_c34074, gALL_rep_c34746). Under the control conditions (RUB-C/I59-C), proteins of translational machinery were identified for up-regulated genes (gALL_c3061, gALL_c16674, gALL_c19094) and photosynthesis for down-regulated genes (gALL_rep_c34074, gALL_rep_c37283, gALL_rep_c50892).

Expression profiles of candidate genes

Among the candidate genes (CGs) identified *in silico* as presenting up- and down-regulation, expression profiles from 20 of them were analysed by qPCR together with the expression of 17 orphan genes (3 of them already studied in *C. canephora* [10, 11, 30, 31]) and *LTP* genes [32]. For all these genes, expression profiles were analysed in plagiotropic buds of Rubi and IAPAR59 under control and drought conditions. These results are presented in separate sections below, according to the observed expression patterns.

Genes with induced expression under drought conditions

Twenty-five genes showing up-regulated expression profiles under drought conditions, mainly in IAPAR59 and to a lesser extent in Rubi, were identified (Fig. 4). This was observed for *CaSTK1* which encodes a putative oxidative stress response serine/threonine protein kinase with 87 % identity with a predicted protein of *Populus trichocarpa* (XP_002299433). In that case, expression of this gene was highly induced by drought in the D^T cultivar IAPAR59. Similar profiles were also observed for the *CaSAMT1* gene encoding a putative S-adenosyl-L-methionine-dependent methyltransferase and the orphan genes *CaUNK2* and *CaUNK3*. The latter gene had no open reading frame but presented high identity (*e*-value $2E^{-45}$) with the SGN-U637447 contig and also with various coffee ESTs mainly found in *C. canephora* cherries at early developmental stages (data not shown).

Expression of the *CaSLP1* gene encoding a putative protein homologous (65 % identity, 74 % similarity) to a protein of *Nicotiana benthamiana* containing a peptidase S8/subtilisin-related domain, was also higher in IAPAR59 than in Rubi under drought conditions. A similar situation was observed for the *CaMAS1*gene encoding a protein of 311 amino acid residues sharing similarities (*e*-value $2E^{-121}$, 66 % identity, 82 %, similarity) with momilactone A synthase-like protein from *Vitis vinifera* (XP_002275768) that contains a secoisolariciresinol dehydrogenase conserved domain.

Table 3 Reads showing differential expression between cultivars and/or treatments

Libraries	EdgeR DEG (% of total clusters)	DEseq DEG (% of total clusters)	Total DEG (% of total clusters)	Up-regulated clusters (% of total clusters)	Down-regulated clusters (% of total clusters)
I59-D/I59-C	209 (0.49 %)	176 (0.41 %)	226 (0.52 %)	165 (0.38 %)	61 (0.14 %)
RUB-D/RUB-C	323 (0.75 %)	306 (0.71 %)	343 (0.80 %)	251 (0.58 %)	92 (0.21 %)
RUB-C/I59-C	173 (0.40 %)	169 (0.39 %)	184 (0.43 %)	104 (0.24 %)	80 (0.19 %)
RUB-D/I59-D	392 (0.91 %)	433 (1.00 %)	490 (1.14 %)	320 (0.74 %)	170 (0.39 %)

Differentially expressed genes (*DEG*) were obtained with the R/Bioconductor packages DEseq and EdgeR. Total DEG values mean the union of DEseq and EdgeR results. The calculation of percentage was based on total of clusters (43,087 clusters). Cultivars (RUB Rubi and I59: IAPAR59) of *C. arabica* and treatments (*C* control and *D* drought) are indicated

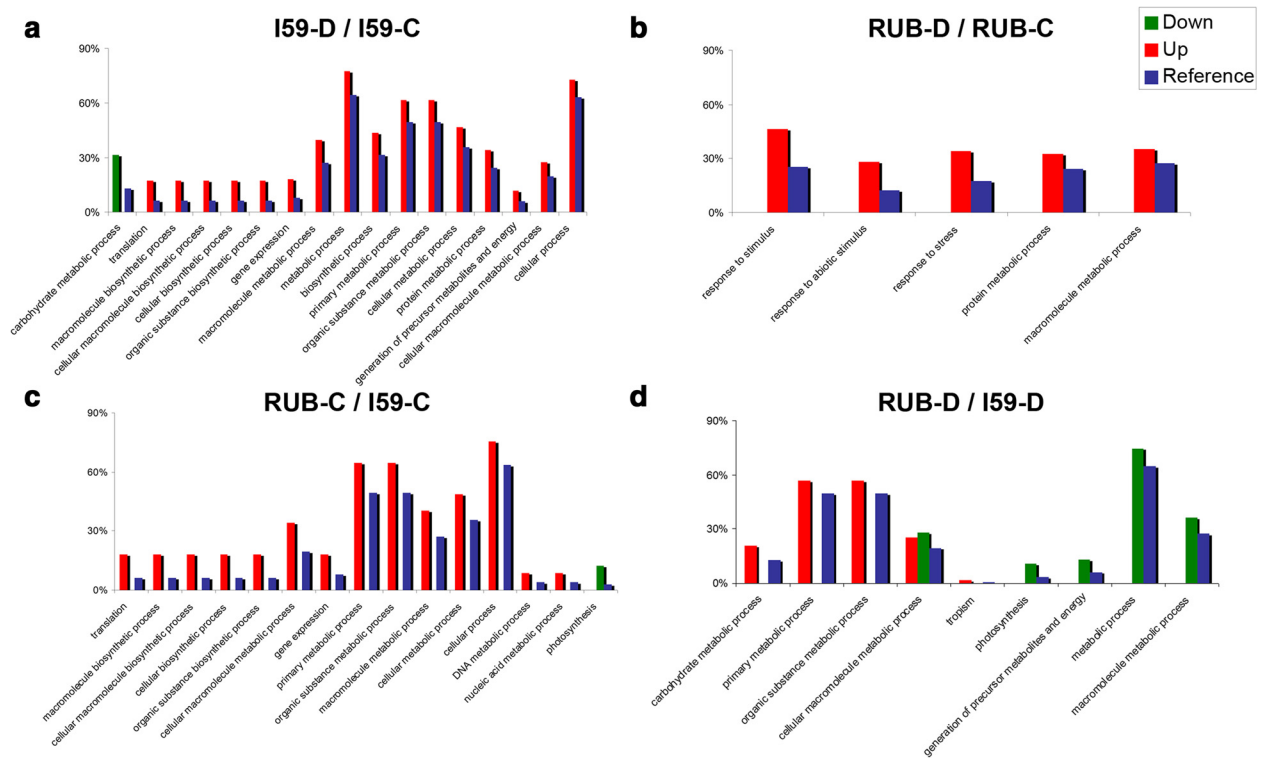


Fig. 3 Gene ontology (GO) enrichment analysis on a list of differentially expressed genes up- and down-regulated under four conditions. The calculation of fold change was based on the ratio of: (**a**) I59-D/I59-C; (**b**) RUB-D/RUB-C; (**c**) RUB-C/I59-C; and (**d**) RUB-D/I59-D. The Y axis indicates the number of genes normalized by the total number of genes used in each comparison from each library. Cultivars (RUB: Rubi and I59: IAPAR59) of *C. arabica* and treatments (C: control and D: drought) are indicated

Similar expression profiles, characterized by high up-regulation under drought conditions particularly in IAPAR59, were observed for the orphan genes *CaUNK1*, *CaUNK4*, *CaUNK5*, *CaUNK8*, and for *CaPSBB* (similar to the gene of *C. arabica* chloroplast genome encoding the photosystem II CP47 chlorophyll apoprotein) and *CaSDC1* encoding a putative protein related (81 % identity, 88 %, similarity) to the adenosylmethionine decarboxylase proenzyme of *Catharanthus roseus*). Expression of the *CaUNK6* gene was also induced under drought conditions but without significant difference in expression between the two cultivars.

Interestingly, the expression profiles of orphan genes *CaUNK7*, *CaUNK9*, *CaUNK10*, *CaUNK15*, *CaUNK16* and *CaUNK17* were similar to that of HSP-encoding gene *CaHSP3* in the sense that gene expression was highly up-regulated under drought conditions in both cultivars. In the case of *CaUNK10,* it is worth noting that expression increased 145- and 88-fold under drought conditions in Rubi and IAPAR59, respectively.

Under drought conditions, expression of the *CaGAS2* gene encoding a putative protein homologous (73 % identity, 86 % similarity) to the arbutin synthase from *Rauvolfia serpentina* (AJ310148), was slightly increased in IAPAR59 but reduced in Rubi. The *CaCAB2*, *CaCHI1* and *CaELIP3* genes encoding a photosystem II light harvesting chlorophyll A/B binding protein of *Gardenia jasminoides* (ACN41907), a class III chitinase of *C. arabica* (ADH10372) and an early light-induced protein (ELIP) of *Glycine max* (NP_001235754), respectively, showed similar profiles but with lower expression in Rubi than in IAPAR59, under control and drought conditions. Lastly, expression of the *CaPP2* gene encoding a putative phloem protein 2 (PP2) of *Vitis vinifera* (XP_002279245) increased under drought conditions in Rubi but was quite stable in IAPAR59 under both conditions.

Expression of type II nsLTP genes

The expression of Type II nsLTP-encoding genes was also monitored using the primer pairs LTP-FT/LTP-R1 (specific to the *CaLTP1* and *CaLTP2* genes from the *C. eugenioides* sub-genome of *C. arabica*, hereafter referred to as *CaCe*), LTP-FT/LTP-R2 (specific to *CaLTP3* genes from the *C. canephora* of *C. arabica*, hereafter *CaCc*) and LTP-F100/LTP-R100 recognizing all homologous genes [32]. No expression of *nsLTP* genes was detected under the control conditions in both cultivars (Fig. 5). However, expression of *nsLTP* genes was highly up-regulated in IAPAR59 but not in Rubi under drought conditions. It is worth noting that the *CaLTP1-CaLTP2*

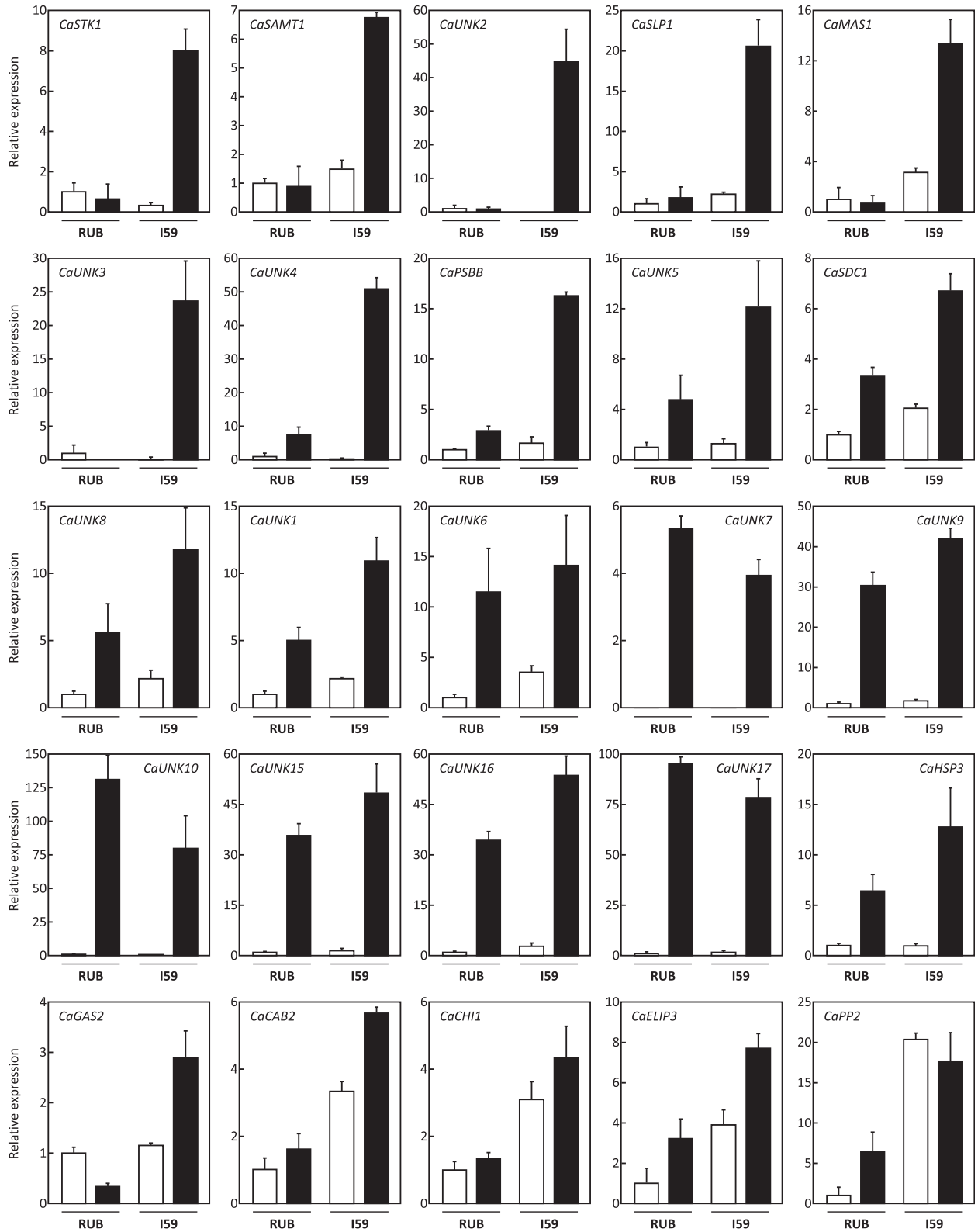


Fig. 4 Expression profiles of genes up-regulated under drought conditions. Gene expression was analysed in plagiotropic buds of Rubi (RUB) and IAPAR59 (I59) cultivars of *C. arabica* grown under control (white isobars) and drought (black isobars) conditions. The gene names are indicated in the histograms. Transcript abundances were normalized using the expression of the *CaUBQ10* gene as the endogenous control. Results are expressed using RUB-C as the reference sample (Relative expression = 1). Values of three technical replications are presented as mean ± SD (bar)

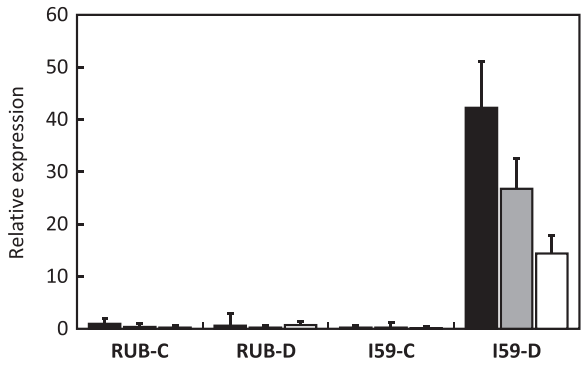


Fig. 5 Expression of *nsLTP* genes. Expression of *CaLTP1-CaLTP2* (*CaCe*: white isobars), *CaLTP3* (*CaCc*: grey isobars) and all (*CaLTP1*, *CaLTP2* and *CaLTP3*: black isobars) genes was analysed by qPCR in plagiotropic buds of Rubi (RUB) and IAPAR59 (I59) cultivars of *C. arabica* grown under control (C) and drought (D) conditions, using the LTP-FT/LTP-R2, LTP-FT/LTP-R1 and LTP-F100/LTP-R100 primer pairs, respectively [37]. Expression levels are expressed in arbitrary units (AU) of *nsLTP* genes using the expression of the *CaUBQ10* gene as the endogenous control and RUB-C (with LTP100 primers) as the reference sample (Relative expression = 1). Values of three technical replications are presented as mean ± SD (bar)

and *CaLTP3* genes were co-expressed in IAPAR59, and that the expression of *CaCc* genes was slightly higher than that of *CaCe* genes.

Drought influences leaf cuticle thickness

Leaf anatomical analyses were also performed, revealing that the abaxial epidermis of IAPAR59 had a thicker cuticle than Rubi under drought conditions (Fig. 6). There was also a strong interaction between genotype and drought conditions (F1, 40 = 16,2). For example, in the D^T cultivar IAPAR59, the abaxial epidermis cuticle thickness greatly increased under drought conditions compared with the control treatment (Table 4). However, no significant variation in abaxial epidermis cuticle thickness could be observed between the control and drought treatments for Rubi leaves.

Genes with reduced expression under drought conditions

The qPCR experiments led to the identification of several genes whose expression was reduced under drought conditions (Fig. 7). In both cultivars, expression of the orphan genes *CaUNK11* and *CaUNK12*, and of the *CaDLP1* gene encoding a putative protein containing a dirigent-like protein domain homologous to the hypothetical protein (CAN61316) of *Vitis vinifera*, was greatly reduced under drought conditions. Expression of the *CaCHI2* gene encoding a protein homologous to the putative chitinase of *Catharanthus roseus* (ADK98562), was 5-fold higher in IAPAR59 than in Rubi under the control conditions but decreased under drought conditions. However, the expression level of the *CaCHI2* gene was similar in IAPAR59 and Rubi under drought conditions. For the genes *CaCHI3* (putative protein related to chitinase-like protein *Artemisia annua* [ABJ74186]), *CaUNK13* and *CaJAMT1* (putative protein containing a methyltransferase domain [pfam03492] found in enzymes acting on salicylic acid, jasmonic acid and

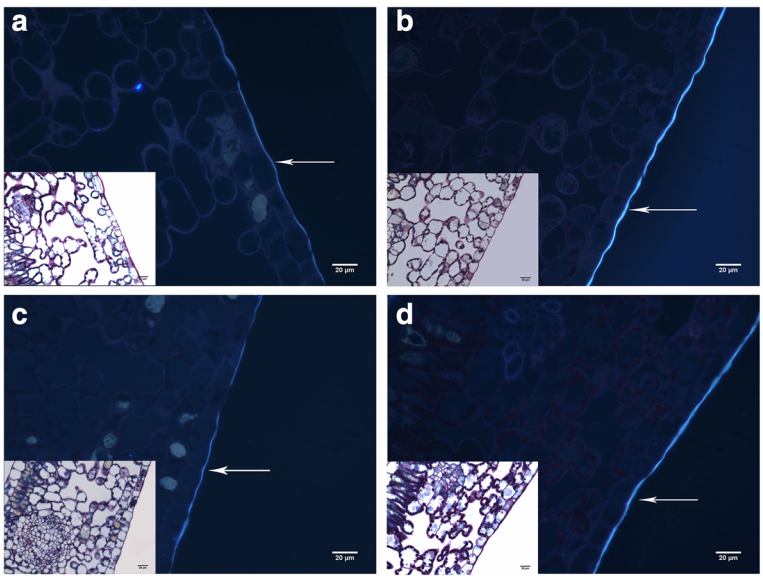


Fig. 6 Comparative analysis of leaf histological cross sections of IAPAR59 (**a** and **b**) and Rubi (**c** and **d**) cultivars of *C. arabica* under control (irrigation: **a** and **c**) and drought (**b** and **d**) conditions. Samples were double stained with Schiff and NBB and observed under wide field (at the bottom left of each image) and fluorescent microscopy (A4 filter). LE = Lower (abaxial) epidermis. The white arrows indicate the fluorescent cuticle. Values of leaf cuticle thickness are given in Table 4. Bars = 20 μm

Table 4 Influence of drought on leaf cuticle thickness

	Cuticle thickness (μm)	
Treatment	IAPAR59	Rubi
Control	$1.49 \pm 0.19^{(a)}$	$1.75 \pm 0.15^{(b)}$
Drought	$1.98 \pm 0.19^{(c)}$	$1.73 \pm 0.28^{(b)}$

Leaves of IAPAR59 and Rubi cultivars of *C. arabica* grown under control (irrigation) and drought conditions were analysed to measure the cuticle thickness of the abaxial faces. Values (in μm) correspond to the average calculated from 11 independent measurements. Those marked with different letters are significantly different (Student-Newman-Keuls mean comparison test, $P < 0.05$)

7-methylxanthine), similar expression profiles were found. In these cases, drought reduced gene expression in both cultivars but expression levels were always higher in IAPAR59 than in Rubi, particularly for *CaJAMT1*.

Gene expression levels of the *CaH2A* (H2A histone protein), *CaGRP2* (putative glycin-rich protein) and *CaUNK14* genes, were similar in Rubi and IAPAR59. For the *CaAEP1* (putative aldose 1-epimerase) and *CaIPS1* (myo-inositol 1-phosphate synthase) genes, gene expression remained high in IAPAR59 under both control and drought conditions, but decreased drastically in Rubi under drought conditions.

Discussion

In this study, we obtained 34.7 Mbp (coverage 6.5x) of sequences with longer reads (mean of 379.2 bp) from plagiotropic shoot apices enriched in meristems and primordium leaves of two cultivars of *C. arabica* under control (irrigation) and drought conditions. These sequences were assembled giving 43,087 clusters (17,719 contigs exclusively from 454-sequencing and 25,368 hybrid contigs formed by 454 and Sanger sequences) with a mean size ≥ 300 bp each. These RNAseq data, which complement those already available in public databases for coffee ESTs (407 million ESTs: dbEST release June 2015), can be considered as innovative and relevant in the sense that they were produced from *C. arabica* tissues (meristems) that have never previously been studied [39].

The transcriptome annotation by Blast2GO provided information based on the nomenclature and organism of

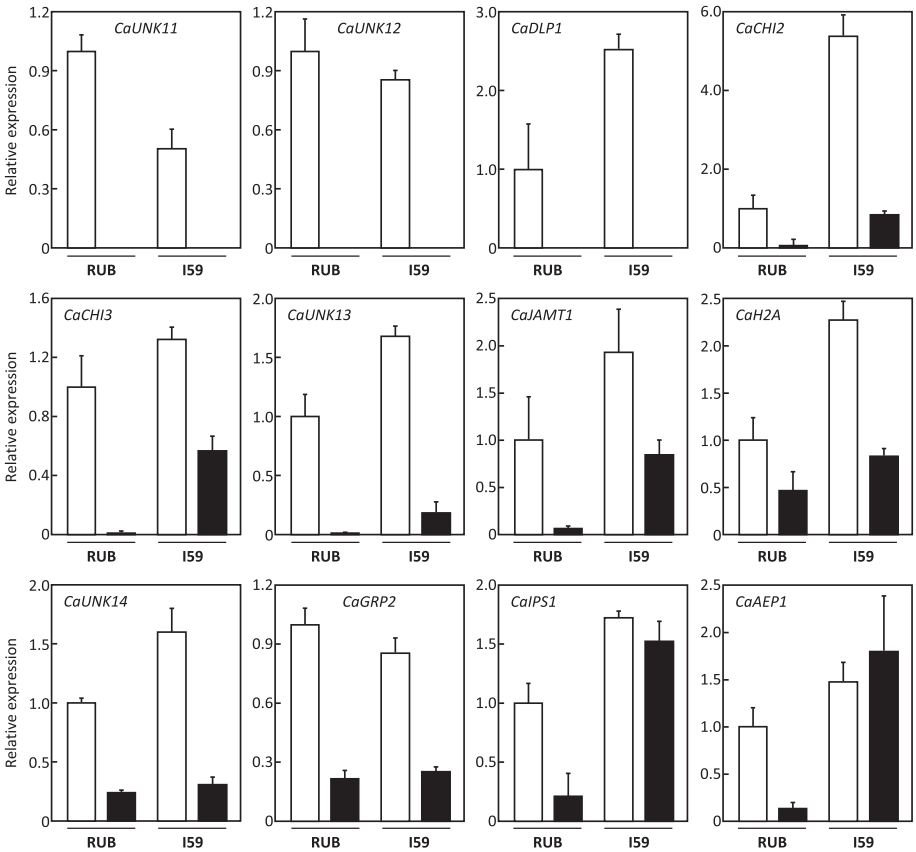


Fig. 7 Expression profiles of genes down-regulated under drought conditions. Gene expression was analysed in plagiotropic buds of Rubi (RUB) and IAPAR59 (I59) cultivars of *C. arabica* grown under control (white isobars) and drought (black isobars) conditions. The gene names are indicated in the histograms. Transcript abundances were normalized using the expression of the *CaUBQ10* gene as the endogenous control. Results are expressed using RUB-C as the reference sample (Relative expression = 1). Values of three technical replications are presented as mean ± SD (bar)

origin of genes in the NCBI/NR database, the enzyme family, a functional analysis of proteins from the InterPro database, and metabolic functions, biological processes and cellular location from gene ontology. Our results showed that a large percentage of transcriptome alignment had 36,965 hits with known function (85.8 %), 1,824 genes with unknown function (4.2 %) both in the NCBI/NR database, and only 1,515 hits in the Interpro database (3.5 %), thereby enabling the identification of most genes. With this analysis, we identified 34,857 genes related to *Coffea* sp. (80.9 % of the total). We also found 1,383 genes from *Solanum* sp., 573 genes from *Populus trichocarpa*, 482 genes from *Vitis vinifera* and 156 genes from *Arabidopsis* sp. Thus, the transcriptome was aligned with several genes from different plant species and these genes may be conserved among these species, including *Coffea* sp. On the other hand, our results also included 2,783 "no-hit" genes (6.5 %), perhaps indicating the presence of unannotated or new genes.

The comparisons of DNA libraries undertaken during this work led to the identification of 1,243 genes (Table 3: ∑ Total DEG %) with differential expression profiles *in silico* between the drought-susceptible (Rubi) and drought-tolerant (IAPAR59) cultivars of *C. arabica* with drought conditions. The expression profiles of these genes, as well as those of other previously identified genes [10, 11, 30–32], were analysed by qPCR in plagiotropic buds (containing meristems and small leaves) taken from control and drought-stressed plants of Rubi and IAPAR59. For most of the CGs identified during this work, in vivo gene expression profiles confirmed those deduced from *in silico* comparisons of DNA libraries. For example, this was the case for the *CaHSP3* (heat shock protein) gene whose up-regulated expression under drought conditions can be considered as a "molecular control" of stress applied to the plants during this study and confirmed by leaf water potential (Ψ_{pd}) measurements. Many ESTs encoding putative HSPs were also found in leaf cDNA libraries of *C. arabica* (SH2) and *C. canephora* (SH3) plants grown under drought conditions [31], heat stress [40], leaf infection by *Hemileia vastatrix* [15, 16] and also during bean development [14].

Our results also identified several genes differentially expressed in plagiotropic buds of IAPAR59 and Rubi, as for the *CaSTK1* gene encoding a putative serine/threonine protein kinase containing a conserved domain (cd06610) of mitogen-activated protein kinases (MAPKs). These kinases are known to have a central role in the transduction of extra- and intracellular signals in plants, including cell division and differentiation, as well as in responses to various types of stress [41]. In *Pisum sativum*, there is evidence that the MAPK cascade is involved in ABA-regulated stomatal activity as well as ABA-induced gene expression in the epidermal peels [42]. In a recent study, Shen et al. [43] showed that the phosphorylation of OsWRKY30 protein by MAPKs is a key step in conferring drought tolerance in transgenic rice. According to our results, higher *CaSTK1* expression under drought conditions in IAPAR59 than in Rubi could enhance the MAPK cascade and therefore be involved in the drought tolerance of IAPAR59. In this cultivar, the over-expression of *CaSAMT1* under drought conditions is also particularly interesting because this sequence encodes a putative S-adenosyl-L-methionine-dependent methyltransferase related to the TUMOROUS SHOOT DEVELOPMENT2 (TSD2) gene. In *Arabidopsis thaliana*, *tsd2* is a pleiotropic mutation that affects leaf, root and shoot meristem development [44]. Expression of a TSD2:: GUS reporter gene has mainly been detected in meristems where this gene is essential for cell adhesion and coordinated plant development. The weaker expression of *CaSAMT1* in Rubi than in IAPAR59 under drought conditions, points to the existence of major developmental differences between these two cultivars. The differential expression in Rubi and IAPAR59 of the *CaSLP1* gene encoding a putative subtilisin-like protein is also worth noting. In *Arabidopsis*, the subtilisin-like serine-protease *SDD1* (*stomatal density* and *distribution*) gene was shown to be strongly expressed in stomatal precursor cells (meristemoids and guard mother cells) [45]. In addition, *sdd1* mutation increased leaf stomatal density (SD) while *SDD1* over-expression led to the opposite phenotype with decreased SD. In *C. arabica*, maximum and minimum average stomatal densities were observed in full sunlight and shaded conditions respectively, providing evidence for the existence of plasticity for this characteristic in this coffee species [46, 47]. Even though no SD were observed between Rubi and IAPAR59 under moderate drought conditions [48], the *CaSLP1* expression profiles presented here do not preclude the involvement of this gene in the genetic determinism of drought tolerance in coffee.

Another interesting response concerned the differential expression of the *CaMAS1* gene encoding a putative protein containing the conserved domain [cd05326]. This domain is also found in secoisolariciresinol dehydrogenase-like proteins catalyzing the NAD-dependent conversion of (-)-secoisolariciresinol to (-)-matairesinol, like the *Arabidopsis* ABA2 protein considered to be one of the key regulators of ABA biosynthesis [49]. Based on the *CaMAS1* expression profiles presented here, it is possible that ABA synthesis was enhanced by drought in plagiotropic buds of IAPAR59 but not (or to a lesser extent) in those of Rubi. This hypothesis is also reinforced by the fact that higher *CaJAMT1* expression was observed in IAPAR59 than in Rubi buds. Indeed, in addition to well-known functions of jasmonates in plant defence mechanisms in response to biotic stress [50], recent studies also demonstrated that methyl jasmonate stimulates ABA

biosynthesis under drought conditions in panicles of *Oryza sativa* [51].

Higher expression of *CaSDC1* (encoding a protein sharing 89 % similarity with the S-adenosyl-L-methionine decarboxylase from *Catharanthus roseus*) under drought conditions in IAPAR59 than in Rubi is also worth noting because this enzyme catalyzes the synthesis of polyamines (e.g. spermine, spermidine and putrescine) involved in stress tolerance in higher plants [52]. In *Theobroma cacao*, ABA and drought induced the expression of *TcSAMDC* increasing spermine and spermidine leaf contents correlated with changes in stomatal conductance [53]. More recently, *SAMDC* over-expression in transgenic rice was also shown to facilitate drought tolerance [54]. Investigation of polyamine levels in plagiotropic buds and leaves of IAPAR59 and Rubi would be of particular interest to see if these compounds are involved in drought tolerance in coffee.

In mature plants, nuclear-encoded early-light inducible proteins (ELIPs) accumulate in response to various stress conditions including ABA or desiccation [55]. These proteins are presumed to protect the chloroplast apparatus from photo-oxidation occurring after stomatal limitation of photosynthesis [56]. In a recent study, transgenic plants of *Medicago truncatula* over-expressing the *Dsp22* gene from *Craterostigma plantagineum* were shown to be able to recover from water deprivation better than wild type plants, thereby reinforcing the idea of using ELIP-encoding genes to improve abiotic stress resistance in crops [57]. Our results clearly highlight the increased expression of the *CaELIP3* (ELIP-like), *CaPSBB* (CP47-like) and *CaCAB2* (PSII Cab proteins) genes, respectively, under drought conditions. Interestingly, the expression levels of all these genes were always higher in IAPAR59 than in Rubi. These results are also in accordance with electronic Northern experiments which showed high accumulation of ELIP and Cab-encoding ESTs in cDNA libraries of *C. arabica* and *C. canephora* subjected to drought [58].

Another surprising result concerned the *CaPSBB* gene that was reverse-transcribed and detected during our qPCR experiments despite the fact that it corresponds to a chloroplast gene [59]. However, preliminary analyses of a whole genome sequence of *C. canephora* revealed the presence of a CP47/like nuclear gene [60]. Interestingly, photosystem II CP47 chlorophyll apoproteins encoding ESTs have also been reported to be expressed in *C. arabica* beans [61], leaves infected by *Hemileia vastatrix* [62] and also in the cDNA libraries (SH2 and SH3) of drought-stressed coffee plants [14, 24, 31], demonstrating increased expression of this gene under biotic and abiotic stress. As CP47 and ELIP proteins are essential for the activity and protection of the photosynthetic apparatus [55], the expression profiles reported here probably reflect a better photosynthetic and physiological status of IAPAR59 compared to Rubi.

Differential expression was also observed for the chitinase-encoding gene *CaCHI1*, with higher expression in IAPAR59 than in Rubi. An opposite situation was observed with respect to the chitinase-encoding genes *CaCHI2* and *CaCHI3*, whose expression was reduced under drought conditions. It is worth noting that the expression of these genes under drought conditions was always higher in IAPAR59 than in Rubi. These results also show that coffee chitinase-encoding genes responded in different ways to drought. A large number of chitinase-encoding ESTs were identified in the BCGP project [24], mainly in the SH2 cDNA library of drought-stressed plants of *C. arabica* var. Catuai [58], but also in the leaves of *C. arabica* infected by leaf rust [62]. Even though chitinases are defence-related enzymes induced by abiotic stress, some evidence also indicates their participation in tolerance to abiotic stress [63]. Even though the roles of pathogenesis-related proteins in abiotic stress are still not fully understood, D^T transgenic plants over-expressing chitinase genes have been obtained [64]. In that sense, the high level of expression for *CaCHI1* in plagiotropic buds of IAPAR59 under both control and drought conditions could have an important function in drought tolerance.

Arbutin is a phenolic glucoside (4-hydroxyphenyl-β-D-glucopyranoside) abundant in the leaves of many freezing- or desiccation-tolerant plants [65] and also present in coffee fruits [66]. In a previous study, down-regulation of the *CcGAS1*gene encoding arbutin synthase was reported in leaves of *C. canephora* under drought conditions [10]. The results presented here clearly demonstrated differential expression profiles for *CaGAS2* between the two cultivars of *C. arabica*. Gene expression increased under drought conditions in IAPAR59 while the opposite was observed in Rubi. Even though the presence of arbutin in coffee leaves has never been demonstrated, further analyses of this metabolite should be performed to investigate the role of this glucoside (and of other phenolic compounds) in preventing cell damage in coffee subject to abiotic stresses.

The *CaPP2* gene (encoding a putative phloem protein 2, PP2) also showed differential expression profiles, with higher expression in IAPAR59 than in Rubi. In higher plants, PP2s are sieve elements (SE) very abundant in the phloem sap. These proteins are believed to play an important role in the establishment of phloem-based defence mechanisms induced by insect attacks and feeding stress [67], but also by wounding and oxidative conditions [68]. The functions of PP2 proteins are still not clear but they could act by forming high molecular weight polymers to close ("SE plugging") the sieve pores caused by external injuries mainly due to biotic stress [69]. When *Arabidopsis* was treated with $HrpN_{Ea}$ (a

proteinaceous elicitor of plant defences produced by gram-negative plant pathogenic bacteria), the suppression of phloem-feeding activities by aphids was attributed to over-expression of the PP2-encoding gene *AtPP2-A1* [70]. Other studies showed that HrpN activated ABA signalling, thereby inducing drought tolerance in *Arabidopsis thaliana* [71]. Based on these results, the involvement of PP2 proteins in plant response mechanisms to abiotic stress can be hypothesized, for example by maintaining (or protecting) the integrity of vessels under drought conditions by forming sieve plate filaments upon oxidation [72]. In that case, higher synthesis of CaPP2 which would be expected to occur in IAPAR59 plagiotropic buds under drought conditions could play a role in drought-tolerance by reducing sap-flow in young leaves and consequently increasing the water use efficiency of this cultivar [48].

Other interesting results concerned the gene expression stability of the *CaAEP1* (putative aldose 1-epimerase) and *CaIPS1* (myo-inositol 1-phosphate synthase) genes observed in IAPAR59 under control and drought conditions, whereas expression of both genes decreased under drought conditions in Rubi. Plant cells use myo-inositol to synthesize a variety of low molecular weight compounds and sugar alcohols such as the galactinol, a key element in the formation of raffinose family oligosaccharides. Nishizawa et al. [73] found that plants with high galactinol and raffinose contents were less susceptible to oxidative stress. In *C. arabica*, up-regulation of *CaGolS* genes involved in galactinol biosynthesis was reported in leaves of plants subjected to severe drought [74]. In addition, drought up-regulated the expression of mannose 6-phosphate reductase (involved in mannitol biosynthesis) in leaves of *C. canephora* [10, 11] and *C. arabica* [75, 76]. Even though little is known about the biochemical mechanisms of drought tolerance in coffee, the accumulation of carbohydrates expected in leaves of drought-stressed plants as a consequence of the up-regulated expression of these genes, could play an important role in the genetic determinism of this phenotype in coffee [77].

In addition to the previously described genes, our results also identified several orphan genes that presented differential expression profiles between the cultivars and treatments, such as *CaUNK2*, *CaUNK3* and *CaUNK4* whose expression was highly induced under drought conditions in IAPAR59 and to a lesser extent in Rubi. Orphan genes are also expected to interact specifically with the environment as a consequence of lineage-specific adaptations to that environment [78].

Interestingly, the expression profiles of the *CaUNK2* and *CaUNK3* orphan genes were very similar to those of Type II nsLTP-encoding genes, with high expression mainly detected under drought conditions in plagiotropic buds of IAPAR59 but not in those of Rubi. Up-regulation of *LTP* genes under drought conditions is well documented in higher plants [79–81]. Lipid transfer proteins (LTPs) are thought to be involved in the transfer of lipids through the extracellular matrix for the formation of cuticular wax [82]. In fact, together with the lipophilic cutin polymer matrix, waxes enter in the composition of cuticle, which forms the first barrier between plants and environmental stresses by limiting non-stomatal water loss and gas exchanges, hence mitigating the effects of drought by controlling water loss associated with epidermal conductance [83]. In *Nicotiana glauca,* LTP genes are predominantly expressed in the guard and epidermal cells and are induced under drought conditions [84], providing evidence that LTP play an important role in the development of drought tolerance. Even though the up-regulation of *CaLTP* genes observed under drought in plagiotropic buds of IAPAR59 cannot explain directly the greater thickness of leaf cuticle observed in this cultivar than in Rubi, these results strongly suggested that lipid metabolism plays a major role in coffee drought tolerance.

As reported in other higher plants, our study also highlighted the differential expression of many genes encoding proteins known to be over-expressed under biotic stress (e.g. chitinases and PP2), by drought. The fact that our experiment was conducted with drought-stressed plants grown under uncontrolled (field) conditions, could explain such a situation. However, it is also probable that these results reflect a biological reality since it is well known that crosstalk exists in higher plants between signalling pathways for biotic and abiotic stress responses [85].

Conclusions

During this work, we produced some new transcriptomic information for *C. arabica* with a total of 34.7 Mbp of sequences assembled into 43,087 clusters (41,512 contigs and 1,575 singletons) from genes expressed in plagiotropic shoot apices enriched in meristems and primordium leaves in D^T (IAPAR59) and D^S (Rubi) cultivars grown under control and drought conditions. Major differences between these plants concerned their phenotypic behaviour (e.g. predawn leaf water potential, Ψ_{pd}) and transcriptome expression profiles. Differences between these plants affected genes of specific pathways such as those involved in abscisic acid biosynthesis, perception and transduction of drought stress, plant development and lipid metabolism. In that sense, the present study increased the number of CGs potentially involved in the genetic determinism of drought tolerance firstly identified in *C. canephora*. Because *C. arabica* is an amphidiploid species (originating from a natural hybridization event between *C. canephora* and *C. eugenioides*), its transcriptome is a mixture of homologous genes expressed from these two sub-

genomes in which *C. eugenioides* is assumed to express genes mainly for proteins involved in basal biological processes (e.g. photosynthesis), while the *C. canephora* sub-genome is assumed to regulate Arabica gene expression by expressing genes for regulatory proteins and adaptation processes [86]. In this genetic context, it is possible that the characteristics of IAPAR59 that enable it to better withstand drought stress than Rubi, really originated from the specific expression of *C. canephora* genes recently introgressed (through the Timor hybrid HT832/2 [19]) in this cultivar of *C. arabica* [33]. Even though this study provides further indications about the way in which different coffee cultivars activate their transcriptomes, additional work is still required to understand how epigenetics and epistasis regulate gene expression in the different coffee sub-genomes (*CaCe* and *CaCc*) in *C. arabica* under drought conditions.

Source of the plant materials and permissions

This work was carried out as part of the scientific cooperation project entitled "Study of genetic determinism of drought tolerance in coffee" (2006–2010) approved between Embrapa and CIRAD. It complied with all institutional, national, or international guidelines. In the frame of this project, field experiments were conducted at the Cerrado Agricultural Research Center (Planaltina-DF, Brazil) with all permissions of partners and in accordance with local legislation.

Ethics approval and consent to participate

Not applicable.

Consent to publish

Not applicable.

Availability of supporting data

The reads were submitted to GenBank and to the BioProject/NCBI database under the accession number PRJNA282394.

Additional files

Additional file 1: Table S1. Summary of Blast2GO automatic annotation of the transcriptome clusters. (XLS 15226 kb)

Additional file 2: Figure S1. Complete bioinformatics pipeline of the transcriptome assembly and automatic annotation methods used in this work. (TIF 105 kb)

Additional file 3: Table S2. Summary of the DEseq/ EdgeR fold changes and *p*-values between the cultivars (I59: IAPAR59 and RUB: Rubi) and between control (C) and drought (D) conditions. These tables also contain Blast2GO automatic annotation of the transcriptome clusters. Table lanes coloured in grey are related to clusters aligned to the new candidate genes tested by RT-qPCR (see Table 1). (XLS 51895 kb)

Abbreviations

EST: expressed sequence tag; qPCR: quantitative polymerase chain reaction.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

GCR and PM measured predawn leaf water potentials and harvested the samples. FdB and TL performed meristem dissections, RNA extractions and cDNA synthesis. LSM, MFC, GAGP and RV were responsible for the bioinformatic processing of the data. JLV, FLM and ML performed the histology and microscopy analyses. PM, ACA and DP selected the candidate genes qPCR-analysed by FAC, NGV, KED, JCA and MGC. GCR, ACA and PM designed the study, drew up the experimental design and implemented it. RV, MFC, GAGP, ACA and PM drafted the manuscript. All the authors read and approved the final version of the manuscript.

Acknowledgments

The authors would like to thank Dr Aymbiré Francisco Almeida da Fonseca from the INCAPER Institute for kindly providing seeds of the Rubi and IAPAR59 commercial cultivars of *C. arabica*. The authors are also very grateful to Drs Antonio Fernando Guerra and Omar Cruz Rocha (Embrapa Cerrados) for their assistance during the field trial experiments. The authors wish also to thank Daphne Goodfellow, as well as Peter Biggins and Cécile Fovet-Rabot (CIRAD-DGDRS) for English revision of the manuscript.

Funding

PM acknowledges financial support from CIRAD (Centre de Coopération Internationale en Recherche Agronomique pour le Développement, Montpellier, France) and the CIRAD ATP project "Analysis of phenotypic plasticity in response to water constraints in perennial plants growing under different field conditions". ACA acknowledges financial support from the Brazilian Coffee R&D Consortium, FINEP and INCT-café (CNPq/FAPEMIG). MFC and GAGP acknowledge financial support from the Center for Computational Engineering and Sciences - FAPESP/Cepid (2013/08293-7). The authors acknowledge the scholarships from the Brazilian agencies CAPES (KED, NGV), Brazilian Consortium of Coffee Research (FAC, JCA) and CAPES-COFECUB Project Sv738-12 (MGC).

Author details

[1]Laboratório de Genômica e Expressão (LGE), Departamento de Genética e Evolução, Instituto de Biologia/UNICAMP, Cidade Universitária Zeferino Vaz, 13083-970 Campinas, SP, Brazil. [2]Embrapa Recursos Genéticos e Biotecnologia (LGM-NTBio), Parque Estação Biológica, CP 02372, 70770-917, Brasilia, DF, Brazil. [3]CIRAD UMR AGAP, F-34398 Montpellier, France. [4]Embrapa Informática Agropecuária, UNICAMP, Av. André Tosello n° 209, CP 6041, 13083-886 Campinas, SP, Brazil. [5]present address: Embrapa Café, INOVACAFÉ, Campus UFLA, 37200-000 Lavras, MG, Brazil.

Received: 14 October 2015 Accepted: 13 April 2016
Published online: 19 April 2016

References

1. Pay E. The market for organic and fair-trade coffee. FAO Rome. 2009.
2. ICO (International Coffee Organization), Statistics of production 2015: http://www.ico.org/prices/po-production.pdf.
3. DaMatta FM, Ramalho JC. Impact of drought and temperature stress on coffee physiology and production: a review. Braz J Plant Physiol. 2006;18: 55–81.
4. Silva EA, Mazzafera P, Brunini O, Sakai E, Arruda FB, Mattoso LHC, et al. The influence of water management and environmental conditions on the chemical composition and beverage quality of coffee beans. Braz J Plant Physiol. 2005;17:229–38.
5. Bunn C, Läderach P, Pérez Jimenez JG, Montagnon C, Schilling T. Multiclass classification of agro-ecological zones for Arabica coffee: an improved understanding of the impacts of climate change. PLoS One. 2015;10, e0140490.
6. Montagnon C, Leroy T. Réaction à la sécheresse de jeunes caféiers *Coffea canephora* de Côte-d'Ivoire appartenant à différents groupes génétiques. Café Cacao Thé. 1993;37:179–90.
7. Pinheiro HA, DaMatta FM, Chaves ARM, Loureiro ME, Ducatti C. Drought tolerance is associated with rooting depth and stomatal control of water use in clones of *Coffea canephora*. Ann Bot. 2005;96:101–8.
8. Beining A, Burkhardt J, Fetene M. Water relations of Ethiopian wild coffee populations: genetic fixation and phenotypic plasticity. In Proceedings of

the 22nd International Scientific Colloquium on Coffee, Campinas, International Scientific Association on Coffee, Paris, CDROM-A105, 2008.
9. Shinozaki K, Yamaguchi-Shinozaki K. Gene networks involved in drought stress response and tolerance. J Exp Bot. 2007;58:221–7.
10. Marraccini P, Vinecky F, Alves GSC, Ramos HJO, Elbelt S, Vieira NG, et al. Differentially expressed genes and proteins upon drought acclimation in tolerant and sensitive genotypes of *Coffea canephora*. J Exp Bot. 2012;63:4191–212.
11. Vieira NG, Carneiro FA, Sujii PS, Alekcevetch JC, Freire LP, Vinecky F, et al. Different molecular mechanisms account for drought tolerance in *Coffea canephora* var. Conilon. Trop Plant Biol. 2013;6:181–90.
12. Lin C, Mueller LA, McCarthy J, Crouzillat D, Pétiard V, Tanksley SD. Coffee and tomato share common gene repertoires as revealed by deep sequencing of seed and cherry transcripts. Theor Applied Genet. 2005;112:114–30.
13. Poncet V, Rondeau M, Tranchand C, Cayrel A, Hamon S, de Kochko A, et al. SSR mining in coffee tree EST databases: potential use of EST-SSRs as marker across *Coffea* genus. Mol Gen Genet. 2006;276:436–49.
14. Vieira LGE, Andrade AC, Colombo CA, Moraes AAH, Metha A, Oliveira AC, et al. Brazilian coffee genome project: an EST-based genomic resource. Braz J Plant Physiol. 2006;18:95–108.
15. Guzzo SD, Harakava R, Tsai SM. Identification of coffee genes expressed during systemic acquired resistance and incompatible interaction with *Hemileia vastatrix*. J Phytopathol. 2009;157:625–38.
16. Fernandez D, Santos P, Agostini C, Bon MC, Petitot AS, Silva CM, et al. Coffee (*Coffea arabica* L.) genes early expressed during infection by the rust fungus (*Hemileia vastatrix*). Mol Plant Pathol. 2004;5:527–36.
17. Traas J, Vernoux T. The shoot apical meristem: the dynamics of a stable structure. Philos T Roy Soc B. 2002;357:737–47.
18. Torres TT, Metta M, Ottenwalder B, Schlotterer C. Gene expression profiling by massively parallel sequencing. Genome Res. 2008;18:172–7.
19. Carvalho CHS, Fazuoli LC, Carvalho GR, Guerreiro-Filho O, Pereira AA, de Almeida SR, et al. Cultivares de café arábica de porte baixo. In: Carvalho CHS, editor. Cultivares de Café: origem, características e recomendações. Brasilia: Embrapa Café; 2008. p. 157–226.
20. Ratter JA, Ribeiro JF, Bridgewater S. The Brazilian cerrado vegetation and threats to its biodiversity. Ann Bot. 1997;80:223–30.
21. Lécolier A, Noirot M, Escoute J, Chrestin H, Verdeil JL. Early effects of the mutation laurina on the functioning and size of the shoot apex in coffee tree and analysis of the plastochron phases: relationships with the dwarfism of leaves. Trees. 2009;23:1043–51.
22. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, et al. Genome sequencing in microfabricated high-density picolitre reactors. Nature. 2005;437:376–80.
23. Ning Z, Cox AJ, Mullikin JC. SSAHA: a fast search method for large DNA databases. Genome Res. 2001;11:1725–9.
24. Mondego JMC, Vidal RO, Carazzolle MF, Tokuda EK, Parizzi LP, Costa GGL, et al. An EST-based analysis identifies new genes and reveals distinctive gene expression features of *Coffea arabica* and *Coffea canephora*. BMC Plant Biol. 2011;11:30.
25. Baudet C, Dias Z. New EST trimming strategy in: Brazilian Symposium on Bioinformatics, 2005. Lect Notes Bioinf. 2005;3594:206–9.
26. Chevreux B, Pfisterer T, Drescher B, Driesel AJ, Müller WEG, Wetter T, et al. Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs. Genome Res. 2004;14:1147–59.
27. Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics. 2005;21:3674–6.
28. Anders S, Huber W. Differential expression analysis for sequence count data. Genome Biol. 2010;11:R106.
29. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26:139–40.
30. Vieira NG. Identificação e caracterização de genes órfãos ("*NO HITS*") de café (*Coffea* spp.). Master Thesis, University of Lavras (UFLA), Lavras-MG, Brazil. 2013. 130p. http://repositorio.ufla.br/handle/1/4928.
31. Vinecky F, da Silva FR, Andrade AC. Análise *in silico* das bibliotecas de cDNA SH2 e SH3 para a identificação de genes responsivos à seca em cafeeiro. Coffee Sci. 2012;7:1–19.
32. Cotta MG, Barros LMG, de Almeida JD, de Lamotte F, Barbosa EA, Vieira NG, et al. Lipid transfer proteins in coffee: isolation of *Coffea* orthologs, *Coffea arabica* homeologs, expression during coffee fruit development and promoter analysis in transgenic tobacco plants. Plant Mol Biol. 2014;85:11–31.
33. Marraccini P, Freire LP, Alves GSC, Vieira NG, Vinecky F, Elbelt S, et al. *RBCS1* expression in coffee: *Coffea* orthologs, *Coffea arabica* homeologs, and expression variability between genotypes and under drought stress. BMC Plant Biol. 2011;11:85.
34. Ramakers C, Ruijter JM, Deprez RH, Moorman AF. Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. Neurosci Lett. 2003;339:62–6.
35. Schwendiman J, Pannetier C, Michaux-Ferriere N. Histology of somatic embryogenesis from leaf explants of the oil palm *Elaeis guineensis*. Ann Bot. 1988;62:43–52.
36. Buffard-Morel J, Verdeil JL, Pannetier C. Embryogenèse somatique du cocotier (*Cocos nucifera* L.) à partir d'explant foliaire: étude histologique. Can J Bot. 1992; 70:735–41.
37. Fisher DB. Protein staining of ribonned epon section for light microscopy. Histochemie. 1968;16:92–6.
38. Sokal RR, Rohlf JF. Biometry: the principles and practice of statistics in biological research. 3rd ed. New York: WH Freeman and Company; 1995. 887 pp.
39. Lashermes P, Andrade AC, Etienne H. Genomics of coffee, one of the world's largest traded commodities. In: Moore H, Ming R, editors. Genomics of tropical crop plants. Berlin: Springer; 2008. p. 203–26.
40. Bardil A, de Almeida JD, Combes MC, Lashermes P, Bertrand B. Genomic expression dominance in the natural allopolyploid *Coffea arabica* is massively affected by growth temperature. New Phytologist. 2011;192:760–74.
41. Mishra NS, Tuteja R, Tuteja N. Signaling through MAP kinase networks in plants. Arch Biochem Biophys. 2006;452:55–68.
42. Burnett EC, Desikan R, Moser RC, Neill SJ. ABA activation of an MBP kinase in *Pisum sativum* epidermal peels correlates with stomatal responses to ABA. J Exp Bot. 2000;51:197–205.
43. Shen H, Liu C, Zhang Y, Meng X, Zhou X, Chu C, et al. OsWRKY30 is activated by MAP kinases to confer drought tolerance in rice. Plant Mol Biol. 2012;80:241–53.
44. Krupková E, Immerzeel P, Pauly M, Schmülling T. The *TUMOROUS SHOOT DEVELOPMENT2* gene of *Arabidopsis* encoding a putative methyltransferase is required for cell adhesion and co-ordinated plant development. Plant J. 2007;50:735–50.
45. von Groll U, Berger D, Altmann T. The subtilisin-like serine protease SDD1 mediates cell-to-cell signaling during *Arabidopsis* stomatal development. Plant Cell. 2002;14:1527–39.
46. Pompelli MF, Martins SCV, Celin EF, Ventrella MC, DaMatta FM. What is the influence of ordinary epidermal cells and stomata on the leaf plasticity of coffee plants grown under full-sun and shady conditions? Braz J Biol. 2010; 70:1083–8.
47. Kufa T, Burkhardt J. Stomatal characteristics in Arabica coffee germplasm accessions under contrasting environments at Jimma, southwestern Ethiopia. Intern J Bot. 2011;7:63–72.
48. Rodrigues GC, Rojas JSD, Roupsard O, Leroy T, Pot D, Moreira MZ, et al. Preliminary results on phenotypic plasticity of coffee (Coffea arabica cv. Rubi and lapar59) plants in response to water constraint under field conditions. In Proceedings of the 23rd International Scientific Colloquium on Coffee, Bali, International Scientific Association on Coffee, Paris, CDROM-PB731, 2010.
49. Schwartz SH, Léon-Kloosterziel KM, Koornneef M, Zeevaart JAD. Biochemical characterization of the *aba2* and *aba3* mutants in *Arabidopsis thaliana*. Plant Physiol. 1997;114:161–6.
50. Santino A, Taurino M, De Domenico S, Bonsegna S, Poltronieri P, Pastor V, et al. Jasmonate signaling in plant development and defense response to multiple (a) biotic stresses. Plant Cell Rep. 2013;32:1085–98.
51. Kim EH, Kim YS, Park SH, Koo YJ, Choi YD, Chung YY, et al. Methyl jasmonate reduces grain yield by mediating stress signals to alter spikelet development in rice. Plant Physiol. 2009;149:1751–60.
52. Alcázar R, Altabella T, Marco F, Bortolotti C, Reymond M, Koncz C, et al. Polyamines: molecules with regulatory functions in plant abiotic stress tolerance. Planta. 2010;231:1237–49.
53. Bae H, Kim SH, Kim MS, Sicher RC, Lary D, Strem MD, et al. The drought response of *Theobroma cacao* (cacao) and the regulation of genes involved in polyamine biosynthesis by drought and other stresses. Plant Physiol Biochem. 2008;46:174–88.
54. Peremarti A, Bassie L, Christou P, Capell T. Spermine facilitates recovery from drought but does not confer drought tolerance in transgenic rice plants expressing *Datura stramonium S*-adenosylmethionine decarboxylase. Plant Mol Biol. 2009;70:253–64.

55. Duque AS, Almeida AM, Silva AB, Silva JM, Farinha AP, Santos D, et al. Abiotic stress responses in plants: unraveling the complexity of genes and networks to survive. In: Vahdati K, Leslie C, editors. Abiotic Stress - Plant responses and applications in agriculture. Rijeka (Croatia): InTech Publisher; 2013. p. 49–101.
56. Hutin C, Nussaume L, Moise N, Moya I, Kloppstech K, Havaux M. Early light induced proteins protect *Arabidopsis* from photoxidative stress. Proc Natl Acad Sci U S A. 2003;100:4921–6.
57. Araújo SS, Duque AS, Silva JM, Santos D, Silva AB, Fevereiro P. Water deficit and recovery response of *Medicago truncatula* plants expressing the ELIP-like DSP22. Biol Plantarum. 2013;57:159–63.
58. Alvarenga SM, Caixeta ET, Hufnagel B, Thiebaut F, Zambolim EM, Zambolim L, et al. *In silico* identification of coffee genome expressed sequences potentially associated with resistance to diseases. Gen Mol Biol. 2010;33:795–806.
59. Samson N, Bausher MG, Lee SB, Jansen RK, Daniell H. The complete nucleotide sequence of the coffee (*Coffea arabica* L.) chloroplast genome: organization and implications for biotechnology and phylogenetic relationships amongst angiosperms. Plant Biotechnol J. 2007;5:339–53.
60. Denoeud F, Carretero-Paulet L, Dereeper A, Droc G, Guyot R, Pietrella M, et al. The coffee genome provides insight into the convergent evolution of caffeine biosynthesis. Science. 2014;345:1181–4.
61. Salmona J, Dussert S, Descroix F, de Kochko A, Bertrand B, Joët T. Deciphering transcriptional networks that govern *Coffea arabica* seed development using combined cDNA array and real-time RT-PCR approaches. Plant Mol Biol. 2008; 66:105–24.
62. Fernandez D, Tisserant E, Talhinhas P, Azinheira H, Vieira A, Petitot AS, et al. 454-pyrosequencing of *Coffea arabica* leaves infected by the rust fungus *Hemileia vastatrix* reveals in planta-expressed pathogen-secreted proteins and plant functions in a late compatible plant-rust interaction. Mol Plant Pathol. 2012;13:17–37.
63. Yu LX, Djebrouni M, Chamberland H, Lafontaine JG, Tabaeizadeh Z. Chitinase: differential induction of gene expression and enzyme activity by drought stress in the wild (*Lycopersicon chilense* Dun.) and cultivated (*L. esculentum* Mill.) tomatoes. J Plant Physiol. 1998;153:745–53.
64. Dana MM, Pintor-Toro JA, Cubero B. Transgenic tobacco plants overexpressing chitinases of fungal origin show enhanced resistance to biotic and abiotic stress agents. Plant Physiol. 2006;142:722–30.
65. Hincha DK, Oliver AE, Crowe JH. Lipid composition determines the effects of arbutin on the stability of membranes. Biophys J. 1999;77:2024–34.
66. Iwai K, Kishimoto N, Kakino Y, Mochida Y, Fujita T. In vitro antioxidative effects and tyrosinase inhibitory activities of seven hydroxycinnamoyl derivatives in green coffee beans. J Agric Food Chem. 2004;52:4893–8.
67. Will T, van Bel AJE. Physical and chemical interactions between aphids and plants. J Exp Bot. 2006;57:729–37.
68. Tjallingii WF. Salivary secretions by aphids interacting with proteins of phloem wound responses. J Exp Bot. 2006;57:739–45.
69. Kehr J. Phloem sap proteins: their identities and potential roles in the interaction between plants and phloem-feeding insects. J Exp Bot. 2006;57:767–74.
70. Zhang C, Shi H, Chen L, Wang X, Lü B, Zhang S, et al. Harpin-induced expression and transgenic overexpression of the phloem protein gene *AtPP2-A1* in *Arabidopsis* repress phloem feeding of the green peach aphid *Myzus persicae*. BMC Plant Biol. 2011;11:11.
71. Zhang C, Bao Z, Liang Y, Yang X, Wu X, Hong X, et al. Abscisic acid mediates *Arabidopsis* drought tolerance induced by $HrpN_{Ea}$ in the absence of ethylene signaling. Plant Mol Biol Rep. 2007;25:98–114.
72. Walz C, Juenger M, Schad M, Kehr J. Evidence for the presence and activity of a complete antioxidant defence system in mature sieve tubes. Plant J. 2002;31:189–97.
73. Nishizawa A, Yabuta Y, Shigeoka S. Galactinol and raffinose constitute a novel function to protect plants from oxidative damage. Plant Physiol. 2008;147:1251–63.
74. dos Santos TB, Budzinski IGF, Marur CJ, Petkowicz CLO, Pereira LFP, Vieira LGE. Expression of three galactinol synthase isoforms in *Coffea arabica* L. and accumulation of raffinose and stachyose in response to abiotic stresses. Plant Physiol Biochem. 2011;49:441–8.
75. Freire LP, Marraccini P, Rodrigues GC, Andrade AC. Análise da expressão do gene da manose 6 fosfato redutase em cafeeiros submetidos ao déficit hídrico no campo. Coffee Sci. 2013;8:17–23.
76. de Carvalho K, Petkowicz CLO, Nagashima GT, Bespalhok Filho JC, Vieira LGE, Pereira LFP, et al. Homeologous genes involved in mannitol synthesis reveal unequal contributions in response to abiotic stress in *Coffea arabica*. Mol Genet Genomics. 2014;289:951–63.
77. DaMatta FM, Chaves ARM, Pinheiro HA, Ducatti C, Loureiro ME. Drought tolerance of two field-grown clones of *Coffea canephora*. Plant Sci. 2003;164:111–7.
78. Tautz D, Domazet-Lošo T. The evolutionary origin of orphan genes. Nat Rev Genet. 2011;12:692–702.
79. Treviño MB, O'Connell MA. Three drought-responsive members of the non-specific lipid-transfer protein gene family in *Lycopersicon pennellii* show different patterns of expression. Plant Physiol. 1998;116:1461–8.
80. Jung HW, Kim W, Hwang BK. Three pathogen-inducible genes encoding lipid transfer protein from pepper are differentially activated by pathogens, abiotic, and environmental stresses. Plant Cell Environ. 2003;26:915–28.
81. Jang CS, Lee HJ, Chang SJ, Seo YW. Expression and promoter analysis of the *TaLTP1* gene induced by drought and salt stress in wheat (*Triticum aestivum* L.). Plant Sci. 2004;167:995–1001.
82. Cameron KD, Teece MA, Smart LB. Increased accumulation of cuticular wax and expression of lipid transfer protein in response to periodic drying events in leaves of tree tobacco. Plant Physiol. 2006;140:176–83.
83. Schweizer P, Felix G, Buchala A, Muller C, Metraux JP. Perception of free cutin monomers by plant cells. Plant J. 1996;10:331–41.
84. Cameron KD, Moskal WA, Smart LB. A second member of the *Nicotiana glauca* lipid transfer protein gene family, *NgLTP2*, encodes a divergent and differentially expressed protein. Funct Plant Biol. 2006;33:141–52.
85. Fujita M, Fujita Y, Noutoshi Y, Takahashi F, Narusaka Y, Yamaguchi-Shinozaki K, et al. Crosstalk between abiotic and biotic stress responses: a current view from the points of convergence in the stress signaling networks. Curr Opin Plant Biol. 2006;9:436–42.
86. Vidal RO, Mondego JMC, Pot D, Ambrósio AB, Andrade AC, Pereira LFP, et al. A high-throughput data mining of single nucleotide polymorphisms in *Coffea* species expressed sequence tags suggests differential homeologous gene expression in the allotetraploid *Coffea arabica*. Plant Physiol. 2010; 154:1053–66.